

У. ЧЕРЧМЕН, Р. АКОФ, Л. АРНОФ

ВВЕДЕНИЕ В ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

ПЕРЕВОД С АНГЛИЙСКОГО

В. Я. АЛТАЕВА, Ю. А. КРУТИКОВА, А. И. ТЕЙМАНА

ПОД РЕДАКЦИЕЙ

А. Я. ЛЕРНЕРА

ИЗДАТЕЛЬСТВО «НАУКА»

ГЛАВНАЯ РЕДАКЦИЯ

ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ

МОСКВА 1968

INTRODUCTION TO OPERATIONS RESEARCH

C. WEST CHURCHMAN
Professor

RUSSELL L. ACKOFF
*Director Operations Research
Group and Professor*

E. LEONARD ARNOFF
*Assistant Director Operations
Research Group and Associate Professor*

Введение в исследование операций. Черчмен У.,
Акоф Р., Арноф Л. Перевод с английского.
Издательство «Наука», Главная редакция физико-мате-
матической литературы, М., 1967, 488 стр.

Исследование операций — бурно развивающаяся область науки, основанная на применении современных разделов математики и тесно связанная с кибернетикой, теорией автоматического управления, экономикой и рядом других наук. Целью исследования операций является получение оптимальных решений задач, возникающих в процессе функционирования или создания сложных систем, включающих людей и машины.

В книге, являющейся общепризнанным классическим трудом по исследованию операций, рассмотрены методологические основы, типичные математические модели и принципы практической реализации результатов операционных исследований. Не привлекая излишне сложный математический аппарат, однако и не постулаясь строгостью и глубиной изложения, авторы дают хорошее представление о круге операционных задач и методах их решения, уделяя особое внимание вопросам постановки задач.

Книга рассчитана на математиков, кибернетиков, экономистов, инженеров, а также руководителей, работающих в различных отраслях народного хозяйства.

Табл. 154. Илл. 123. Библ. 440 назв.

О Г Л А В Л Е Н И Е

От редактора русского перевода	7
Предисловие к английскому изданию	

Ч А С Т Ь I

ВВЕДЕНИЕ

Глава 1. Общая характеристика исследования операций	12
Глава 2. Операционное исследование системы как единого целого	26
Глава 3. Комплексный подход к решению задачи контроля качества	58

Ч А С Т Ь II

ЗАДАЧА

Глава 4. Анализ организации	67
Глава 5. Постановка задачи	94
Глава 6. Определение весовых коэффициентов целей	117

Ч А С Т Ь III

МОДЕЛЬ

Глава 7. Построение и решение модели	134
--	-----

Ч А С Т Ь IV

МОДЕЛИ УПРАВЛЕНИЯ ЗАПАСАМИ

Глава 8. Простейшие модели управления запасами	163
Глава 9. Модели управления запасами при нескольких уровнях цен	189
Глава 10. Модели управления запасами при наличии ограничений	199

Ч А С Т Ь V

МОДЕЛИ РАСПРЕДЕЛЕНИЯ

Глава 11. Линейное программирование	215
Глава 12. Задача о назначении	263
Глава 13. Некоторые примеры задач распределения	282

ОГЛАВЛЕНИЕ

ЧАСТЬ VI

МОДЕЛИ МАССОВОГО ОБСЛУЖИВАНИЯ

<i>Глава 14. Модели очередей</i>	<i>299</i>
<i>Глава 15. Задержки транспорта на пунктах сбора пошлин</i>	<i>319</i>
<i>Глава 16. Модели упорядочения</i>	<i>345</i>

ЧАСТЬ VII

МОДЕЛИ ЗАМЕН

<i>Глава 17. Модели замен</i>	<i>370</i>
---	------------

ЧАСТЬ VIII

СОСТЯЗАТЕЛЬНЫЕ МОДЕЛИ

<i>Глава 18. Теория игр</i>	<i>398</i>
<i>Глава 19. Модели торгов</i>	<i>426</i>

ЧАСТЬ IX

ИСПЫТАНИЕ, КОРРЕКТИРОВКА И ПРИМЕНЕНИЕ МОДЕЛИ

<i>Глава 20. Данные для испытания</i>	<i>439</i>
<i>Глава 21. Корректировка и применение решения</i>	<i>452</i>

ЧАСТЬ X

АДМИНИСТРАТИВНОЕ УПРАВЛЕНИЕ ОПЕРАЦИОН- НЫМИ ИССЛЕДОВАНИЯМИ

<i>Глава 22. Отбор и подготовка персонала и организация операционных исследований</i>	<i>475</i>
<i>Предметный указатель</i>	<i>485</i>

ОТ РЕДАКТОРА РУССКОГО ПЕРЕВОДА

Происходящая сейчас научно-техническая революция вызвала появление нового объекта исследований в области управления, получившего название *большие системы*. Общая теория управления большими системами пока еще находится на начальном этапе развития, поэтому любая попытка добиться практического результата приводит к необходимости ограничивать область исследований конкретным классом больших систем. Один из самых значительных классов больших систем объединяет системы организационного управления, решающие задачи руководства системами, включающими людей. Главнейшие из этих задач состоят в определении целей функционирования организации, анализе ее деятельности и выборе решений, направленных на достижение поставленных целей. При этом возникает необходимость согласования работы различных подразделений и отдельных исполнителей, входящих в организацию с тем, чтобы направить их на достижение общей цели организации в целом.

Исследование операций представляет собой комплекс научных методов, предназначенных для решения указанных задач организационного управления. Методы исследования операций, возникшие во время второй мировой войны и бурно развивавшиеся в течение последних двух десятилетий, приобретают все большее значение. Они не только сформировались в новое и важное научное направление, но также стали эффективным средством решения широкого круга практических задач в промышленности, экономике и военном деле.

Литература на русском языке, излагающая совокупность методов исследования операций, пока еще весьма ограничена. Она содержит лишь несколько монографий, касающихся относительно узких вопросов, и ряд популярных брошюр. Предлагаемая русскому читателю книга является первым систематическим изложением, подводящим итоги достижений в этой области на начальном этапе ее развития. Хотя с момента выхода книги в свет прошло уже несколько лет, эта работа, ставшая классической, неоднократно переиздававшаяся за рубежом и переведенная на многие языки, представляет большой интерес для советского читателя. В ясной и доступной форме изложены в ней основные идеи и методы решения широкого круга операционных задач. Несомненной заслугой авторов является объединение в достаточно стройную систему многочисленных разрозненных методов решения разнообразных операционных задач на единой методологической основе. Авторам удалось найти удачное соотношение между теоретическими и прикладными аспектами управления.

Стремясь сделать книгу доступной и полезной для широкого круга читателей, включая операционистов, ученых, работающих в смежных областях, и «потребителей» — руководителей организаций, авторам пришлось ограничиться примерами относительно простых моделей и использованием не слишком сложного математического аппарата. Это позволило

без излишнего усложнения проиллюстрировать типичные постановки задач и принципы их решения.

К сожалению, авторам не удалось избежать повторений и некоторой разнохарактерности изложения, обычных, впрочем, для коллективных трудов. В книге не получили отражения некоторые новые классы операционных моделей, например сетевые модели, модели поиска и т. п., предложенные после выхода этой книги. Некоторые разделы (например, управление запасами, стратегия торгов) изложены относительно примитивно. Тем не менее большой фактический материал, ясное и доступное изложение, большое внимание к методологическим аспектам проведения операционных исследований делают эту книгу весьма полезной и интересной.

Советский читатель, разумеется, без труда заметит, что ряд примеров, которыми оперируют авторы, и некоторые точки зрения руководителей являются специфическими для экономики капиталистических стран и не могут быть непосредственно перенесены в условия нашей социалистической экономики. Тем не менее можно полагать, что появление на русском языке этого капитального труда будет не только способствовать распространению в Советском Союзе операционных методов решения задач организационного управления, но также будет стимулировать проведение научных исследований в этой весьма актуальной области.

А. Лернер

ПРЕДИСЛОВИЕ К АНГЛИЙСКОМУ ИЗДАНИЮ

Книга эта возникла из лекций по «Краткому курсу исследования операций», который ежегодно (начиная с 1952 г.) читается в Кейсовском технологическом институте. Этот курс, а соответственно и книга, преследует две цели:

1. Дать потенциальному заказчику операционных проектов общее представление об исследовании операций, необходимое для понимания возможностей и средств этой дисциплины.

2. Дать будущему исполнителю таких проектов общий обзор этой области и теоретическую базу, на которую он мог бы в дальнейшем опираться при накоплении знаний, необходимых для компетентного применения методов исследования операций. И в том и в другом случае общая задача заключается не в том, чтобы составить перечень «приемов», а в том, чтобы раскрыть принципы научного подхода к исследованию операций.

Для потенциального заказчика наибольший интерес представляют части I, II, III и X, а также разбор примеров, приводимых в других частях. Будущий исполнитель должен внимательно проработать весь материал.

Необходимо подчеркнуть, что книга представляет собой не более как введение в теорию. Она не предназначается в качестве справочника для специалиста по исследованию операций, опытного операциониста. При изложении материала технические вопросы по возможности упрощались, но так, чтобы это не вело к искажениям. Книга не требует очень высокой математической подготовки. Для чтения частей I, II, III и X вполне достаточно элементарных математических знаний, даже если они приобретены в «далеком прошлом», при условии, что читателя не страшат символика и абстрактные рассуждения. Для некоторых разделов частей IV — IX может понадобиться более глубокое знакомство с математикой, в частности с основами математического анализа.

Авторы включили в книгу некоторые главы и разделы, подготовленные другими лицами. Глава 3 написана Лорингом Дж. Миттеном (Loring G. Mitten), глава 4 — Ван Корт Хэаром (Van Court Hare), разбор примера в главе 4 — Р. Дж. Д. Гиллисом (R. J. D. Gillis), главы 8 и 18 — Элайэзером Неддором (Eliezer Naddor), главы 10 и 17 — Бертрамом Э. Рифасом (Bertram E. Rifas), главы 12 и 16 — Рэмом Васвани (Ram Vaswani), глава 14 — Уолтером Р. Ван Ворхисом (Walter R. Van Voorhis), глава 19 — Лоренсом Фридманом (Lawrence Friedman), разбор примера на осуществление решения в главе 21 — Элизабет А. Смолл (Elizabeth A. Small) и глава 22 — Джозефом Ф. Макклоски (Joseph F. McCloskey). Пример 1 в главе 6 частично перепечатан из статьи Пола Стиллсона (Paul Stillson), а глава 15 — из статьи Лесли С. Эди (Leslie C. Edie). Перепечатка сделана с любезного разрешения названных лиц и редакторов журналов «Индустриэл Куолити Контрол» («Industrial Quality Control») и «Оперейшнз Рисерч» («Operations Research»).

Этот материал был отредактирован и организован в единый связный текст. В ряде случаев изменения первоначального текста были очень значительны, поэтому авторы принимают на себя всю ответственность за возможную неточность или непоследовательность изложения.

Авторы особенно подробно рассматривают применение операционных методов исследования к решению промышленных задач, поскольку это непосредственно связано с их работой. Авторы хорошо понимают, что исследование операций широко применялось и применяется при решении военных задач, но, как показано в гл. 15, эта наука таит в себе огромные, еще не использованные возможности приложения для решения задач в сфере общественно-полезного производства. Фактически двое из авторов этой книги затратили немало времени и сил на попытку применить операционные методы исследования для решения задач городского планирования и организационных задач профсоюзов. Авторы уверены, что дальнейшее развитие исследования операций приведет к распространению этих методов на все области административного руководства. Но пока вряд ли можно оспаривать тот факт, что военная и промышленная администрация проявляют больше заинтересованности в поощрении и использовании операционных проектов, чем органы гражданского управления.

Авторы чрезвычайно обязаны следующим лицам за их конструктивную критику во время подготовки книги: У. У. Куперу (W. W. Cooper), Элфреду У. Джоунзу (Alfred W. Jones), Дж. С. Минасу (J. S. Minas), Джону Ф. Муту (John F. Muth), Леону Прицкеру (Leon Pritzker), Ричарду С. Раднеру (Richard S. Rudner) и Максу А. Вудбери (Max A. Woodbury). Неоценимую редакторскую помощь оказали Беверли Бонд (Beverly Bond), Ричард Э. Дил (Richard E. Deal) и Артур Дж. Яспен (Arthur J. Yaspán).

Особую признательность за одобрение и энергичную поддержку этого и других начинаний операционной группы авторы хотели бы выразить главе нашего отделения Клею Г. Холлистеру (Clay H. Hollister), деканам Элмеру Г. Хатчиссону (Elmer H. Hutchisson) и Карлу Б. Макичрону (Carl B. McEachron).

Авторы хотели бы также поблагодарить Кэрол Мару Придо (Carol Mara Prideaux) и Грейс Уайт (Grace White) за их работу по подготовке рукописи.

*С. Уэст Черчмен
Рассел Л. Акоф
Э. Ленард Арноф*

Кливленд, Огайо
25 октября 1956 г.

Часть I

ВВЕДЕНИЕ

В первой части этой книги рассматривается предмет и особенности исследования операций. В главе 1 дается определение и общая характеристика предмета. Центральным в этом обсуждении является понятие *цели* операционного исследования, которая заключается в достижении системного (всестороннего) подхода к задаче. Это положение наглядно проиллюстрировано на примере, изложенном в главе 2. В главе 3 обсуждается пример, который убедительно раскрывает причины, требующие обязательного включения в операционную группу представителей различных дисциплин. Вопросы конкретного состава самой операционной группы рассмотрены в главе 22.

Глава 1

ОБЩАЯ ХАРАКТЕРИСТИКА ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

ОТЛИЧИТЕЛЬНЫЕ ЧЕРТЫ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

Никакая наука не рождается в один прекрасный день. Каждая наука появляется в результате совпадения все возрастающего интереса к определенному классу задач и уровня развития научных принципов, методов и средств, с помощью которых оказывается возможным решить эти задачи. Исследование операций не является исключением. Его исторические корни так же глубоки, как корни науки и административного руководства. Но свое название исследование операций получило только в 1940 году *. Первоначально исследование операций возникло во время второй мировой войны в Англии, а затем разработка этого направления была быстро подхвачена в США. Таким образом, впервые исследование операций было применено к военным задачам, а уже после войны — к задачам финансового, промышленного и гражданского управления. Сперва этот процесс шел быстрее в Англии, пока в 1951 году операционные исследования не завоевали прочных позиций в американской промышленности и стали бурно развиваться в США.

Хотя работа в той области, которая впоследствии получила название исследования операций, началась в военных организациях, ее развитие и оформление в самостоятельную дисциплину может быть описано по аналогии с хорошо известным процессом развития производства. До промышленной революции производство было представлено главным образом мелкими предприятиями, где все руководство осуществлял один хозяин, который делал закупки, планировал и направлял работу, сбывал продукцию, нанимал и увольнял рабочих и т. п. Впоследствии произошло разделение административных функций. Появились руководители производственных отделов, отделов сбыта, финансовых отделов, отделов кадров и другие. Усиливающаяся механизация, сопровождаемая частичной автоматизацией, привела к еще более быстрому росту промышленного производства, что вызвало дробление операций и дальнейшее расчленение административных функций. Так, производственные отделы оказались разделенными на более мелкие группы, занимающиеся вопросами эксплуатации и ремонта, контроля качества, планирования, снабжения, хранения готовой продукции и т. п.; эти группы нередко возглавляют лица, занимающие особые административные должности. Появление многочисленных отдельных операций неизбежно потребовало создания административных должностей по руководству этими операциями. В настоящее время все меньшее число руководителей производственных отделов имеет непосредственную связь с производственными цехами. Они фактически превратились в работников, выполняющих чисто организационные функции.

Одновременно с ростом дифференциации и расчленения административных функций возрастал интерес ученых к задачам, возникающим в различных производственных подразделениях промышленной организации. В частности, ученые стали все более пристально заниматься вопросами промышленного производства, и в результате их совместных усилий появился ряд новых отраслей прикладных наук со своими специальностями инженеров-механиков, инженеров-химиков, инженеров промышленного строительства,

* Краткую историю исследования операций можно найти в книге Трефтена (Trefethen) [8].

инженеров по статистическому контролю качества, специалистов по организации производства. В других областях появились специалисты по исследованию конъюнктуры рынка, по промышленной экономике, эконометрии, психологии трудовых отношений, промышленной социологии и аналогичным прикладным научным дисциплинам.

В этот период дифференциации и расчленения административных функций начал возникать и формироваться новый класс задач управления, задач, которые могут быть названы *задачами организационно-управленческого характера*. Эти задачи возникли как прямое следствие функционального разделения труда на предприятии, которое привело к необходимости организационной деятельности. В любой организации отдельное ее специализированное подразделение (управление, отдел или группа) выполняет определенную часть общей работы. Выполнение каждой части работы необходимо для достижения общих целей организации. Однако в результате такого разделения труда у каждого специализированного подразделения возникают свои собственные цели. Например, производственный отдел обычно ставит своей целью максимальное снижение себестоимости продукции при максимальном росте объема ее выпуска. Отдел сбыта стремится минимизировать расходы на продажу единицы продукции при увеличении общего объема сбыта. Финансовый отдел стремится оптимизировать политику капиталовложений фирмы. Отдел кадров старается нанять хороших работников при минимальной оплате и удерживать их на своем предприятии и т. д. Все эти цели не всегда согласуются; в действительности они очень часто приходят в прямое противоречие друг с другом.

Рассмотрим в качестве примера взаимоотношения различных отделов по вопросу стратегии управления запасами в общем случае. Производственный отдел заинтересован в возможно более длительном и непрерывном серийном производстве, ибо такое производство сокращает затраты на наладку оборудования, а следовательно, снижает общие производственные расходы; в то же время оно может привести к большому затовариванию готовой или неоконченной продукцией при сравнительно узкой номенклатуре. Отдел сбыта стремится произвести поставку одновременно возможно большего ассортимента товаров. Отсюда он заинтересован как в разнообразии товаров, так и в большом их количестве. Ему также хотелось бы, чтобы производственный отдел был более гибок и мог принимать небольшие срочные индивидуальные заказы. Финансовый отдел стремится свести к минимуму запасы, так как он заинтересован в снижении процента капиталовложений, не дающих прибыли. Отдел кадров стремится закрепить за предприятием рабочую силу, а этого можно достигнуть только в том случае, если в период делового спада продолжать производить продукцию, хотя и хранить ее на складе, и т. д.

Стратегия в отношении запасов отражается на деятельности каждого специализированного подразделения промышленной организации. Наиболее благоприятная стратегия для одного отдела очень редко оказывается наиболее благоприятной и для другого. Задача, следовательно, формулируется следующим образом: какая стратегия в отношении запасов будет наиболее благоприятной для организации в целом? Это задача организационного типа, поскольку (а) она связана с проблемой оптимизации работы всей организации в целом и (б) затрагивает противоречивые интересы отдельных специализированных подразделений данной организации.

Задачи подобного типа могут также возникать и действительно возникают даже в пределах какого-либо одного из подразделений промышленной организации. Например, относительно проблемы запасов различные

интересы могут столкнуться даже в пределах производственного отдела, а это значит, что может возникнуть задача организационного типа для руководителя данного отдела. Например, одно из подразделений производственного отдела может быть заинтересовано в снижении затрат на переоборудование, что, однако, может сказаться на увеличении объема запасов некоторых видов продукции до размеров, превышающих вместимость существующих складских помещений. Это вызовет столкновение интересов внутри производственного отдела, и разрешение этого конфликта мы назовем задачей организационного типа для данного отдела.

Необходимо отметить, что расхождение целей отдельных подразделений одной организации не является недостатком. Если большая группа людей пытается выполнить определенную задачу, то вряд ли они могут действовать так, как на их месте поступал бы один человек. И бессмысленно стараться разрабатывать для крупной промышленной организации такие планы, которые бы предполагали, что каждый знает и может оценить работу любого другого. В таких случаях, по-видимому, единственным решением может быть распределение функций. Таким образом, задачи организационного типа возникают в результате необходимости разделения функций. Их решение очень редко принимает форму совета, обращенного к отдельным подразделениям: «Давайте постарайтесь понять проблемы друг друга». Оно требует прежде всего чрезвычайно тонкого понимания соотношения целей подразделения и общих целей организации. Необходимо создавать у подразделений заинтересованность в достижении своих собственных целей, ибо чрезмерная забота об общих целях может привести к бесплодности всех усилий. Поэтому, когда мы говорим о «всеобщем оптимуме», мы имеем в виду стратегию, которая принимает во внимание необходимость разделения функций внутри организации.

С появлением задач организационного типа возникла должность консультантов при администрации. Роль консультантов заключается в том, чтобы помочь решению возникающих организационных задач, используя свой опыт по решению задач аналогичного типа в других областях. Используемый ими метод состоит в нахождении того, что является общим для определенных задач данного класса при анализе предлагаемых решений. Естественно, что постепенно начали предприниматься попытки отыскать общую структуру (построить «модель») этих решений и общие принципы проверки правильности найденных структур. Эти попытки привели к необходимости научного подхода к исследованию задач организационного типа, хотя надо сказать, что научные достижения использовались в этой области время от времени и до появления исследования операций.

Во время второй мировой войны военные власти обратились к ряду ученых с призывом помочь решению стратегических и тактических задач. Многие из этих задач принадлежали именно к тому типу, который мы называли организационными задачами. Представители различных отраслей науки были сведены в рабочие группы, перед которыми ставились задачи оптимизации использования ресурсов. Это были первые операционные группы.

Целью исследования операций, дисциплины, возникшей в связи с развитием промышленных организаций, является обеспечение руководства научной основой для решения задач, связанных с вопросами взаимодействия различных подразделений организации в интересах общих целей всей организации. Решение, которое представляется наиболее выгодным для всей организации в целом, считается *оптимальным решением*; решение, которое наиболее выгодно для одной или более частей организации, назовем *субоптимальным*. Задача установления критериев оптимальности решения

сама по себе чрезвычайно сложна и очень специфична, поэтому более подробно она будет рассмотрена в главе 5.

При проведении операционных исследований стараются отыскать наилучшие решения для возможно большей части всей организации. Например, при попытке разрешить задачу ремонта на предприятии операционными методами стремятся учесть влияние различных вариантов политики проведения ремонтов на производство в целом. Можно пойти даже дальше и попытаться определить, как это влияние на данную фирму в свою очередь отражается на всей промышленности и т. д. По возможности также предпринимается попытка учесть, как воздействие на производственный отдел скажется на других отделах и на фирме в целом. В исследовании операций стремятся учитывать взаимодействия или причинно-следственные связи настолько, насколько это представляется существенным. Однако в любом конкретном случае объем операционных исследований практически обычно ограничен, либо в связи с тем, что доступ на более высокие уровни организации закрыт, либо в связи с ограничениями во времени, деньгах и ресурсах. Это необходимо иметь в виду при чтении книги. Всегда существует расхождение между идеальными устремлениями и реальными возможностями. Тем не менее основная цель исследования операций сводится к отысканию оптимальных решений задач организационного типа с учетом функционирования всей организации в целом.

Такой общий характер цели операционных исследований является примером системного подхода, поскольку под «системой» подразумевается взаимосвязанный комплекс функционально соотнесенных компонентов. Так, фирма является системой «человек — машина». Но не все системы включают человеческий или социальный компонент. Автомобиль, например, состоит из таких функциональных частей, как мотор, передача, радиатор, генератор. Все эти и другие части собраны в единый механизм, который может удовлетворять ряд потребностей. Эффективность функционирования каждой части зависит от отлаженности механизма в целом, а эффективность работы всего механизма зависит от того, как функционирует каждая его отдельная часть.

Проблемы оптимизации конструкции механических систем сходны с проблемами оптимизации систем «человек — машина», но не тождественны им. Оба типа систем связаны с противоречиями интересов. Покупатели автомобилей хотят иметь надежные, экономичные, удобные, просторные, изящные машины, обладающие высокой скоростью. Но нельзя в одинаковой степени удовлетворить все эти требования одновременно. Поэтому задача конструирования требует оптимизации с учетом некоторого множества по крайней мере частично противоречивых целей. Проблема «разделения труда» в промышленных организациях отлична от вопросов функционирования компонентов в механических системах, так как в системах, куда входят люди, существует серьезная проблема создания заинтересованности для работы различных подразделений по выполнению соответствующих функций.

Применение научных методов к построению механических систем и систем «человек — машина» иногда называют анализом систем и ставят знак равенства между ними и исследованием операций. В определенном смысле конструирование и оценка боевых систем и систем связи является операционным исследованием. Но данная книга написана с ориентацией на системы, включающие человека, поскольку именно на это делается упор в практике применения исследования операций в сфере деловой и производственной деятельности.

Системный подход к задачам не означает, что сформулированная в наиболее общем виде задача должна быть решена в одном исследовательском проекте. Как бы это ни было желательно, практически реализовать это удается очень редко. На практике общая задача обычно решается по частям в определенной последовательности. Во многих случаях общая задача не может быть сформулирована заранее, но решение на одном этапе позволяет определить следующий шаг. Например, система контроля качества может потребовать установления наиболее экономичного количества выпускаемой продукции по видам изделий. После того как оно будет определено, может оказаться, что различные виды изделий в данном количестве не могут быть произведены на имеющемся оборудовании в отведенное для этого время. Этот вывод приводит к постановке новой задачи, решение которой повлияет на решение, полученное на предыдущем этапе. Короче говоря, хотя одно-временная оптимизация всех частей системы чрезвычайно желательна, практические ограничения, как правило, требуют последовательной оптимизации отдельных частей системы с параллельной взаимной корректировкой локальных оптимумов для определения оптимального решения всей задачи в целом.

Из утверждения о том, что исследование операций стремится описать возможно большую часть системы, не следует, что нужно начинать с исследования системы в целом. Большинство операционных исследований начинается с известных задач ограниченного масштаба. Но в дальнейшем границы исследования расширяются, насколько позволяют обстоятельства. По существу, объем исследования является мерилom, определяющим одну из особенностей исследования операций. Соответственно операционное исследование очень часто начинается с тех же задач, с которых мог бы начать инженер-механик, инженер промышленного строительства, инженер-химик или исследователь конъюнктуры рынка, но оно редко заканчивается задачами такого типа. Эта сторона исследования операций будет проиллюстрирована на примере, разобранном в главе 2.

Для исследования операций характерно, что при решении каждой проблемы возникают все новые и новые задачи. Отсюда ясно, что применение операционных методов неэффективно, если решаются узкие, ограниченные задачи. Наибольшая выгода может быть получена при непрерывном преемственном исследовании, т. е. при последовательном переходе от одной задачи к другой. Это вытекает из самого определения целей исследования операций, которое приводилось выше.

Одной из наиболее существенных черт исследования операций является стремление найти оптимальное решение, определить оптимальную стратегию, конструкцию. Задача заключается не в том, чтобы найти решение лучше существующего, а в том, чтобы найти самое лучшее решение из всех возможных. Такое решение не всегда удается получить из-за ограничений, накладываемых современным состоянием науки, недостатком времени, средств или возможностей. Но усилия при исследовании операций непрерывно направлены на поиск оптимального решения или решения, возможно более близкого к нему.

В некоторых случаях при проведении операционного исследования не удается определить оптимальное решение в связи с тем, что одна из существенных характеристик системы не может быть оценена в пределах ограничений, наложенных на задачу. Например, может оказаться **невозможным** оценить в денежном выражении потери времени, в течение которого клиенту приходится ждать, пока его обслужат. Однако и в этом случае можно найти оптимальное решение для каждого значения, которое может принять время

ожидания. Тот, кто отвечает за принятие решения, в такой ситуации должен, опираясь на свой здравый смысл, приписать определенную оценку времени ожидания. После этого можно отыскать оптимальное решение с учетом принятой оценки. При этой процедуре можно также выявить количественную зависимость между значением критерия и функционированием системы. Например, исследование операций может показать, какова общая величина ожидаемых потерь для данного канала обслуживания при каждом возможном значении времени ожидания, затраченного клиентом. В любом случае окончательное решение зависит от тех, кто руководит операциями, а не от тех, кто их исследует. Операционная группа может только рекомендовать решения или основу для выбора определенного решения. Она может также помочь в осуществлении решения, когда оно принято.

Суммируя изложенное, можно сказать, что исследование операций в самом общем смысле может быть охарактеризовано как применение научных принципов, методов и средств к задачам, связанным с функционированием организационных систем, с целью предоставить тем, кто управляет этими системами, оптимальные решения. Данная работа ограничивается вопросами применения исследования операций к задачам организационно-управленческого характера, возникающим в крупных организациях.

Это не значит, что мы собираемся игнорировать другие виды анализа систем, но цель построения модели, например, системы производства будет заключаться в том, чтобы дать исследователю возможность понять, как производство связано с организационной деятельностью. Поэтому мы начинаем наше обсуждение в главе 4 с рассмотрения подходов к формулированию задачи, что в основе своей является попыткой подойти к задаче с возможно более общей точки зрения, учитывая всю систему в целом. Читатель при знакомстве с дальнейшим изложением должен помнить об этой главе, когда будет рассматриваться конкретные особенности решения той или иной специфической задачи. Так, внимательный читатель будет иметь в виду не просто решение задач, связанных с запасами как таковыми, а решение организационных задач, в которых вопросы запасов играют существенную роль. Мы далее увидим, что модель запасов является частным методом, а то, как она связана с общим исследованием, составляет общий принцип. В процессе чтения глав, рассматривающих различные модели, читатель должен непрерывно спрашивать себя, как данная модель может быть применена к конкретной системе, с которой он хорошо знаком. Таким образом читатель избежит неверного восприятия изложения и не будет подходить к нему как к описанию отдельных приемов и случаев.

КОМПЛЕКСНЫЙ ПОДХОД

Исследование операций выделилось из других наук, многое у них позаимствовав. Это наблюдается при «рождении» каждой новой теоретической дисциплины. И всегда затруднительно отделить новую область от тех, из которых она выделилась, так как существует пересечение задач, методов и принципов. Со временем происходит полная дифференциация, и постепенно тех, кто работает в новой области, перестают одолевать таким вопросом, как: «А чем это отличается от того-то и того-то?» Быстрое развитие исследования операций как самостоятельной дисциплины со своим собственным названием свидетельствует о ее растущем признании. Но дифференциация от других областей еще далеко не завершена.

Пересечение принципов, методов и средств исследования операций и других областей в значительной степени объясняется тем, как проводились

и до сих пор проводятся операционные исследования, выполняемые группами ученых, отдельные члены которых представляют различные теоретические и прикладные дисциплины. Нередко, например, можно встретить группу из математиков, физиков, психологов и экономистов, которые работают вместе над задачами оптимизации политики капиталовложений. Продуктивность таких групп, состоящих из представителей различных специальностей, при решении класса организационных задач, составляющих предмет исследования операций, не случайна.

Когда ученый сталкивается с задачей нового типа, он, как и любой другой на его месте, старается прежде всего определить сущность задачи и установить, приходилось ли ему ранее иметь дело с задачами подобного вида в иных конкретных условиях, в частности в своей специальной области. Если он устанавливает аналогии с задачами, с которыми встречался в своей области, он может посмотреть, насколько применимы методы решения, которыми он воспользовался бы при решении своих специальных задач, к той новой задаче, с которой он столкнулся. Таким образом, он вводит в круг возможных подходов к задаче те методы и принципы, которые вне связи с его специальностью и опытом могли бы даже и не прийти никому на ум. И когда одну задачу решают специалисты из различных областей, то круг возможных подходов к ее решению, естественно, расширяется.

Так, например, специалист в области электроники, изучая проблему производства и регулирования запасов, может быстро понять, что колебания объема запасов являются функцией времени, которое проходит между колебаниями на рынке и приспособлением к ним уровня производства. Фактически он рассматривает задачу как задачу по конструированию управляемой системы, в которой необходимая информация относительно изменения рыночной конъюнктуры поступает быстро и безошибочно в пункт управления производством, откуда могут быть посланы команды на изменения в производственном процессе, снижающие некоторые расходы. Он фактически перевел задачу на язык теории автоматического регулирования, и он знает, как решать такого рода задачи. Пример этот вовсе не гипотетический.

С другой стороны, инженер-химик может иначе посмотреть на ту же задачу и попытаться сформулировать ее в терминах теории непрерывных процессов, и, если ему это удастся, он сможет применить к решению этой задачи имеющиеся в его распоряжении готовые методы.

Какой из возможных подходов окажется наиболее продуктивным, зависит от обстоятельств. Исследовательская группа рассматривает предложенные варианты и выбирает один из них или выдвигает совершенно новый, в котором сочетаются черты предложенных ранее подходов.

Одна из основных идей создания комплексных операционных групп заключается в том, чтобы применить к решению данной задачи наиболее передовую научную методику или разработать новую, которая была бы более эффективна в отношении решения подобных задач, чем любая из уже имеющихся. По существу, эта идея сводится к тому, что ни в одной голове, взятой в отдельности, не может разместиться столько потенциально полезной научной информации, сколько в нескольких, взятых вместе.

Еще одним существенным достоинством такого подхода является то, что большинство систем «человек — машина» требует подхода с различных точек зрения, например с точки зрения физической, биологической, психологической, социологической, экономической и инженерной. Эти различные аспекты системы могут быть лучше поняты и проанализированы теми, кто специализируется в той или иной области. Те, кто непосредственно управ-

ляет данной системой, могут не иметь достаточного представления о какой-то одной или нескольких ее сторонах, а следовательно, не располагать полной картиной ее функционирования. Другими словами, для того чтобы представлять систему как единое целое, недостаточно знать ее части и их взаимодействие; необходимо также видеть различные аспекты функционирования системы. В комплексной группе количество аспектов, которые могут быть подробно рассмотрены, возрастает. Это положение будет проиллюстрировано при рассмотрении примера, изложенного в главе 3.

ПРИНЦИПЫ, МЕТОДЫ И СРЕДСТВА ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

Поскольку в исследовании операций определенные классы задач встречаются все чаще и чаще, они, естественно, подвергаются более тщательному изучению. В результате для многих классов часто встречающихся задач разработаны новые принципы подхода или предложена модификация прежних. Постепенно арсенал принципов, методов и средств, разработанных или приспособленных специально для целей операционного исследования, значительно вырос. Он достиг уже такого уровня, когда одному человеку трудно быть осведомленным относительно всех нововведений. Это уже само по себе привело к важным последствиям.

Десять лет назад любой, кто интересовался этой областью, обладал творческим складом ума и хорошей научной или инженерной подготовкой, мог легко стать операционистом. Для этого не нужно было специальной подготовки или образования. Теперь возможность такого свободного перехода в область исследования операций сужается, поскольку в связи с быстрым прогрессом в этой области требуется все больше и больше времени, чтобы быть в курсе тех принципов, методов и средств, которыми она в настоящее время располагает.

Вместе с тем с увеличением необходимого объема знаний появляется возможность преподавать исследование операций как отдельный предмет. Во многих университетах, колледжах и технических институтах вводятся соответствующие курсы, а в некоторых из них окончившим курс предлагается совершенствоваться дальше по программе, которая дает возможность получения ученой степени. Именно в результате интенсивного развития этой области как самостоятельной, со своими собственными принципами, методами и средствами, появилась возможность написания этой книги.

Здесь мы в основном сосредоточим свое внимание на все возрастающем объеме знаний, требуемых в исследовании операций. Мы не будем останавливаться на описании большого числа различных средств, необходимых работникам в этой области для решения своих задач. В частности, мы не намерены излагать здесь те математические и статистические сведения, которые требуются для успешного проведения операционного исследования. Мы не будем затрагивать такие области, как анализ затрат, экономика, прогнозирование, использование вычислительной техники, которые также важны для практического использования исследования операций. Все это человек, специализирующийся в этой области, безусловно, должен знать, и тот факт, что эти вопросы не включены в данную книгу, не следует рассматривать как принижение их значения. Весь этот материал невозможно поместить в одной книге, да для этого и нет необходимости. По всем этим вопросам существует вполне доступная литература. Но вот чего в ней нет, так это общего введения, знакомящего читателя с принципами, идеями, методами, которые были разработаны в исследовании операций или появились сперва в других областях, а затем были приспособлены для

новых целей. Эту задачу мы и ставили перед собой, когда создавали эту книгу.

Термины «средства», «методы» и «принципы», которые часто в различных научных работах используются взаимозаменяемо, в данной книге строго разграничены. В том понимании, как они употребляются в данной работе, они связаны между собой примерно так же, как средства, используемые при строительстве здания, методы использования этих средств и разработка принципов строительства, которые предусматривают использование данных методов и средств надлежащим образом. Например, таблица случайных чисел является научным средством. То, как средство используется (например, метод Монте-Карло, который будет рассмотрен в главе 7), является научным методом. Исследовательская идея, которая требует использования метода Монте-Карло и таблиц случайных чисел, является научным принципом. Аналогично этому математический анализ есть не что иное, как научное средство; использование математического анализа для нахождения оптимального значения переменной в математической модели системы является научным методом, а замысел применения математической модели для оптимизации системы является научным принципом.

Хотя и справедливо утверждение о том, что все науки пользуются некоторыми общими принципами, методами и средствами, также справедливо и то, что каждая отрасль науки использует и свои собственные, специфические, только ей свойственные принципы, методы и средства, в чем отражается особенность предмета, который с их помощью изучается. И сама наука может считаться настолько развитой, насколько она сумела развить принципы, методы и средства, приспособленные к нуждам изучения ее специфического предмета.

Учебники по таким давно определившимся областям науки, как физика и химия, останавливаются на вопросах принципов лишь вскользь. Их главным образом интересуют методы и средства. В данной книге основное внимание уделяется принципам. Это делается потому, что в новой области исследований вопрос подхода является более важным, чем вопрос использования того или иного метода или средства. До того, как исследование операций начало вырабатывать свои методы и средства, оно все равно было полезным вследствие силы своего принципа решения задач. Как мы уже говорили выше, читатель должен всегда рассматривать тот или иной метод только как некоторый способ решения более общей задачи, а не как нечто ценное само по себе. Таким образом он сумеет избежать как опасности ограничиться одним или несколькими отдельными методами, так и излишней привязанности к ним. Непредвзятость в отношении конкретных методов, наряду с широтой представления об их полезности, в сочетании с умением оценить задачу в целом являются основой твердого принципа в науке.

ЭТАПЫ ОПЕРАЦИОННОГО ИССЛЕДОВАНИЯ

Еще десять лет назад от человека, занимающегося вопросами исследования операций, трудно было бы добиться описания хода операционного исследования. Сейчас его трудно удержать от этого. Если записать мнение каждого операциониста относительно порядка проведения операционного исследования, то мнения в некоторых отношениях будут отличаться. Но вместе с тем обнаружится и много общего. В частности, большинство согласится, что основными этапами выполнения операционного проекта являются:

1. Постановка задачи.
2. Построение математической модели изучаемой системы.

3. Нахождение решения с помощью модели.
4. Проверка модели и полученного с ее помощью решения.
5. Построение процедуры подстройки решения.
6. Осуществление решения.

Каждый из этих этапов будет предметом особого рассмотрения в последующих главах, но представляется полезным дать читателю некоторое предварительное представление о них, приведя краткую характеристику отдельных этапов.

Постановка задачи. Задача должна быть сформулирована с точки зрения заказчика и исполнителя операционного проекта. Заказчик — это лицо (или группа лиц), которое руководит исследуемыми операциями, или, как принято говорить, принимает решение. При формулировании задачи с точки зрения заказчика должен быть проведен анализ системы, находящейся под его управлением, а также его целей и возможных вариантов действий. Должно быть установлено, как исследуемые решения могут повлиять на тех людей, цели и действия которых также подлежат учету. Подход, который мы определили как системный, тесно связан с установлением непосредственных целей. Проводя операционные исследования, необходимо учесть как можно более широкий круг выдвигаемых целей. В самых общих словах, задача исследования заключается в том, чтобы определить, какой из возможных вариантов действий наиболее эффективен в отношении множества связанных с проблемой целей. Следовательно, при формулировании задачи исследования должен быть установлен некоторый критерий эффективности и определена его пригодность.

Построение математической модели. Модель выражает эффективность исследуемой системы как функцию множества переменных, из которых по крайней мере одно поддается управлению. Операционная модель в самом общем виде может быть представлена уравнением

$$E = f(x_i, y_i),$$

где E — критерий эффективности системы, x_i — управляемые переменные системы и y_i — неуправляемые переменные. Ограничения, наложенные на переменные, могут быть выражены в дополнительной системе равенств или неравенств.

Нахождение решения с помощью модели. Существуют два метода получения оптимального решения (или некоторого приближения к нему) с помощью модели: аналитический и численный. Аналитические процедуры сводятся к использованию математической дедукции. Это требует использования различных разделов математики, таких, как математический анализ и матричная алгебра. Аналитические решения получаются в абстрактном виде, т. е. подстановка чисел вместо символов обычно производится уже после того, как будет получено решение.

Численные процедуры состоят в принципе в подборе различных значений для управляемых переменных модели, сопоставлении полученных данных и выборе того набора значений, который дает наиболее выгодное решение. Такие процедуры могут варьироваться в широком диапазоне от простого метода проб и ошибок до сложных итераций. Процедура итераций состоит в том, что в результате последовательных проб производится попытка подойти к оптимальному решению. Кроме того, итерация обычно дает некоторый набор правил, которые позволяют опознать оптимальное решение, когда оно найдено.

Некоторые выражения в модели не могут быть численно определены с достаточной точностью вследствие каких-либо математических или чисто практических причин. Во многих подобных случаях для получения примерной оценки предложений может быть применен особый вид случайных выборов, называемый методом Монте-Карло.

Проверка модели и решения. Модель всегда лишь частично отображает действительность. Модель можно считать хорошей, если, несмотря на свою неполноту, она может точно предсказывать влияние изменений в системе на общую эффективность всей системы. Адекватность модели может быть проверена путем определения степени точности предсказания влияния этих изменений. Решение может быть оценено путем сопоставления результатов, полученных без использования данного решения, и результатов, полученных при его применении. Эти оценки могут производиться ретроспективно с использованием ранее полученных данных или путем практических испытаний и предварительных проверок. Проверка требует очень тщательного анализа данных с целью определения, какие из них являются существенными и какие нет.

Построение процедуры подстройки решения. Решение, полученное на модели, действительно только до тех пор, пока неуправляемые переменные сохраняют свои значения и соотношения между переменными в модели остаются постоянными. Когда же значение одной или более неуправляемых переменных либо одно или более отношений между переменными существенно изменилось, само решение «выходит из-под контроля» и возможность управления им теряется. Существенность изменения зависит от того, насколько решение отклоняется от истинного оптимума при изменившихся условиях, и от затрат на изменение действующего решения. Значит, для того чтобы установить процедуру подстройки решения, необходимо разработать средства определения того, когда возникают существенные изменения; кроме того, должны быть установлены правила такой модификации решения, которая бы учитывала эти изменения.

Осуществление решения. Проверенное решение должно быть представлено в виде ряда рабочих процедур, которые могут быть легко поняты и применены теми, кто будет отвечать за их осуществление. При этом должны быть определены и соблюдены необходимые изменения в существующих процедурах и ресурсах.

Перечисленные этапы почти никогда, за редкими исключениями, не проводятся в указанном порядке. Более того, некоторые из этих этапов могут осуществляться одновременно. Например, во многих операционных проектах формулирование задачи продолжается фактически вплоть до завершения всего исследования. В процессе исследования, как правило, различные направления работы непрерывно взаимодействуют друг с другом.

ТИПИЧНЫЕ КЛАССЫ ПРОЦЕССОВ И ЗАДАЧ

Методы и средства, которые рассматриваются ниже, распадаются на определенные классы в зависимости от типа процесса, к которому они могут быть применены. В большинстве случаев эти процессы уже имеют общепринятые названия; в отдельных случаях таких названий пока нет, и авторы

были вынуждены приписать этим классам некоторые наименования, которые они сочли достаточно подходящими.

Эти типы процессов и классы связанных с ними задач, которые мы будем рассматривать, можно кратко охарактеризовать следующим образом.

Процессы создания и хранения запасов. Процессами, связанными с проблемой запасов, принято называть такие процессы, которые требуют либо обоих, либо одного из двух следующих решений: (а) сколько заказывать (производить или покупать) и (б) когда заказывать. Эти решения требуют регулирования запасов, что связано с затратами по одной или нескольким из следующих статей: заказ или переоборудование; дефицит или задержка; изменение уровня производства или закупок. В качестве средств, применяемых к решению этих задач, используются уравнения экономического размера партии, а также аппарат линейного, динамического и квадратичного программирования.

Процессы распределения. Эти процессы возникают, когда (а) существует ряд операций, которые необходимо выполнить, и ряд различных путей их выполнения, (б) нет в наличии ресурсов или средств, обеспечивающих выполнение каждой из этих операций наиболее эффективным образом. Задача в таком случае заключается в соединении операций и ресурсов таким образом, чтобы добиться максимального общего эффекта. Как ресурсы, так и операции могут быть конкретно заданы. Если задано что-либо одно из двух, то задача заключается в том, чтобы определить, при каком втором в комбинации с заданным первым можно в конечном итоге добиться наибольшего эффекта.

Средствами, которые в практике работы наиболее часто использовались при решении задач распределения, являются линейное и другие виды математического программирования.

Процессы обслуживания. Эти процессы связаны с наличием клиента, требующего обслуживания. За исключением очень редких случаев, либо клиенту, либо обслуживающему, либо как тому, так и другому приходится ждать. Все виды ожидания связаны с расходами. Задача заключается в том, чтобы регулировать появление клиентов или определять объем и организацию обслуживания с тем, чтобы свести к минимуму оба эти вида ожидания и связанные с ними расходы.

К задачам по определению объема обслуживания и времени появления клиентов (составление расписания) применима *теория очередей*. К задачам по определению порядка обслуживания клиентов применима *теория расписаний*. И наконец, к задачам определения распределения и группировки каналов и других элементов процессов обслуживания в систему обслуживания применима *теория балансирования линий*.

Процессы замены. Процессы замены распадаются на два основных класса в зависимости от вида износа оборудования: либо оно устаревает и становится несовременным (менее эффективным) в результате длительного срока службы или появления нового оборудования (например, станки), либо не устаревает, но полностью выходит из строя, «гибнет» (например, электролампы).

Для устаревающих видов оборудования задача заключается в установлении времени замены с тем, чтобы сократить расходы на новое оборудование, на ремонт и эксплуатацию старого или расходы, связанные со снижением

производительности оборудования. Для видов оборудования, которые полностью выходят из строя, задача заключается в том, чтобы определить, какие элементы следует заменить (например, все, кроме тех, которые были установлены на прошлой неделе) и как часто их менять с тем, чтобы снизить расходы на оборудование, на замену частей и другие затраты, связанные с выходом их из строя.

Задачи эксплуатации и ремонта могут рассматриваться как частный случай задач замены, поскольку здесь требуется, как правило, замена отдельной части, а не всего комплекса оборудования. Отсюда как к задачам замены оборудования, так и к задачам ремонта применим один и тот же подход.

Состязательные процессы. Состязательные процессы — это процессы, в которых эффективность решения одной стороны может оказаться сниженной в связи с решением другой стороны. Наиболее часто встречающаяся в проектах операционного исследования ситуация носит название *игра*. Игра характеризуется количеством участников, правилами игры, сформулированными с учетом всех возможных допустимых действий, некоторым множеством конечных состояний (выигрыш, проигрыш, ничья и т. п.) и наградами или потерями, вызываемыми этими результатами. Основные методы, связанные с решением этого класса задач, известны как *теория игр*.

Другим видом состязательного процесса является такая ситуация, при которой имеет место *аукционный торг*. Она отличается от игры следующими признаками: (а) количество участников обычно неизвестно, (б) количество партий, как правило, не ограничено, (в) награды и проигрыш заранее неизвестны и могут лишь быть примерно оценены, (г) результат состязания (выигрыш или проигрыш) также неизвестен и может быть только оценен. Разработка *теории торгов* еще только началась, но уже имеются некоторые полезные средства для решения этого рода задач.

Комбинированные процессы. Реальные системы очень редко характеризуются только одним из вышерассмотренных процессов. Так, например, задачи по управлению производством обычно включают сочетание процессов создания запасов, распределения и обслуживания. Или же, если взять задачу замены отдельных элементов, которые выходят из строя, то она, как правило, связана с проблемой запасов, а задача торгов может оказаться связанной с задачей распределения ресурсов между несколькими возможными контрактами, на которые могут быть сделаны ставки.

Обычная процедура исследования комбинированных процессов состоит в последовательном решении задач. Мы знаем, что во многих случаях не удается достичь оптимального решения даже при последовательной, повторяющейся корректировке. Отсюда ясно, что перед исследователем операций со все возрастающей необходимостью встает проблема отыскать сочетание абстрактных процессов и построить модели, основанные на взаимодействии нескольких рассматриваемых здесь процессов. Удовлетворение этой потребности все больше и больше привлекает внимание научной мысли.

Следует отметить также, что шесть видов рассмотренных здесь процессов не исчерпывают всех ситуаций, с которыми приходится сталкиваться при решении операционных задач. Но они охватывают большинство ситуаций, с которыми до сих пор приходилось иметь дело в практике исследования операций. Однако можно ожидать, что в будущем обнаружится и будет подвергнуто математическому анализу гораздо большее число встречающихся ситуаций.

Названия абстрактных моделей не должны сковывать воображения читателя. Модели запасов применимы к задачам наличных денег, оборотного капитала и рабочей силы. Модели очередей могут быть применены к решению некоторых задач управления запасами. Воображение является столь же эффективным средством успешного научного анализа, как и любая другая способность разума. Читатель проработает эту книгу с тем большей для себя пользой, чем менее предвзято он к ней подойдет и чем больше ассоциаций она у него вызовет.

КРАТКОЕ ЗАКЛЮЧЕНИЕ

Было показано, что исследование операций появилось как самостоятельная дисциплина в связи с развитием организаций, в которых административные функции разделены по видам и уровням руководства. Необходимость научного изучения задач организационного характера, задач, которые связаны с взаимодействием функциональных подразделений организаций, и возможность, предоставленная ученым военной администрацией во время второй мировой войны для разработки новых методов подхода к решению подобных задач, привели к возникновению исследования операций как самостоятельной области знаний.

Исследование операций, возможно, еще слишком молодая область, чтобы ей можно было дать законченное классическое определение. Тем не менее нами предложено следующее рабочее определение: *исследование операций состоит в применении научных принципов, методов и средств к задачам, связанным с функционированием систем, с целью предоставить тем, кто отвечает за управление данными системами, оптимальные решения поставленных задач*. В этой книге основное внимание уделяется системам «человек — машина» в промышленных организациях.

Было показано, что использование операционных групп (члены которых являются специалистами в различных областях) дает возможность применения различных принципов, методов и средств. Исследование операций положило начало разработке подхода, который обеспечивает наиболее эффективное решение организационно-управленческих задач. Процедуру этого подхода можно разбить на следующие этапы:

1. Постановка задачи.
2. Построение математической модели рассматриваемой системы.
3. Нахождение решения с помощью модели.
4. Проверка модели и полученного с ее помощью решения.
5. Построение процедуры подстройки решения.
6. Осуществление решения.

Хотя комплексные операционные группы дают возможность применять к решению конкретных задач разнообразные существующие методы и средства, были разработаны новые методы и средства, а прежние были видоизменены применительно к типичным классам задач, связанным со следующими пятью процессами: управление запасами, распределение, обслуживание, замена оборудования и состязательные процессы. Каждый из этапов исследования операций и каждый класс задач будет более подробно рассмотрен в последующих главах.

Из печати вышел ряд статей общего характера по исследованию операций. Большинство из тех работ, которые появились до 1954 года, приведены в прекрасной библиографии, опубликованной в «Operations Research for Management». Некоторые из более поздних статей подобного типа даются в библиографии в конце главы.

Теперь мы перейдем к рассмотрению некоторых примеров для иллюстрации отдельных важных характерных черт исследования операций, которые подверглись обсуждению в этой главе.

ЛИТЕРАТУРА

1. Camp Glen D., The Science of Generalized Strategies and Tactics, Textile Res. J., XXV, no. 7, 629—634 (July 1955).
2. Herrmann Cyril C. and Magee John F., Operations Research for Management, Harv. Busin. Rev., 31, no. 4, 100—112 (July-Aug. 1953).
3. Hurni M. L., Observations on Operations Research, J. Opns. Res. Soc. Am., 2, no. 3, 234—248 (Aug. 1954).
4. Hurni M. L., The Purpose of Operations Research and Synthesis in Modern Business, Management Consultation Service, General Electric Co., New York, June 24, 1955.
5. Johnson Ellis A., Operations Research in Industry, Proceedings of Operations Research Conference, Society for Advancement of Management, New York, 1954.
6. Smiddy Harold F. and Naum Lionel, Evolution of a Science of Managing in America, Mgmt. Sci., 1, 1—31 (Oct. 1954).
7. Solow Herbert, Operations Research in Business, Fortune, LIII, no. 2, 128 ff. (Feb. 1956).
8. Trefethen Florence N., A History of Operations Research, in Joseph F. McCloskey and Florence N. Trefethen (eds.), Operations Research for Management, The Johns Hopkins Press, Baltimore, 1954.

Глава 2

ОПЕРАЦИОННОЕ ИССЛЕДОВАНИЕ СИСТЕМЫ КАК ЕДИНОГО ЦЕЛОГО

ВВЕДЕНИЕ

Ниже приводится пример, при изложении которого основное внимание уделяется взаимодействию различных сторон производственного процесса и роли этих связей при исследовании, направленном на решение организационных задач.

В принципе большинство руководителей и исследователей вряд ли стали бы отрицать необходимость системного подхода к задаче. Однако, к сожалению, существует печальное расхождение между принципами и практикой. Постоянная занятость административного работника, как правило, не дает ему возможности проанализировать все варианты своих решений и оценить их возможные последствия. И даже тогда, когда позволяет время, он редко может систематически применить определенный метод, чтобы убедиться, что он изучил действительно все, что связано с предлагаемым решением задачи.

Принципы исследования, рассматриваемые в этой книге (особенно в частях II, III и IX), предназначены для обеспечения глобального решения организационных задач. При изложении данного примера, который служит в основном целям иллюстрации системного подхода в его самом общем виде, отдельные частности этих принципов не рассматриваются.

Изложение примера тем не менее преследует и другие цели. При детальном рассмотрении принципов, методов и средств в дальнейшем очень удобно иметь наглядный пример, на который можно ссылаться. А поэтому разбор примера проводится несколько более подробно, чем если бы имелась в виду только задача проиллюстрировать системный подход. Многие подробности приведены для того, чтобы служить материалом для ссылок при последующем обсуждении.

ХАРАКТЕРИСТИКА ФИРМЫ И ОПЕРАЦИОННОЙ ГРУППЫ

Начнем с общего описания фирмы, в которой проводились операционные исследования. Фирма известна главным образом как производитель металло-режущих станков. Эта фирма, самая крупная в мире по производству станков одного типа, выпускает более 50% всей продукции страны. Фирма производит несколько других типов станков, часть из которых связана, а часть не связана с ее основной продукцией. Продажная цена различных моделей основной продукции составляет от 10 тыс. до 40 тыс. долларов. Во время проведения исследования годовой выпуск продукции фирмы составил сумму в 50 млн. долларов. Фирма имеет два предприятия, на которых работает 3500 чел. Занятость рабочих на фирме была наибольшей во время войны, но и в период проведения операционных исследований (1952—1953 гг.) фирма почти достигала своей максимальной производственной мощности.

На первом совместном заседании администрации фирмы и сотрудников операционной группы Кейсовского технологического института представители администрации заявили, что их прежде всего интересует возможность применения операционных методов для анализа деятельности фирмы. Хотя они не хотели конкретно указать задачу, с решения которой можно было бы начать, одна задача, в частности, их особенно волновала. Здесь нет необходимости детально анализировать эту задачу, она будет рассмотрена только в связи с причинами, по которым она не была избрана операционной группой. Кратко эту задачу можно сформулировать следующим образом.

Уровень основного производства станков был относительно постоянным, поскольку существовал большой портфель невыполненных заказов. Предприятия работали фактически на полную производственную мощность. Поступали дополнительные заказы, но в меньшем количестве, чем производственные возможности фирмы. Таким образом, портфель заказов сокращался. При сохранении спроса на том же уровне фирма примерно через год выполнила бы все заказы. А еще через год, если бы фирма продолжала выпускать основную продукцию в таком же количестве, возникло бы перепроизводство. Одно из второстепенных изделий, не связанных с основным видом продукции, выпускалось на значительно более низком уровне, и сбыт его сдерживался, с тем чтобы не принимать заказы на большее количество, чем фирма производила в то время. Вместе с тем руководители фирмы понимали, что сбыт этой второстепенной продукции можно увеличить, проявив больше инициативы в поисках рынка. Поставленная задача заключалась в следующем: когда фирма должна начать сокращать основное производство станков и увеличивать производство своей второстепенной продукции и какими темпами должно проводиться и то и другое? Вопрос администрации фактически сводился к тому, как найти оптимальное соотношение между двумя параллельными видами производства и стратегией сбыта.

Операционная группа попросила, чтобы выбор задачи был отложен по следующей причине. Поставленный вопрос предполагал, что единственный способ увеличить производство второстепенного вида продукции заключается в сокращении основного производства станков. Другими словами, этот вопрос подразумевал, что изготовление станков осуществляется с максимальной производительностью или по крайней мере что производственная мощность не может быть эффективно увеличена. Представлялось нецелесообразным начинать работу группы в фирме, основываясь на таком сильном допущении. Представлялось также, что для того, чтобы убедиться в правильности этого предположения, членам группы необходимо знать о деятельности фирмы гораздо больше, чем они знали. Поэтому представители

администрации и члены операционной группы пришли к заключению, что должно быть отведено несколько недель на ознакомление группы с делами фирмы с тем, чтобы в процессе предварительного изучения работы фирмы группа и определила ту задачу, с которой она могла бы начать исследование.

Была создана группа из трех человек, в которую вошли два представителя из операционной группы факультета организации производства при Кейсовском институте и один представитель фирмы. Представитель фирмы был специалистом по исследованию финансовых операций. Он состоял при главном бухгалтере фирмы в качестве консультанта по анализу финансовых проблем, и по роду своей деятельности ему часто приходилось иметь дело со многими сложными вопросами, связанными с различными сторонами работы фирмы. Хотя прежде он не был знаком с операционными исследованиями, его большой опыт и интерес к принципам решения задач делали его идеальным членом группы.

Для большего удобства работы группа разместилась в отделе главного бухгалтера, хотя подчинялась она непосредственно вице-президенту фирмы. Главный бухгалтер помогал в установлении контактов с работниками фирмы.

В процессе проведения исследования количество членов группы менялось. На некоторых этапах работы ее численность доходила до четырех представителей института, нескольких аспирантов и девяти сотрудников фирмы. С всем протяжении исследования проводились регулярные консультации с другими сотрудниками операционной группы при Кейсовском институте, преподавателями и самыми различными работниками фирмы.

ПЕРИОД ОЗНАКОМЛЕНИЯ

Как было указано выше, начальный этап исследования состоял в предварительном знакомстве с фирмой. Прежде всего было организовано подробное обследование основного предприятия и административных отделов. Группа попросила разрешения ознакомиться со структурой фирмы и запросила схему ее организации. Фирма не питала особого пристрастия к схемам, и поэтому эта просьба натолкнулась на некоторые трудности. Но после того, как схема была получена и были выяснены некоторые связанные с ней вопросы, оказалось, что между схемой и фактической системой руководства деятельностью фирмы наблюдаются значительные расхождения. Группе необходимо было прежде всего знать характер операций, которыми занималась фирма, и, во-вторых, как организовано и как осуществляется управление этими операциями.

Поэтому группа решила изучать фирму как организованную систему связи, которая управляет процессом производства (в главе 4 мы определим основные положения, которые могут лечь в основу подобного исследования). Каков же в таком случае основной источник информации, которая идет по каналам системы? Очевидно, это заказчик, потребитель продукции. Как же фирма получает информацию относительно запросов потребителя? Через своих агентов по сбыту. Таким образом, свое знакомство с работой фирмы группа начала с отдела сбыта. Она выяснила, как агенты по сбыту отбирают потенциальных заказчиков, в какого вида контакты с потребителями они вступают, как они отчитываются в своей деятельности, как выполняются заказы и определяются прогнозы на будущее и т. д. Затем был изучен процесс переработки поступающей информации в различных секциях этого отдела. Далее группа выяснила, как эта информация преобразуется и поступает в систему производства, как она вызывает поступление сырья и как

в итоге выпускается конечная продукция, направляемая заказчику. Через две недели были собраны горы данных и кипы документации. Несколько дней ушло на определение существа этих сложных процессов и изображение их в виде рабочей блок-схемы управления и потока материалов (см. рис. 2.1). Для целей данного описания нет необходимости объяснять эту схему во всех деталях, но, вероятно, целесообразно рассмотреть одну часть цепи, чтобы наглядно показать, насколько подобная схема облегчает понимание процесса прохождения информации внутри фирмы.

Сначала рассмотрим словесное описание одного из участков потока информации. Отдел планирования производства получает от диспетчерской группы график сборки на каждый месяц. Этот график определяет количество и модели станков, которые надлежит собрать в каждый из последующих пяти месяцев. Для каждой намеченной к сборке модели в отделе планирования производства имеется полный перечень необходимых деталей. Далее, для каждой детали в том же отделе ведутся карты учета, в которых указано, сколько соответствующих деталей имеется на складе, в производстве или заказано на других предприятиях. Для любой детали существует четыре возможные ситуации:

1. Деталь производится фирмой и находится на складе или в производстве.

2. Деталь производится фирмой, но ее нет ни на складе, ни в производстве.

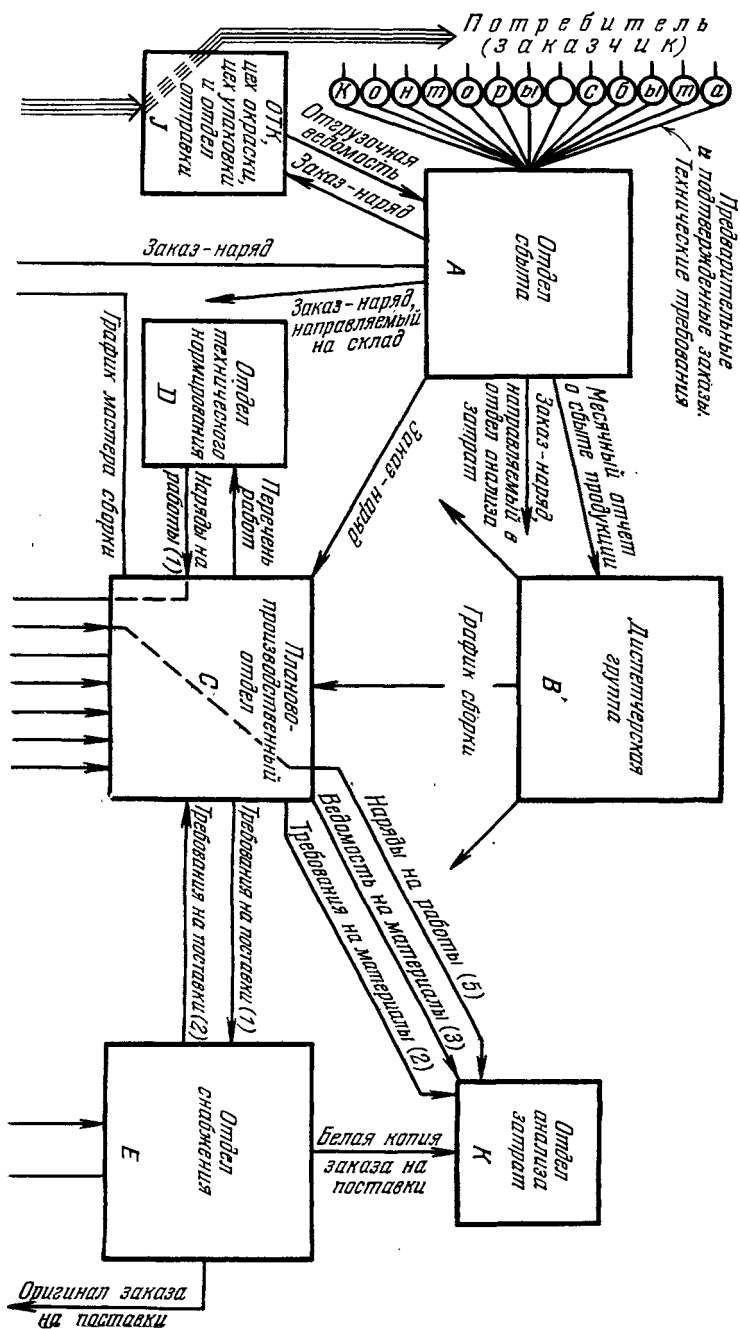
3. Деталь закупается и находится на складе, или на нее сделан заказ.

4. Деталь закупается, но ее нет на складе и на нее не оформлен заказ.

Рассмотрим здесь только одну, четвертую ситуацию. Для каждой модели станка, указанной в графике, отдел планирования подготавливает список деталей, которых нет на складе. Этот список называется требованием на заказ. Требование направляется в отдел снабжения, который подготавливает семь экземпляров заказа на каждую требуемую деталь и возвращает наряд в плановый отдел как уведомление о том, что заказ принят и размещен. Эта информация переносится на карты, где отмечается наличие соответствующих деталей. Оригинал заказа направляется поставщику. Одна копия поступает в отдел анализа затрат, который впоследствии использует эту информацию при определении себестоимости единицы продукции. Три копии направляются в отдел приемки. Оставшиеся две копии хранятся в отделе снабжения для облегчения проверки в случае задержки поставок.

Когда отдел приемки получает детали, заказанные у поставщика, он возвращает одну копию заказа в отдел снабжения. Отдел снабжения помещает все свои копии заказа в архив: его дело сделано. Отдел приемки пересылает детали вместе с двумя оставшимися копиями заказа на склад. Работник склада отмечает получение заказа на одной из копий и отправляет ее обратно в отдел поставок, работа которого теперь также закончена и зафиксирована. Когда детали поступают на хранение, последняя копия заказа направляется со склада в плановый отдел. В этом отделе отмечается поступление деталей на картах учета имеющихся на складе деталей. Цепь, таким образом, замкнулась.

Обратите внимание, как тот же процесс изображен на рис. 2.1,— следить за ним легко и просто. На схеме также показаны цепи для каждой из остальных возможных ситуаций вместе с другими стадиями процесса. Этот анализ потока информации и управления постоянно использовался не только операционной группой: представители администрации также нашли эту схему очень полезной и прибегали к ней при обсуждении организации фирмы и при ознакомлении новых служащих с ее работой.



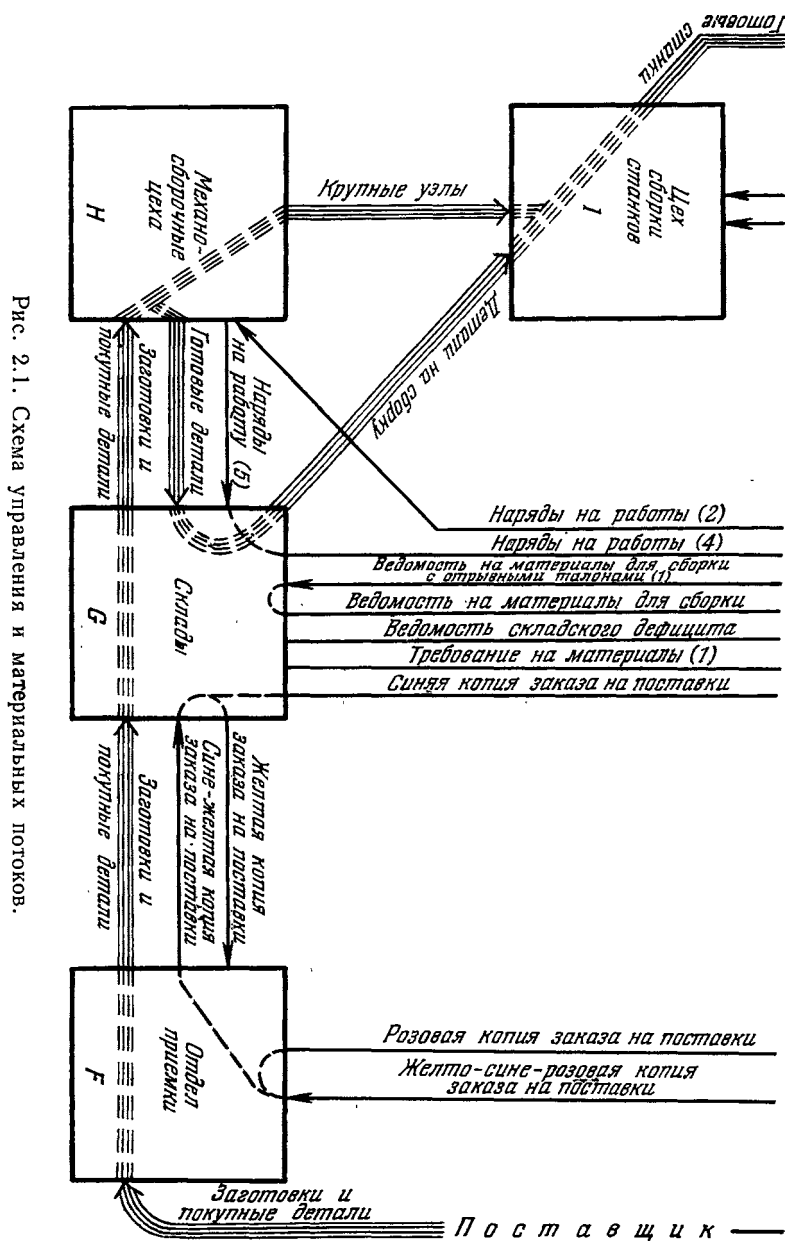


Рис. 2.1. Схема управления и материальных потоков.

В процессе сбора информации, необходимой для только что приведенного вида анализа системы, операционная группа пришла к выводу, что имеется задача, которая является общей для всех отделов. И, пожалуй, неудивительно, что такой задачей оказалась задача управления запасами. Некоторые высказывали мнение, что запасы фирмы слишком велики. Другие считали, что они слишком малы. Но так или иначе все думали о запасах.

Операционная группа получила документы по учету запасов, проведенному в конце предыдущего года, и для удобства пользования представила эти данные в виде таблицы 2.1. Вертикальные столбцы означают виды запасов, а горизонтальные строки — тип продукции, для которой предназначаются эти запасы. Стоимость запасов в каждой клетке была переведена в проценты (приблизительно) по отношению к общей стоимости всех запасов, которая составляла 11 млн. долларов.

Таблица 2.1 выявила ряд данных, которые в общем уже были известны, например то, что 65% запасов относилось к основной продукции. Она также помогла обнаружить некоторые не столь явные факты: 29% всех запасов приходилось на детали, готовые или находящиеся в производстве. На этом

Таблица 2.1

Распределение запасов в процентах

Вид изделий	Сырье	Закупаемые детали	Готовые крупные узлы	Готовые подузлы	Готовые детали	Узлы в производстве	Детали в производстве	Прочие	Всего
A *	0,4	9,9		3,3	16,8	18,1	12,0	5,0	65,5
B	0,1	0,0+			0,0+		0,2	0,0+	0,3
C	0,1	0,6		0,4	2,1	2,5	1,0	0,2	6,9
D **	0,1	0,6		0,0+	2,6	2,5	2,6	0,2	8,6
E								0,2	0,2
F	0,0+	0,4	1,3	0,1	3,1	1,6	0,3	4,8	11,6
G		2,7		0,0+	0,2	0,4	0,2	3,4	6,9
Всего	0,7	14,2	1,3	3,8	24,8	25,1	16,3	13,8	100,0

* Основной станок.
 ** Второстепенный продукт.

основании, а также в связи с тем, что проблема запасов оказалась удобным средством проникнуть в деятельность фирмы, операционная группа рекомендовала в качестве первой задачи изучение запасов деталей.

Группа вновь встретилась с представителями администрации, показала им схему управления и потока материалов на предприятии, обсудила ее вместе с ними и предложила поставить задачу изучения запасов деталей. Представители администрации приняли это предложение, и группе была предоставлена «свобода действий».

Планирование производства деталей. Исследование началось с опроса работников фирмы относительно существа задачи управления запасами деталей. Обычная формулировка ответа была примерно следующей: задача заключается в определении минимального запаса деталей, необходимого для сохранения существующего темпа сборки и отгрузки. Группа не была удовлетворена такой постановкой задачи, так как это означало, что уровень

прибылей от сбыта продукции примерно постоянный или что если даже он колеблется, то эти колебания не отражаются существенно на уровне запасов. Если, как считала группа, объем запасов связан с издержками производства, то тогда, по-видимому, он должен определяться не как наименьшая величина, необходимая для сохранения данного объема сбываемой продукции, а как величина, которая может быть использована для получения возможно большей прибыли при данном объеме сбыта. Такое рассуждение заставило группу переформулировать задачу как задачу планирования производства деталей таким образом, чтобы минимизировать общие производственные затраты, в том числе и расходы, связанные с запасами.

Что нужно для производства отдельной детали? Прежде всего, сырье, стоимость которого определяется закупочными ценами и расходами на транспорт. Затем сырье хранится, и на это идут дополнительные расходы. После этого наступает очередь планирования, при котором определяется будущая судьба материалов. Планирование, осуществляемое штатом служащих, также требует расходов. Далее, цеха должны быть подготовлены для изготовления детали. В последующем материал подвергается обработке и ожидает перехода от одной операции к другой. Наконец, наступает период хранения готовых деталей и оформления необходимой документации. На основании предварительного исследования группа пришла к выводу, что изменения в планировании производства не могут существенно повлиять на запасы сырья и полуфабрикатов. Это положение было принято для упрощения задачи, а затем проверено. Но об этом позже.

Приведенное нами нестрогое описание процесса изготовления детали требует уточнения, которое было достигнуто при изучении оперативного планирования производства деталей и при определении и описании существенных переменных этого процесса. Графики производства деталей составлялись ежемесячно (другими словами, плановый период составлял один месяц). Однако не все детали производились каждый месяц. Из общего количества примерно в 18 тыс. различных видов деталей, производимых на предприятии, только одна треть (6 тыс.) производилась каждый месяц.

Удобно было найти некоторый интервал времени, относительно которого можно было бы подсчитать затраты. Для расчетов был выбран годовой период. Математическая модель процесса производства, которая была впоследствии разработана, является общей в том смысле, что период для расчета затрат может быть установлен любой продолжительности. Модель и процесс ее построения будут рассмотрены в главе 7. В модели учтены три существенных вида затрат:

1. Наладка оборудования для выпуска каждой партии (переменные затраты на изготовление каждой детали).
2. Расходы на сырье и на обработку каждой детали (постоянные * затраты на каждую деталь).
3. Затраты на хранение, выраженные в процентах от общей стоимости детали из расчета месячного хранения.

Смысл по крайней мере некоторых из этих расходов далеко не очевиден, так что рассмотрим их подробнее.

Рассмотрим сначала расходы на подготовительно-заключительные операции для каждой партии. Термин «партия» относится ко всем деталям, которые изготавливаются при одной наладке оборудования, применяемого

* Эти затраты не являются постоянными в абсолютном смысле, но их колебания чрезвычайно малы по сравнению с затратами, которые мы называли переменными. Как будет показано ниже, некоторые расходы на сырье фактически менялись.

при их изготовлении. Размеры партии могут меняться. Другими словами, размер партии определяется числом деталей, изготовленных в результате непрерывной последовательности операций. Затраты на подготовительно-заключительные операции включают четыре основных компонента.

1. Подготовка на уровне администрации. Прежде чем произвести какую-либо работу в цехе, плановый отдел должен спланировать производство, а технологический отдел должен подготовить необходимые чертежи и документацию.

2. Подготовка на уровне цехов. Эти расходы включают конкретную наладку оборудования, предназначенного для изготовления данных деталей или производства соответствующих операций, стоимость брака, который неизбежен при проведении наладки для запуска партии, и затраты на организацию контроля качества.

3. Окончание работ на уровне цехов. С этим связаны расходы по размещению изготовленных деталей на складе и оформлению необходимой документации.

4. Окончание работ на уровне администрации. Сюда относятся расходы, связанные с расчетами, производимыми отделом анализа затрат.

Естественно, что работа по определению стоимости подготовительно-заключительных операций по производству отдельной детали не относится к числу легких. В данном случае она потребовала кропотливого труда и тесного контакта с рядом отделов. Однако она оказалась весьма полезной, ибо в результате был поставлен очень важный вопрос. В системе бухгалтерского учета не определялась стоимость каждой детали. Не следует ли бухгалтерии это делать? Новый финансовый ревизор фирмы использовал этот вопрос для подкрепления своего плана превратить бухгалтерский учет из пассивной регистрации данных за отчетный период в процесс, дающий действенные данные для руководства. Работа операционной группы еще раз привлекла внимание к необходимости функционального, или оперативного, ведения бухгалтерского учета. Работа группы отчасти способствовала поддержке идеи ревизора, ибо было показано, как можно использовать регрессионный анализ для разделения постоянных и переменных затрат и как можно применить методы статистического контроля качества для непрерывного регулирования этих расходов.

Разработка методов регулирования или по крайней мере методов определения изменений средних значений переменных затрат очень существенна для любого операционного проекта. Любое принятое решение может считаться действительным только в том случае, если использовались реальные данные по затратам. Но расходы меняются. Поэтому должна быть разработана процедура для поддержания постоянства средних расходов и для столь быстрого обнаружения изменений, чтобы в решение можно было внести соответствующую коррективу. Обычно это можно выполнить, используя принципы статистического контроля качества.

Затем были практически определены расходы на подготовительно-заключительные операции по выпуску партии каждой отдельной детали путем изучения среднего расходуемого времени, выраженного через затраты в каждом из соответствующих отделов. Результаты этого исследования сведены в таблицу 2.2. Нетрудно видеть, что если среднее число партий в месяц уменьшается, то наблюдается уменьшение в расходах по планированию производства. Отсюда, как указывалось раньше, задача заключается не только в том, чтобы уменьшить объем запасов по отношению к объему сбыта, но и в том, чтобы минимизировать расходы по планированию производства. Эти расходы могут быть снижены за счет увеличения числа различных

Таблица 2.2

Затраты на зарплату, синьки и бумагу

Зарплата	
Работникам планово-производственного отдела	0,87
Работникам отдела нормирования	0,20
Работникам отдела анализа затрат	0,38
Работникам склада	0,10
Синьки	0,10
Бумага	0,05
Итого	\$ 1,70

деталей, производимых в одной партии; тем самым снижается число партий в год и увеличиваются запасы.

Теперь перед нами встает задача баланса: один фактор (затраты на вложения в запасы) должен уравнивать другой (расходы на подготовительно-заключительные операции). Это типичный случай организационной задачи, который рассматривался в предыдущей главе. Предваряя выводы, скажем, что, как было установлено, расходы на подготовительно-заключительные операции оказались весьма существенными и что правильное их соотношение с расходами на хранение может потенциально дать фирме годовую экономию в затратах на планирование производства деталей примерно в 40 тыс. долларов. Это не будет экономией прямых затрат, так как не подразумевает сокращения работников по планированию. Экономия достигается за счет того, что выигрывается время, которое может быть затрачено на выполнение других работ, прежде не выполнимых из-за недостатка штатов и помещения.

Следующая важная статья затрат представляет собой комбинацию расходов двух видов: расхода на сырье и на производственные операции, которые вначале рассматривались отдельно. Выяснилось, что удобнее их объединить, поскольку и те и другие являются постоянными расходами на деталь. Первый компонент составляют расходы на сырье, необходимое для изготовления детали. Второй компонент — расходы на труд, вкладываемый в работу над каждой деталью, плюс накладные расходы. Накладные расходы, которые входят как в расходы на подготовительно-заключительные операции, так и в расходы на сырье и производство, было нелегко определить и распределить. Удовлетворительная оценка этих расходов была получена при их выражении как функции человеко-часов непосредственного труда, затраченного на выполнение операций.

Третья статья затрат — это расходы на хранение. Группа провела изучение расходов, связанных с работой одного из складов фирмы. Учитывались арендная плата, оплата за отопление и освещение, расходы на складское хозяйство, содержание административно-управленческого аппарата, зарплата складских рабочих и потери при хранении. К этим расходам были добавлены расходы на заем капитала, вложенного в запасы. Это дало цифру немногим большую, чем 1% в месяц на каждый доллар, вложенный в запасы. Для большей безопасности, наряду с этой цифрой в 1%, использовалась и пессимистическая оценка в 2%. Результат такой поправки будет рассмотрен ниже.

Общие накладные расходы и затраты на хранение рассматривались фактически как расходы, меняющиеся прямо пропорционально изменению количества выпущенных деталей. Но в этих расходах есть элементы, которые

оказываются постоянными при определенных масштабах или уровнях производства. Было установлено, однако, что результаты, полученные в конечном итоге, не отличались сколько-нибудь значительно при различных вариантах распределения этих затрат. Вследствие этого применялся наиболее простой способ распределения этих расходов (накладных расходов и расходов на хранение), а именно по времени (на час непосредственно затрачиваемого труда) и по вложенным долларам соответственно.

Уравнение планирования. Была разработана модель производственного процесса, в которой общегодовые затраты на производство каждой детали были выражены как функция от размера партии и, следовательно, от числа партий равного размера для данной детали из расчета на год. В уравнение входят четыре основных элемента затрат: (1) расходы на сырье, (2) расходы на производство, (3) расходы на хранение полуфабрикатов, (4) расходы на хранение готовых деталей. Из-за краткости производственного цикла расходы на хранение полуфабрикатов составили лишь небольшую долю процента общих расходов, в связи с чем эти расходы не были включены в уравнение, так что наиболее существенным оказалось годовое приращение общих производственных затрат. Задача заключается в установлении такого размера партии для каждой детали, при котором минимизируются эти годовые затраты. С помощью аналитического аппарата, подробно описанного в главе 7, было получено следующее уравнение для определения оптимального размера партии:

$$R = \frac{L}{\sqrt{\frac{L P c_2}{c_1 (2 - P)}}}.$$

В этом уравнении R — оптимальное количество деталей в партии, L — число деталей, требуемых в месяц, P — расходы на хранение готовой продукции, выраженные в процентах к количеству капитала, вложенного в одну деталь, c_1 — расходы на подготовительно-заключительные операции для одной партии, c_2 — расходы на сырье и изготовление одной детали.

Практическая проверка. После того как было выведено это уравнение, группа еще раз встретилась с представителями администрации фирмы. Детально математические выкладки не обсуждались, но рассматривались заложенные в них идеи. Эта встреча прояснила многие вопросы, связанные с определением затрат. Представители администрации пришли к выводу, что было бы неплохо испытать модель. Плановый отдел предложил, чтобы уравнение, определяющее размеры партии, было применено первоначально к 23 деталям, которые они сами выберут. Никакого определенного принципа при выборе этих деталей не применялось. Они были отобраны просто потому, что их производство представляло трудности при планировании.

После того как детали были отобраны, группа подсчитала общий годовой прирост затрат на производство отдельной детали, расход времени на подготовку производства при существующем графике, а также вычислила общий годовой прирост затрат и время на подготовку производства при оптимальном размере партий. Результаты показали необыкновенно большие потенциальные возможности сокращения как расходов, так и времени производства. Но для того, чтобы добиться этого сокращения, необходимо было более чем вдвое увеличить запасы готовых деталей. Иными словами, увеличение размера партий, а следовательно и увеличение запасов готовой продукции, дает существенную экономию средств и времени.

Следующее заседание с представителями администрации привело к обоюдному выводу, что полученные результаты указывают на необходимость более систематического изучения более представительной выборки деталей. Для этой цели был взят узел, состоящий из 112 деталей. Было проведено исследование для получения оценки чувствительности производственных затрат к размеру партий. Полученные результаты показали, что оптимальное планирование (по сравнению с фактическим) сократило бы прирост производственных затрат на 3,5%, а время подготовки производства — на 70%. Администрация сочла эти выводы достаточно серьезными, чтобы организовать дальнейшее исследование с целью определения тех мероприятий, которые потребовались бы для реализации этой потенциальной экономии затрат и времени.

Некоторым, возможно, показалось бы, что на этом и должна бы закончиться задача исследования. Но фактически это было в полном смысле лишь самое начало программы выполнения операционного проекта. Наиболее сложные аспекты задачи исследования возникли лишь тогда, когда вся работа сосредоточилась на практическом осуществлении производства при оптимальных размерах партий. И действительно, из всех упоминавшихся эта задача наиболее близка типичной организационной задаче.

На этом этапе работы операционная группа столкнулась с рядом вопросов, ответы на которые позволили бы найти общее решение задач минимизации затрат на производство деталей. Вопросы эти сводились к следующему.

1. Какая (или в каком количестве) необходима информация для снижения производственных затрат при использовании выведенного уравнения оптимального размера партии и каково будет влияние погрешностей в этой информации?

2. Какие дополнительные изменения необходимо провести в существующей процедуре планирования, чтобы применить уравнение планирования?

3. Могут ли какие-либо из операций, на которые влияет объем производства, быть изменены так, чтобы при использовании указанного уравнения увеличить их производительность?

4. Какие условия, принятые при формулировании предлагаемой процедуры планирования, имеют тенденцию к изменениям и как должна меняться процедура планирования, если такие изменения происходят?

Благодаря подробной информации о функционировании системы (которая была накоплена за период общего ознакомления) эти общие вопросы получили более конкретную формулировку, которую можно представить следующим образом.

1. Какой процент ошибки и какого типа ошибки в оценке расходов повлечет за собой увеличение производственных расходов, если применять уравнение планирования?

2. Откуда можно брать капитал для увеличения запасов готовой продукции и как это скажется на кредитах фирмы?

3. Как может быть осуществлен переход к партиям большего размера без дефицита деталей во время переходного периода?

4. Как можно обеспечить полное использование производственных мощностей при ежемесячном выпуске меньшего числа больших по размеру партий?

5. Сколько потребуется дополнительного помещения для хранения и как его можно обеспечить?

6. Как изготовление и обработка деталей по заказам на запчасти могут быть наиболее эффективно совмещены с предлагаемой процедурой планирования производства деталей для сборки?

7. Как может быть улучшен процесс заготовки сырья в свете предлагаемых изменений в производстве деталей?

8. Что можно сделать, чтобы свести к минимуму устаревание деталей, которые будут храниться в больших количествах?

9. Может ли планирование сборки быть улучшено с учетом предлагаемых изменений в производстве деталей?

10. Как можно процедуру планирования производства приспособить к ситуации (которая рано или поздно наступит), когда спрос не будет извештен и постоянен, а будет лишь предсказуем и переменен?

11. Как должно производиться фактическое определение размера партии каждого вида деталей в отдельности?

Прежде чем рассмотреть, как ответить на поставленные, а также и связанные с ними вопросы, следует остановиться особо на одной стороне результатов проведенного на данном этапе исследования.

Как читатель, по-видимому, помнит, задача, поставленная на самой первой встрече с представителями администрации, требовала увеличения производства второстепенной продукции и сокращения производства основного станка. Результаты первого этапа исследования показали потенциальную возможность экономии времени производства станков примерно на 150 тыс. человеко-часов в год. Этого времени, если его правильно использовать, было бы достаточно для того, чтобы добиться желаемого увеличения выпуска второстепенной продукции без изменения уровня производства основного станка. Таким образом, ответ на первоначально поставленные вопросы был получен при системном подходе к задаче.

Теперь перейдем ко второму этапу решения задачи.

ВЛИЯНИЕ ВОЗМОЖНЫХ ОШИБОК НА ОЖИДАЕМУЮ ЭКОНОМИЮ РАСХОДОВ

При расчете экономических размеров партий должны были быть оценены расходы на подготовку производства и непосредственно производственные затраты. Эти оценки были основаны на принятых нормативах, и, естественно, в них была допущена некоторая погрешность. Ошибку нельзя было достаточно точно определить. Вследствие этого необходимо было подойти к решению задачи с другого конца. Вопрос был поставлен так: могут ли оценки затрат быть настолько далеки от истинных, что существующая практика оказалась бы более совершенной, чем рекомендуемая стратегия, если бы истинные затраты были известны? Анализ показал, что тогда оценки как постоянных, так и переменных расходов должны были бы в среднем быть на 10% меньше истинной величины, и только в этом случае предлагаемая процедура будет более дорогостоящей, чем существующая. Однако ясно, что если бы затраты оценивались с такой ошибкой, то фирма давно бы уже обанкротилась (из-за отсутствия прибыли). Фирма могла быть, таким образом, уверенной, что ошибки в оценках этих двух величин не ведут к риску чрезмерных потерь.

А как обстоит дело с расходами на хранение? При подходе к этому вопросу был принят тот же путь. Анализ показал, что годовые расходы на хранение должны превысить по крайней мере 42% средней стоимости деталей, находящихся на складах, чтобы рекомендуемая процедура оказалась более дорогостоящей, чем существующая. Поскольку фирма уже указала, что ее наиболее пессимистические оценки расходов составили 24%, в этой части не оставалось никаких опасений.

ПОТРЕБНОСТИ В КАПИТАЛОВЛОЖЕНИЯХ И КРЕДИТ

Увеличение размера партий, а следовательно и объема запасов, требует дополнительных капиталовложений. В связи с этим возникли три вопроса.

1. Сколько денег потребуется на хранение дополнительных запасов и когда эти деньги могут потребоваться?

2. Какие затраты связаны с привлечением дополнительных капиталовложений?

3. Как увеличение капиталовложений в запасы может повлиять на кредит фирмы?

Представитель фирмы в группе был достаточно компетентен, чтобы дать квалифицированные ответы на эти вопросы. Была произведена оценка требуемой суммы денег исходя из предположения, что деловая активность фирмы остается на том же уровне. Исследование показало, что капиталовложения могут быть обеспечены при тех же затратах, что и прежде. Далее, изучение деятельности других компаний и порядка установления процента на кредит показало, что кредит фирмы вряд ли подвергнется какому-либо воздействию в связи с ожидаемым и необходимым увеличением дополнительно привлекаемого капитала.

На этом этапе исследования было относительно ясно, однако, что деловая активность не сможет сохраниться на том же уровне, а будет падать. Было установлено, что если активность будет падать так, как ожидается, то не потребуются никакого дополнительного капитала. Другими словами, выяснилось, что если произойдет предсказываемый спад и график производства будет построен на основе оптимальных размеров партий, то уровень запасов останется практически прежним. В обычной ситуации при спаде деловой активности произошло бы и соответствующее сокращение запасов.

ПОРЯДОК ПЕРЕХОДА НА ОПТИМАЛЬНЫЕ РАЗМЕРЫ ПАРТИЙ

Следующий важный комплекс задач требовал определения порядка перехода работы цехов на производство партий большего размера. Иными словами, необходимо было разработать план действий для преодоления трудностей переходного периода, которые возникали в связи с тем, что предприятие работало почти на максимальной мощности, производя ежемесячно 6 тыс. различных видов деталей в количестве, обеспечивающем трехмесячную потребность. Предлагаемая же процедура планирования требовала производства в среднем ежемесячно 2 тыс. типов деталей, покрывающих девятимесячную потребность. Но в любой месяц требуется производство 6 тыс. различных деталей. Поэтому, если выбрать для производства только 2 тыс. типов деталей, то снабжение остальными 4 тыс. деталей прекратится.

Возможны три подхода к решению задачи.

1. Привлечь дополнительную рабочую силу и оборудование и тем самым увеличить производственную мощность.

2. Заключить контракт на производство некоторых деталей с другими фирмами.

3. Преодолевать затруднения постепенно, рассчитывая на имеющиеся производственные мощности.

Все три возможности подверглись оценке. Прогнозы указывали на то, что спрос на рынке будет сокращаться и, следовательно, расширение производства с дополнительным привлечением капитала было бы не оправдано.

Убытки по контрактам с другими фирмами вызвали бы дополнительные издержки, которые бы дали во время переходного периода чистый убыток. Следовательно, все указывало на то, что наиболее реальным выходом может быть постепенный переход в пределах возможностей предприятия.

Это решение поднимало следующие вопросы.

1. Как должен проводиться переход?

2. Как долго он продлится?

Прежде всего, было решено расклассифицировать модели станков по степени возможного устаревания и начать переход с тех моделей, вероятность устаревания которых наименьшая. Во-вторых, должна была быть подсчитана потенциальная экономия издержек, связанная с переходом от существующей практики к планированию оптимального размера партий; детали также должны были быть расклассифицированы соответствующим образом. Эта классификация деталей определяла очередность, в соответствии с которой должен осуществляться переход к новой системе их производства. Для того чтобы определить объем дополнительной продукции, на которую можно дать задания цехам в любой данный месяц, было необходимо определить величину и характер машинного времени сверх того, которое требуется при обычном производстве. Параллельно и независимо от операционной группы такая работа проводилась в отделе планирования и была направлена на использование аппаратуры IBM для преобразования производственного графика в оценки загрузки оборудования. Поскольку соответствующая процедура в то время не была еще разработана (теперь она существует), необходимо было полагаться на оценку дополнительной загрузки оборудования, допустимой в любой данный месяц, которую рассчитывали работники самого отдела. Используя эту информацию, операционная группа оценила, что для полного перехода на производство по экономичным размерам партий потребуется от 2 до 3 лет.

ДЕФИЦИТ И ХРАНЕНИЕ

При исследовании путей перехода стало ясно, что нельзя составить производственный план, который был бы рассчитан на полное использование всего машинного времени. Некоторые участки будут простаивать в ожидании поступления продукции с других участков. Задача распределения загрузки оборудования усложнялась, далее, также тем, что в первый месяц после перехода на работу по новой системе возникали затруднения в связи с непредвиденным дефицитом. Необходимо было составить срочные задания по ликвидации этого дефицита, которые были спущены в цеха сверх нормальной нагрузки. Проблема дефицита была очень серьезной и вызвала в отделе планирования немало беспокойства. С помощью на первый взгляд не связанного с этой проблемой исследования, которое будет описано позднее, операционная группа способствовала сокращению числа срочных заданий, направляемых в цеха. При этом выяснилось, что проблема дефицита связана с проблемой дополнительных складских помещений.

Увеличение запасов потребовало увеличения складских помещений. Фирма была заинтересована в том, чтобы не арендовать дополнительные складские помещения. Поэтому было проведено предварительное обследование существовавшего порядка хранения готовой продукции. Стало очевидно, что дополнительные помещения для хранения могут быть получены путем изменения системы хранения или путем перепланировки складских помещений. Казалось, что это может создать серьезное препятствие на пути реализации нового плана.

В то время как группа изучала вопрос дефицита и хранения, нам задали вопрос, как можно спланировать производство запасных частей для удовлетворения заказов по ремонту оборудования. Новые детали требовались не только для сборки станков, но и для замены деталей, вышедших из строя. Фирма предпочитала принимать большие заказы на запасные части. Вошло в практику добавлять к обычному количеству деталей, предназначенному для сборки, дополнительное число деталей, для того чтобы обеспечить возможные заказы на запчасти, которые могут быть получены в период производства очередной партии. Решение относительно увеличения задания с учетом заказов на запчасти принималось на основании прошлой практики и здравого смысла.

ЗАПЧАСТИ ДЛЯ РЕМОНТА ОБОРУДОВАНИЯ

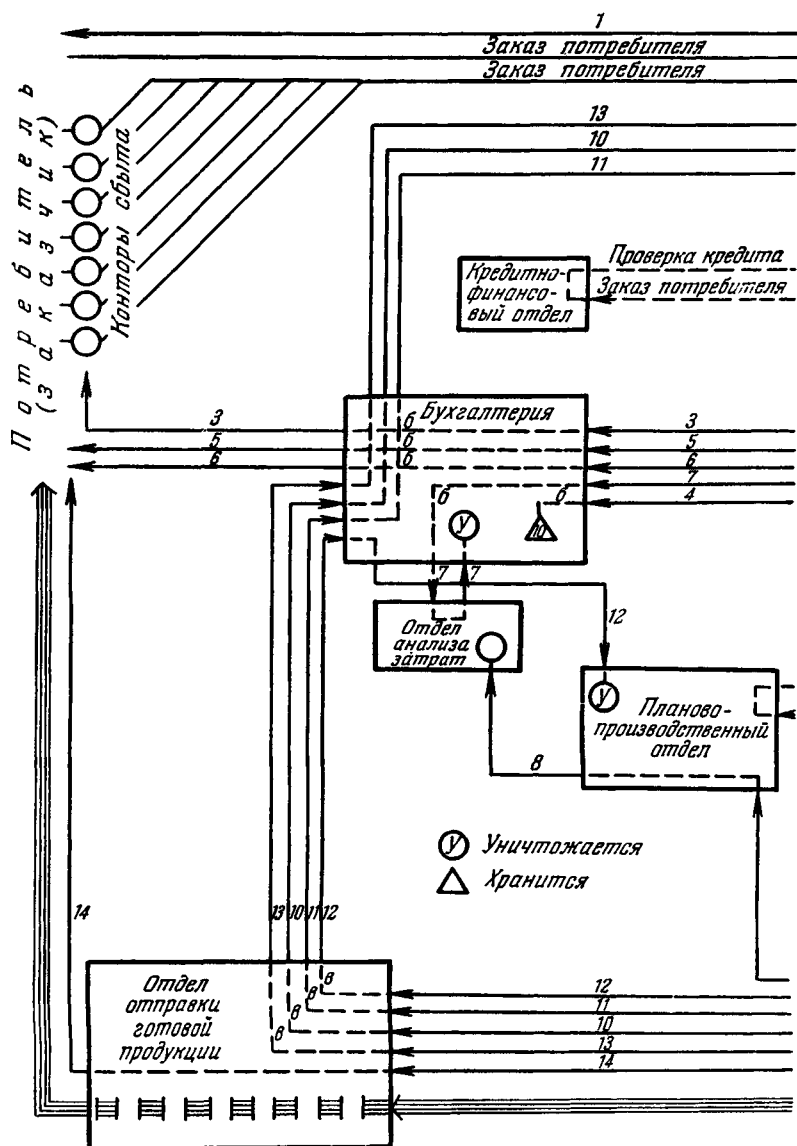
Группа провела исследование распределения заказов на запчасти по некоторой выборке деталей и вскрыла два примечательных фактора: (1) распределение спроса на отдельные детали сильно различалось, (2) колебание заказов по месяцам было также значительно.

Была разработана простая процедура определения сроков и объема заказов на запчасти. Прежде всего, детали классифицировались по степени их важности для работы станка. Класс 1 состоял из тех деталей, без которых станок не мог работать. В класс 2 вошли те детали, без которых станок может работать, но с низкими показателями. Класс 3 включал те детали, отсутствие которых может, хотя не обязательно, затруднить действия рабочего, но никак не ограничивает использование станка. С помощью ответственных за оформление заказов на запчасти был определен риск дефицита в соответствии с данной классификацией. Планирование деталей класса 1 (важных деталей) должно было осуществляться таким образом, что только в 27 из 10 тыс. намечаемых периодов можно было ожидать дефицита деталей данного класса. Для класса 2 риск был установлен в количестве 5 на 100 намечаемых периодов, а для класса 3 — в количестве 30 на 100 периодов. На основе этих приемлемых цифр и информации о времени производственного цикла соответствующей детали стало возможным определить для каждой детали уровень запасов, при котором необходимо размещать очередной заказ на запчасти.

Объем заказов определялся путем добавления к потребностям сборки по данной детали среднемесячного спроса на нее как на запчасть. Иными словами, не требовалось выпуска специальной партии запчастей, а они включались в общие цифры, отражающие потребности сборки. Эта процедура была проверена на выборке деталей. Было установлено, что сокращение расходов будет незначительным, но очень сложный процесс превращался в обычный, и работники планового отдела таким образом в значительной мере избавлялись от излишних хлопот.

Во время этого исследования, однако, становилось все яснее, что оформление документации и распределение материалов для производства запчастей никогда не увязывались с основным производственным процессом выпуска станков. Группа поэтому предложила провести дополнительное исследование выполнения заказов на запчасти, для того чтобы посмотреть, желательно ли такое объединение, и если желательно, то как оно может быть осуществлено.

Предложение было принято, и для выполнения этой работы состав группы был расширен. Кроме основных членов в группу был включен представитель каждого подразделения, связанного с этим процессом. Фирма



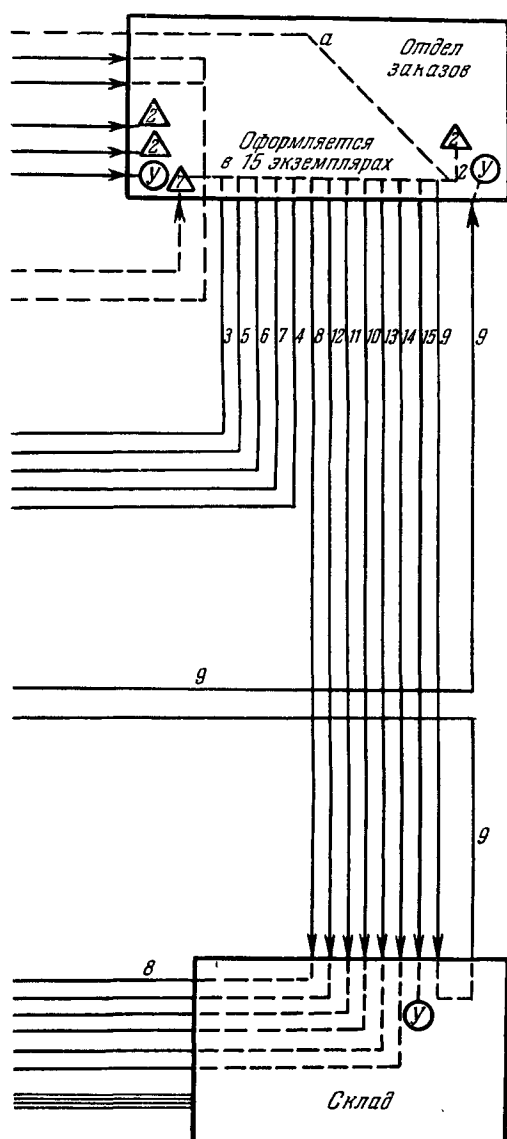


Рис. 2.2. Схема прохождения заказа на запчасти на предприятии станкостроительной фирмы для случая, когда все заказанные части имеются в наличии.

Существенно отметить, что ход выполнения заказа зависит прежде всего от кредитно-финансового отдела, а затем от конторы склада. Картотеки учета деталей ведутся в отделе заказов, планово-производственном отделе, на складе и в бухгалтерии. Схема построена на основании обследования на предприятии документооборота по заказу.

1 — подтверждение получения заказа, 2 — учетная ведомость отдела заказов, 3 — извещения конторам сбыта, 4 — учетная ведомость бухгалтерии, 5 — оригинал накладной, 6 — дубликат накладной, 7 — анализ затрат, 8 — требования на материалы, 9 — сопроводительная накладная, 10 — накладная отдела отправки, 11 — отгрузочная ведомость отдела заказов, 12 — отгрузочная ведомость планово-производственного отдела, 13 — накладная мастера, 14 — упаковочная ведомость, 15 — запрос: а — хранится до получения № 11, б — направляется после получения №№ 10—13, в — направляется после отгрузки.

также привлекла специального консультанта по вопросам систем и процедур. В целом группа насчитывала теперь около десяти человек.

Был проведен анализ прохождения информации при выполнении заказов на изготовление запчастей. Выяснилось, что существовало два вида выполнения этих заказов в зависимости от того, имелись ли готовые детали в наличии на складе или нет. Поток информации для этих двух случаев показан на рис. 2.2 и рис. 2.3. Анализ этих случаев помог определить, каким образом эти процессы могли быть упрощены, для того чтобы сократить объем документации и количество различных видов бумаг. Но эта возможность не привлекла интереса группы в той мере, как два других аспекта системы.

Во-первых, анализ показал, что работники склада брали запчасти из запасов по распоряжению отдела заказов и уведомляли об этом плановый отдел только после того, как выдача была произведена. Подобная практика могла привести к уменьшению запасов ниже уровня, требуемого для обеспечения месячного графика, а отсюда мог возникнуть дефицит, который обнаруживался бы лишь при поступлении заявок на детали для сборки.

Таким образом, фактически регулирование выдачи деталей со склада для выполнения заказов на запчасти находилось в ведении отдела заказов, в то время как выдача деталей для сборки находилась в ведении отдела планирования. Это разделение ответственности в значительной степени являлось причиной дефицита. Очевидно, плановый отдел должен был прежде всего решать, выполнять ли заказы на запчасти или нет, и в критических случаях решать, следует ли исполнять заказы на запчасти или удовлетворять потребности сборки. Принятая процедура в неявном виде приписывала больший вес задержке выполнения заказов на запчасти по сравнению с дефицитом для сборки.

Во-вторых, анализ выполнения заказов на некоторые запчасти, отсутствовавшие на складе, показал, что по распоряжению работника склада имеющиеся детали могли быть выданы со склада, погружены на тележки и перевезены во временное хранилище в ожидании поступления других деталей, необходимых для полной комплектации по данному заказу. Это временное хранение (хранение «А») требовало очень большой площади. Вместо этого можно было применить другой порядок, когда данные детали выделяются для отправки в списках хранения, которые ведет плановый отдел, а выдача со склада производится только тогда, когда имеется в наличии весь комплект требуемых деталей. Такая процедура дает два основных преимущества.

1. Она сократит, а в конце концов сведет на нет хранение «А» и таким образом освободит эту площадь для хранения готовой продукции.

2. Она сократит число физических операций по перемещению деталей, указанных в заказе на запчасти.

Исходя из этих соображений был разработан новый процесс выполнения заказов на запчасти. Изменения в этом процессе начали проводиться в жизнь сразу же несколькими отделами. Уже через короткий промежуток времени можно было наблюдать существенные результаты.

Оставалось сделать некоторые уточнения. Для этого была выделена подгруппа, которая занялась тщательным пересмотром системы. Новая процедура выполнения заказов на запчасти представлена на рис. 2.4 и рис. 2.5. Как можно видеть, в результате сосредоточения ответственности в плановом отделе оказалось возможным не только облегчить затруднения с дефицитом деталей и упростить обработку документации, но также и облегчить задачу хранения деталей, предназначенных для сборки.

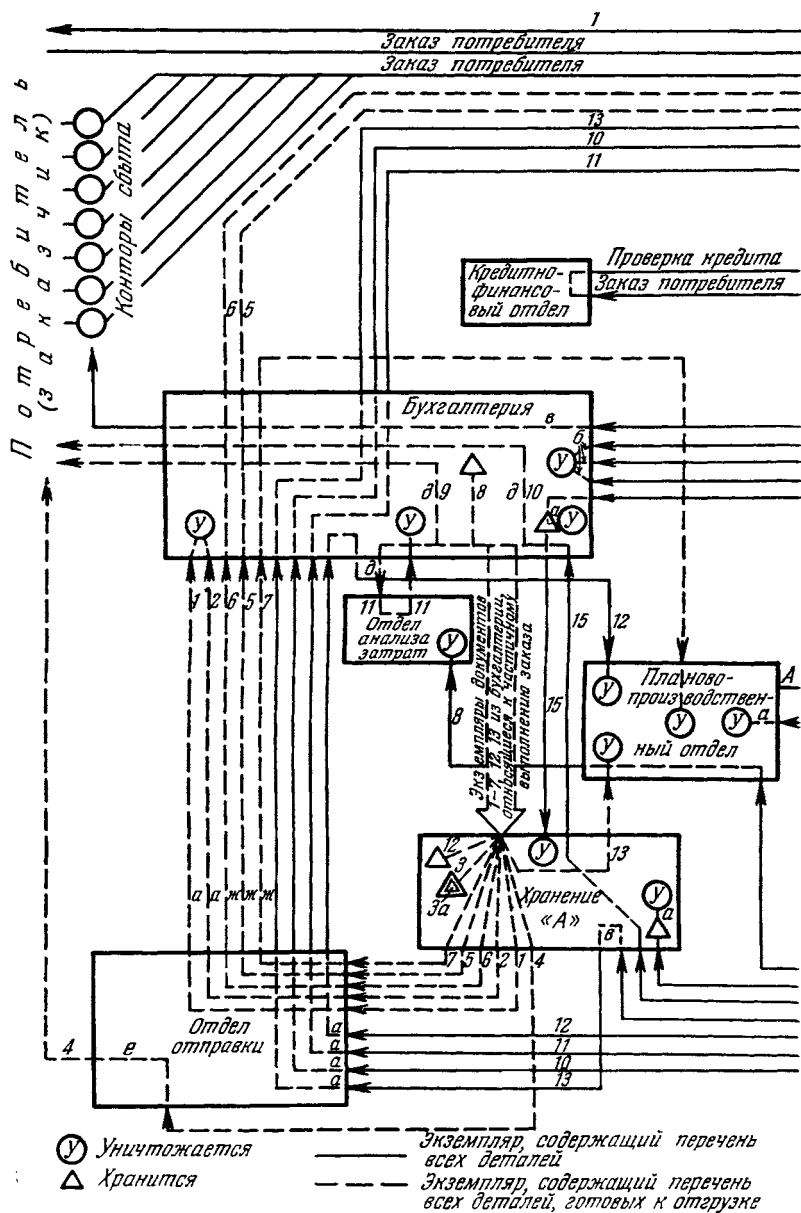
Вернемся снова к основной задаче построения плана производства деталей для сборки.

Закупка и хранение сырья. При оценке результатов предложений по изменению операций необходимо было посмотреть, могут ли они каким-либо образом сказаться на закупках сырья, и если да, то как именно. Было ясно, что с увеличением размера партий закупки можно проводить реже, но в больших количествах. Было организовано исследование с целью установить, дадут ли закупки в больших количествах какие-нибудь выгоды с точки зрения цен. Ответ был отрицательный. Для большинства материалов фирма не могла воспользоваться никакими преимуществами по скидкам на цены при закупке в больших количествах, поскольку и при увеличении закупок фирма оставалась сравнительно незначительным потребителем этого сырья. Но было установлено, что при закупке сырья в больших количествах могут быть существенно сокращены транспортные расходы. Например, фирма нанимала грузовики для перевозки крупного литья. Значительно большее количество литья может быть перевезено при незначительных дополнительных издержках.

Наиболее существенным результатом изучения влияния производства по экономичным размерам партий на сырье было установление того факта, что при сохранении объема закупок может быть достигнута значительная экономия при хранении сырья. Это является следствием более крупных и менее частых выдач сырья со склада. Было показано, что по материалам, заказываемым раз в полгода, средний объем запасов может быть сокращен на 55%. По материалам, заказываемым раз в году, возможное среднее сокращение запасов составило 36 %. Это сокращение имеет и дополнительное значение, так как одновременно ведет к сокращению требуемых для хранения сырья площадей, которые могут в случае необходимости быть использованы для хранения готовых деталей.

Старение. Металлические детали, учитывавшиеся при исследовании, не портятся при хранении, но могут устаревать в связи с изменениями в конструкции. Хотя эта проблема не представляла серьезных затруднений для фирмы, плановый отдел пытался выяснить, не может ли эта проблема при увеличении запасов стать более серьезной. Исследование показало, что среднее время, необходимое для разработки новой конструкции, превышало 9 месяцев, что совпадало со средней продолжительностью предложенного производственного цикла. Вследствие этого было установлено правило, обязывающее технический отдел уведомлять плановый отдел о начале работ, которые в конечном итоге могут сказаться на конструкции той или иной детали. В карту учета хранения этой детали тогда может быть записано, что запрещается выпуск деталей, превышающий объем, необходимый для удовлетворения трехмесячных потребностей.

График сборки. В практику фирмы входила сборка нескольких станков каждой модели ежемесячно. Это означало, что каждый месяц требовалось некоторое количество деталей всех типов (за исключением тех, которые используются только для очень специальных приспособлений). Конечно, можно собирать каждую модель только через два, три, а то и более месяцев. Такое изменение в графике в свою очередь отразится на порядке выдачи деталей со склада. Анализ показал, что для деталей, которые использовались только в одной модели, средние запасы (а следовательно, и общие производственные расходы) могут быть еще более сокращены. Такие сокращения,



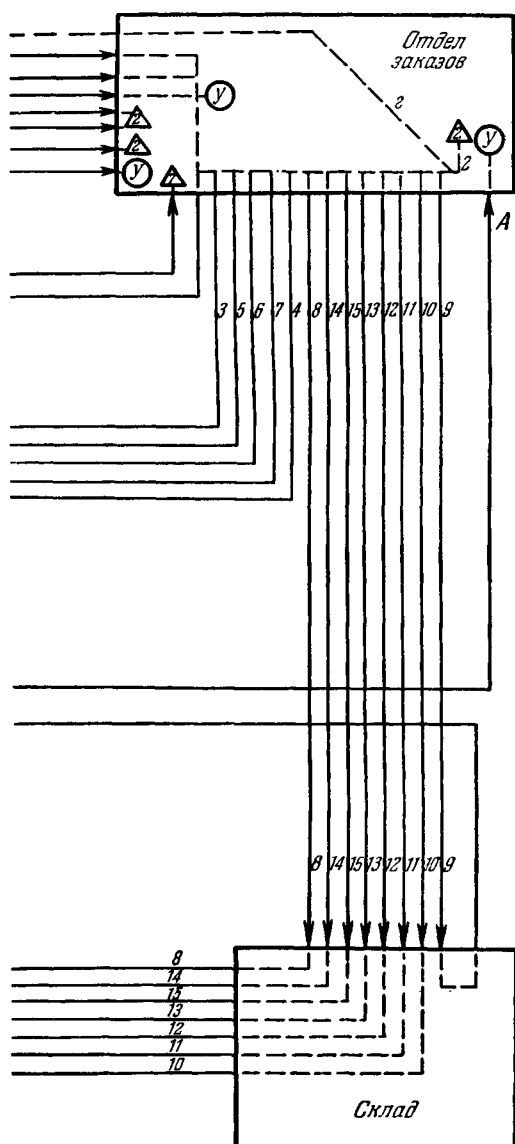
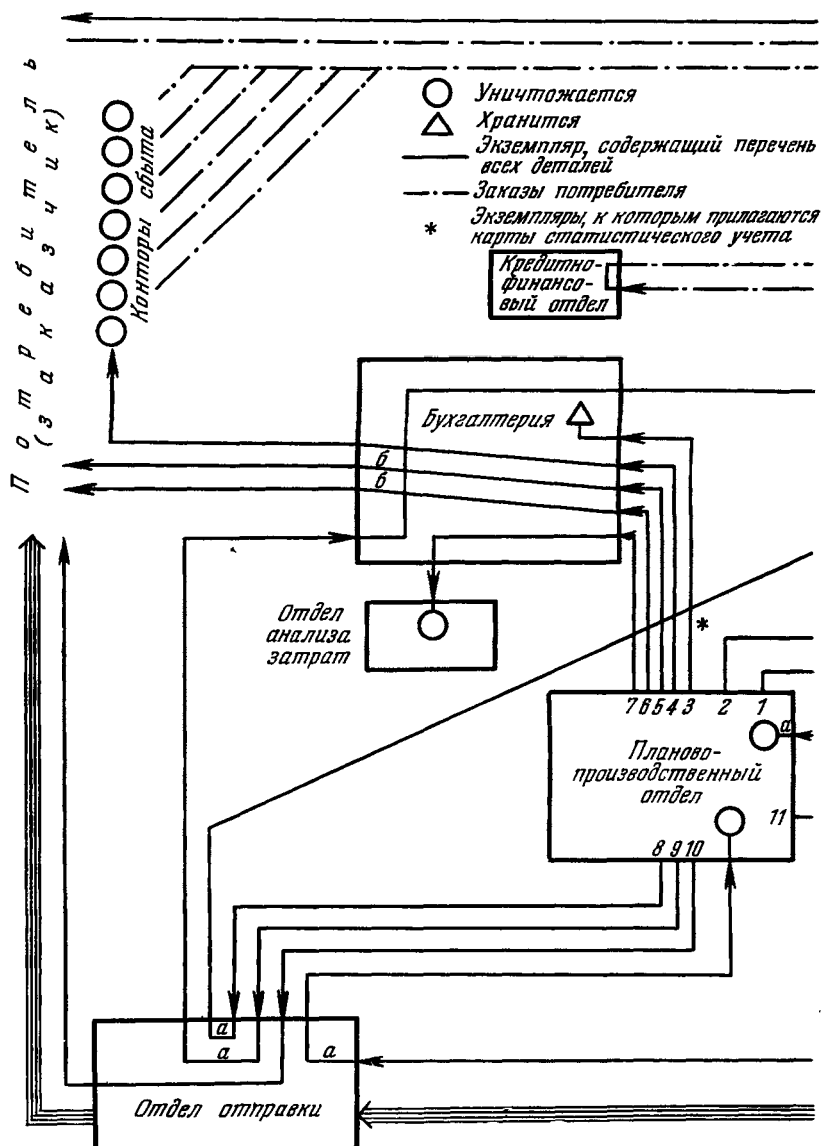


Рис. 2.3. Схема прохождения заказа на запчасти на предприятии станкостроительной фирмы для случая, когда заказанных частей нет в наличии.

Как и на рис. 2.2, при такой схеме контора склада управляет ходом выполнения заказа, что приводит к недостаткам, описанным в тексте. Готовые детали, входящие в комплект заказа, направляются на хранение «А» до поступления недостающих до полного комплекта деталей.

1 — экземпляр упаковочной ведомости ОТК, 2 — экземпляр упаковочной ведомости отдела отправки, 3 — экземпляр упаковочной ведомости склада, 4 — экземпляр упаковочной ведомости для заказчика, 5 — экземпляр упаковочной ведомости бухгалтерии, 6 — экземпляр упаковочной ведомости отдела заказов, 7 — экземпляр упаковочной ведомости планово-производственного отдела, 8 — ведомость бухгалтерии, 9 — оригинал накладной, 10 — дубликат накладной, 11 — наряд с расценками, 12 — наряд-задание, 13 — наряд на начисление премий, А — график выполнения заказа: а — с учетом полной комплектации заказа, б — после поступления № 15, в — после поступления первой комплектующей детали, г — после поступления № 5, 6, 7 (пунктирные линии), д — направляется вместе с материалом, ж — направляется после отгрузки каждой части комплекта.



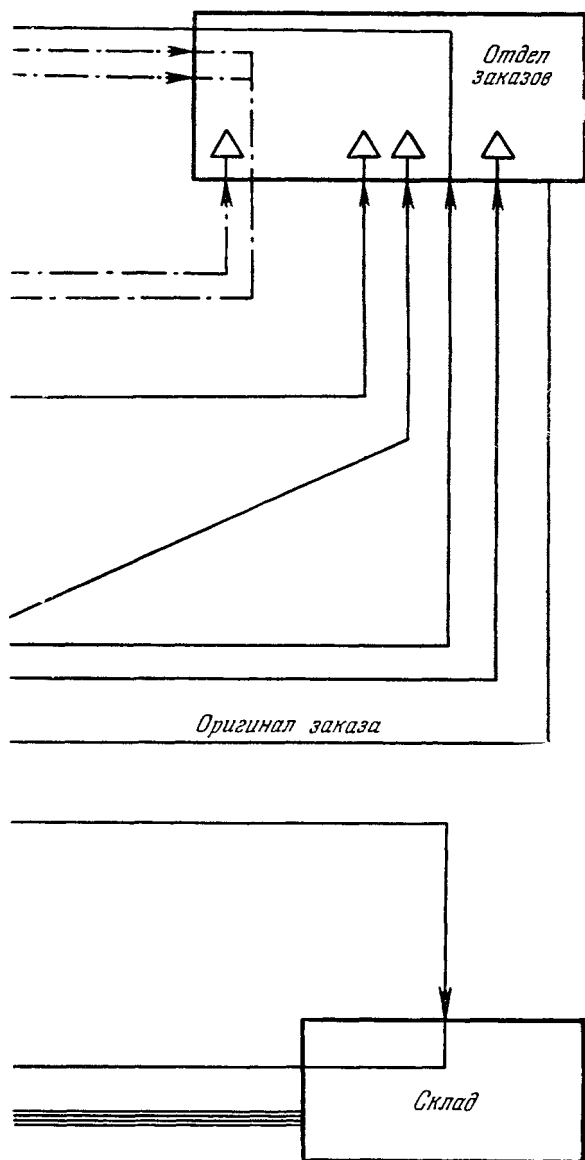
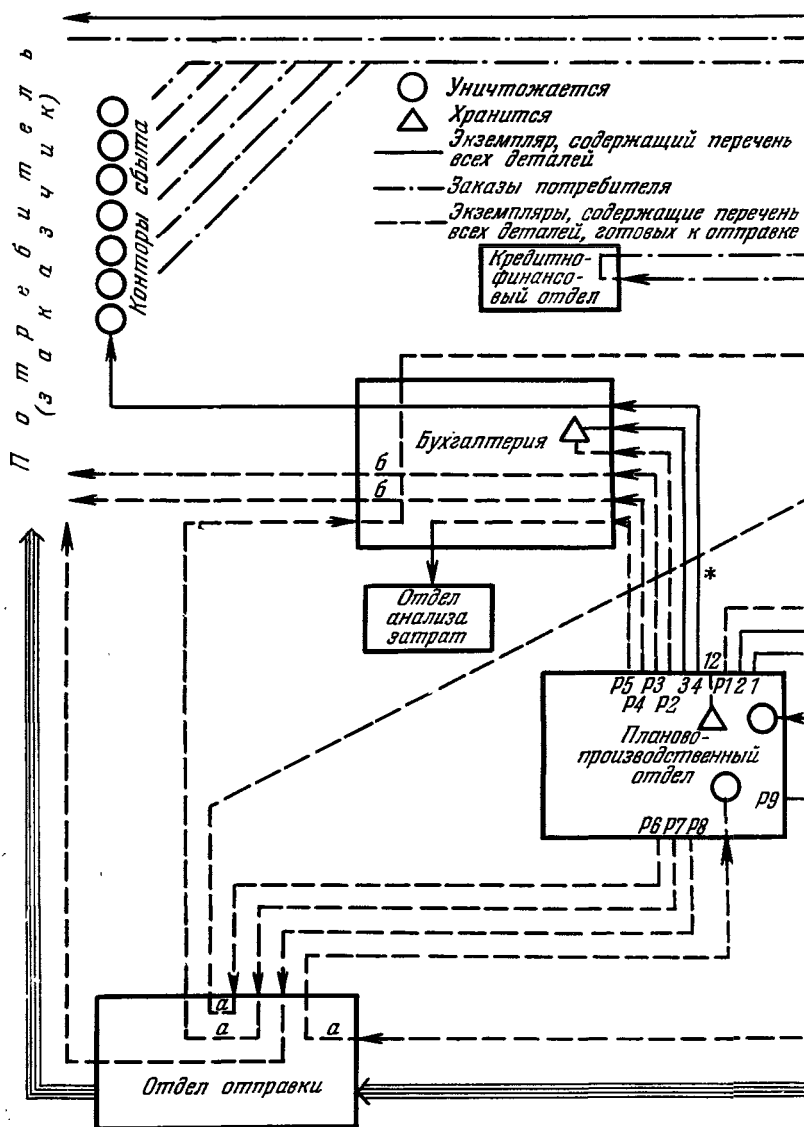


Рис. 2.4. Усовершенствованная схема прохождения заказа на запчасти на предприятиях станкостроительной фирмы для случая, когда все заказанные части имеются в наличии

В этой новой схеме ход выполнения заказа зависит прежде всего от кредитно-финансового отдела, а затем от планово-производственного. Таким образом, объем документации сокращен и все учетные документы сосредоточены в планово-производственном отделе.

1 — учетная ведомость отдела заказов, 2 — подтверждение получения заказа, 3 — учетная ведомость бухгалтерии, 4 — извещение конторам сбыта, 5 — накладная, 6 — дубликат накладной, 7 — экземпляр отдела анализа затрат, 8 — экземпляр мастера, 9 — отгрузочная ведомость, 10 — упаковочная ведомость, 11 — требование складу, 12 — сопроводительный экземпляр планово-производственного отдела, а — направляется после отгрузки, б — направляется после получения № 9.



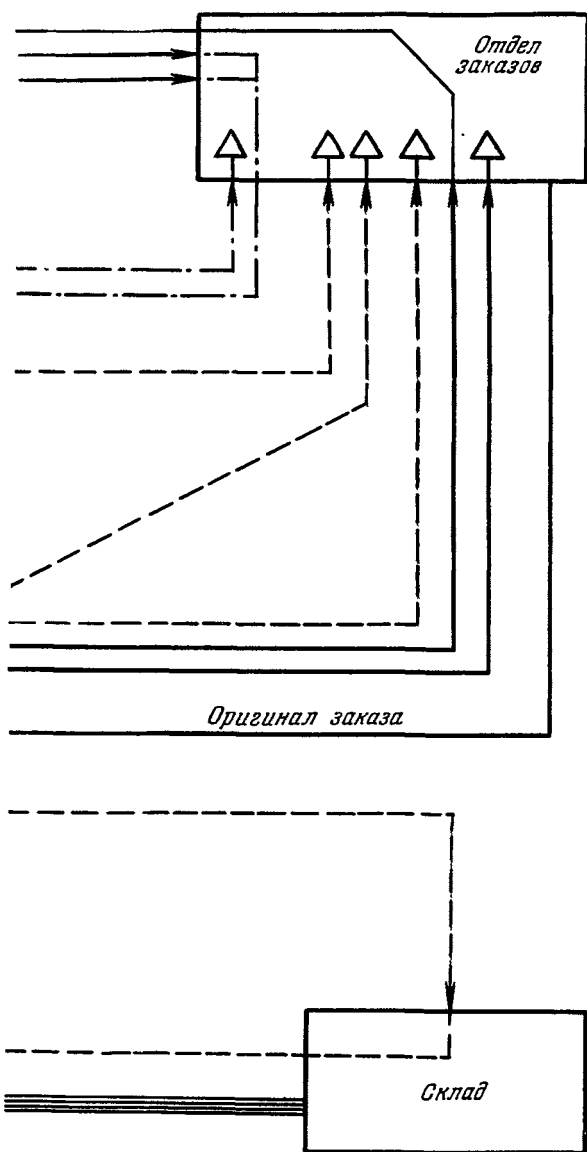


Рис. 2.5. Усовершенствованная схема рис. 2.3, устраняющая отмеченные недостатки.

Управление перенесено в планово-производственный отдел, где сосредоточена вся учетная документация. Хранение «А» ликвидировано, объем документации сокращен. Эта схема разработана по типу электрических схем регулирования.

1 — учетная ведомость отдела заказов, 2 — подтверждение получения заказа, 3 — учетная ведомость бухгалтерии, 4 — извещения конторам сбыта, Р1 — учетная ведомость отдела заказов, Р2 — учетная ведомость бухгалтерии, Р3 — оригинал накладной, Р4 — дубликат накладной, Р5 — экземпляр отдела анализа затрат, Р6 — экземпляр мастера, Р7 — отгрузочная ведомость, Р8 — упаковочная ведомость, Р9 — требования склада, 12 — сопроводительный экземпляр планово-производственного отдела, а — направляется после отгрузки, б — направляется после получения № 9, * — экземпляры, к которым прилагаются карты статистического учета.

однако, могут быть произведены только для небольшого процента видов деталей, поскольку число деталей, применяемых только в одной какой-либо модели, невелико. Для того чтобы можно было использовать эту возможность небольшого сокращения расходов, необходимо значительно усложнить планирование операций по сборке и по производству деталей. Расходы, связанные с необходимостью усложнения планирования, были подсчитаны и сопоставлены с возможным сокращением расходов на запасы. В результате было принято решение не рекомендовать каких-либо изменений в планировании сборки или в производстве деталей.

Оперативно-производственное планирование при переменном спросе. До сих пор исследование велось применительно к известным потребностям на детали, необходимые для сборки. Естественно, что фирма интересовалась также тем, насколько применимо планирование производства по оптимальным размерам партий в условиях по крайней мере частично известного спроса. Уравнение экономического размера партий можно приспособить для использования в этих условиях, если получить объективную и надежную оценку вероятного сбыта. Были проведены исследования с целью определить, существует ли связь между месячным или трехмесячным сбытом станков и какими-либо показателями, относящимися к сырью, изделиям, связанным с данным производством, и общим экономическим условиям в стране. Хотя некоторые зависимости были найдены, ни одна из них не давала надежной оценки краткосрочного сбыта.

До войны в Корее каждый отдел сбыта составлял месячные и трехмесячные прогнозы по данным отдельных агентов. Было установлено, что из двух этих видов прогноза трехмесячные оценки были более надежными, но в то же время смещенными. Иначе говоря, эти предсказания переоценивали фактический сбыт. С учетом этой погрешности были получены объективные оценки, которые имели отклонение от фактического сбыта за три месяца в среднем на 18%.

Анализ тенденций рынка. Была произведена бесполезная попытка получить лучший прогноз, установив тенденции, которые могут быть использованы для оценки сбыта на будущее, на основании накопленных статистических данных. Оценки, полученные таким образом, оказались вдвое менее надежными, чем оценки, определенные по индивидуальным оценкам агентов. Однако они были несмещенными. Было установлено, что предполагаемый сбыт отклоняется случайным образом от основной линии, выражающей тенденцию рынка. В результате был предложен метод прогнозирования изменений в тенденциях рынка.

По двенадцати точкам, взятым за последовательные трехмесячные промежутки времени (подсчитанные помесечно), методом наименьших квадратов был построен график изменения тенденции рынка. Было подсчитано среднеквадратичное отклонение этих двенадцати величин относительно графика. Затем на график были нанесены еще две линии над и под основной линией на расстоянии двух среднеквадратичных отклонений. Затем основная линия и ограничения были экстраполированы. Каждый месяц таким образом графически определялись объемы сбыта на последующие три месяца.

Если цифра сбыта выходила за ограничения, предсказывалось изменение в основной тенденции сбыта и находилась заново основная линия и ограничения, которые подсчитывались на основании последнего набора точек, выразивших новую тенденцию. Если таких точек оказывалось менее шести,

полный подсчет новой тенденции и ее ограничений производился заново каждый месяц до тех пор, пока в новой тенденции не были получены шесть точек. Этот процесс продолжался до тех пор, пока какая-либо точка не выходила за новые ограничения (см. рис. 2.6).

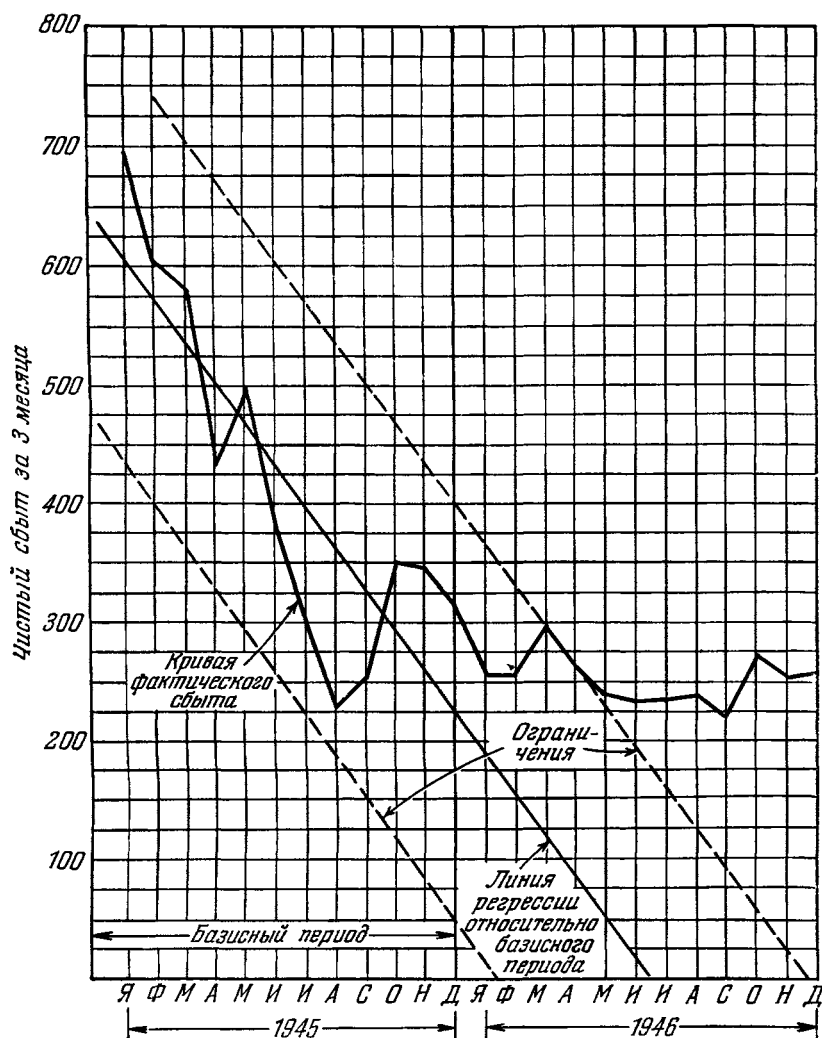


Рис. 2.6. Построение графика тенденций рынка.

Эта процедура была проверена ретроспективно более чем за семилетний период, и каждое изменение в тенденции рынка было правильно предсказано до его появления. С другой стороны, не было предсказано ни одного изменения, которое бы впоследствии не произошло.

Один из работников отдела сбыта предложил модификацию расчета, в которой использовались как предсказания, так и фактический сбыт и в которой ограничения тенденции выражались в процентах от значений основной линии. Основная линия наносилась методом, описанным выше. Предсказания (с поправками на систематическую ошибку) за 12 периодов

были затем нанесены на график, и было определено отклонение каждого предсказания от основной линии. Эти отклонения были выражены в процентах от значений (выше или ниже по вертикали от предсказанного) на основной линии. Было подсчитано среднеквадратичное отклонение этих значений в процентах. Затем на расстоянии трех среднеквадратичных отклонений были нанесены ограничения. Границы ограничений не были параллельны

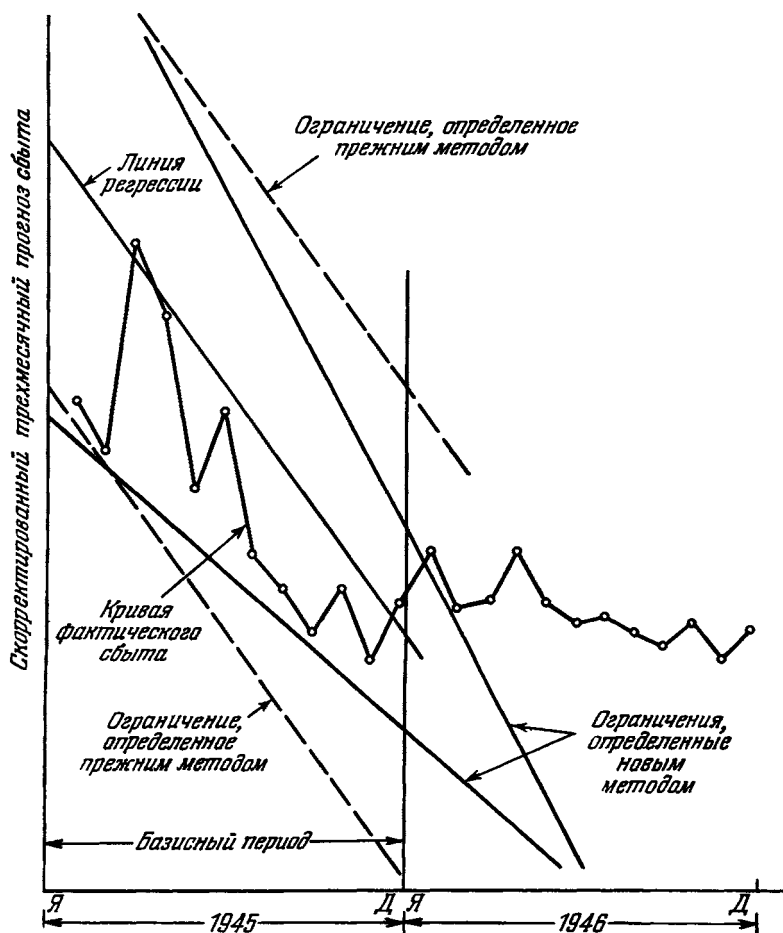


Рис. 2.7. Построение ограничений для графика тенденций рынка.

основной линии, поскольку они наносились на определенном расстоянии от нее, выраженном в процентах. Поэтому они приближались к основной линии в области малых значений и расходились в области больших значений (см. рис. 2.7). Основная линия и ограничения экстраполировались, и на график наносились скорректированные предсказания на последующие три месяца. Когда исследуемое значение выходило за пределы установленных ограничений, предсказывалось изменение в основной тенденции и определялась новая тенденция. Эта процедура позволила предсказать (и в каждом случае правильно) изменения в основной тенденции рынка на 2—5 месяцев раньше, чем ранее описанный метод.

Первый метод был затем применен к анализу рынка, не связанного с деятельностью данной фирмы (при использовании скользящего годового

среднего), с такими же хорошими результатами. В обоих случаях метод хорошо зарекомендовал себя в применении к анализу как статистических, так и текущих данных.

Долгосрочный прогноз. Исследование долгосрочного прогнозирования обнаружило полезное соотношение между одним из общих показателей по стране и годичным сбытом фирмы. Возник вопрос: можно ли разбить годовой сбыт на две его составляющие: первичная продажа станков и продажа для замены устаревшего оборудования? Предполагалось, что в сбыте станков для замены наблюдается большая стабильность, чем в первичном сбыте, а следовательно, если их можно предсказывать отдельно, то можно произвести более надежное прогнозирование минимального сбыта на следующий год.

Замена станков. При учете сбыта не указывалось, была ли продажа первичной или с целью замены. Операционная группа предложила провести выборочный опрос заказчиков с целью установить путем анкетирования, какой процент закупок они производили на протяжении ряда лет с целью замены изношенных станков. Однако представители администрации фирмы сочли, что важность проблемы не столь велика и не оправдывает расходов на проведение подобного опроса. Группе было предложено рассмотреть возможность извлечения соответствующей информации из имеющихся данных.

Группа начала работу с попытки ответить на вопрос: почему владелец предприятия заменяет один станок другим? Выпускаемые фирмой станки не приходят в полную негодность, они не «перегорают», как лампочки. Конечно, их производительность уменьшается, и это снижение производительности может быть предотвращено или отсрочено путем замены деталей. Но, как было установлено, объем ремонтных работ со временем все увеличивается. Поэтому, как представлялось группе, замена станков должна происходить тогда, когда владелец предприятия решает, что непрерывный ремонт оборудования будет обходиться дороже, чем покупка нового. Если допустить, что это так, то отсюда следует, что снижение объема ремонтных работ указывает на предстоящую близкую смену оборудования. Этот вывод указал на опосредствованный метод исследования замены станков.

Группа взяла случайную выборку из заказов на запчасти, полученных в 1940, 1944 и 1948—1952 годах включительно. В год принималось таких заказов до 92 тыс. Из каждого взятого для обработки заказа извлекалась следующая информация:

1. Год покупки станка, для которого заказывались запчасти.
2. Стоимость заказа в долларах.
3. Год, в котором был получен заказ.

Полученные заказы были затем расклассифицированы по годам закупки станков. Затем была подсчитана общая стоимость каждого полученного подкласса и преобразована в среднюю сумму (в долларах) затраченную на станок. Затем эти данные были нанесены на график с нарастающим итогом, как показано на рис. 2.8. Каждый график показывал годовое количество станков на тот год, с которого прием заказов на запчасти начинал уменьшаться. Число лет между взятым годом и годом уменьшения заказов на ремонт был определен на семи графиках. Полученные результаты приводятся в таблице 2.3.

Эти данные не поддавались интерпретации, пока не было осознано, что количество лет до начала спада на заказы запчастей не имеет особого значения, а важно, сколько времени проработал станок до года такого спада.

Таблица 2.3

Год размещения заказа на запчасти	1952	1951	1950	1949	1948	1944	1940
Число лет до начала уменьшения заказов на запчасти	17	20	17	19	15	16	22

Специальный анализ показал, что надежным показателем использования станков является индекс, приведенный в справочнике Федерального бюро

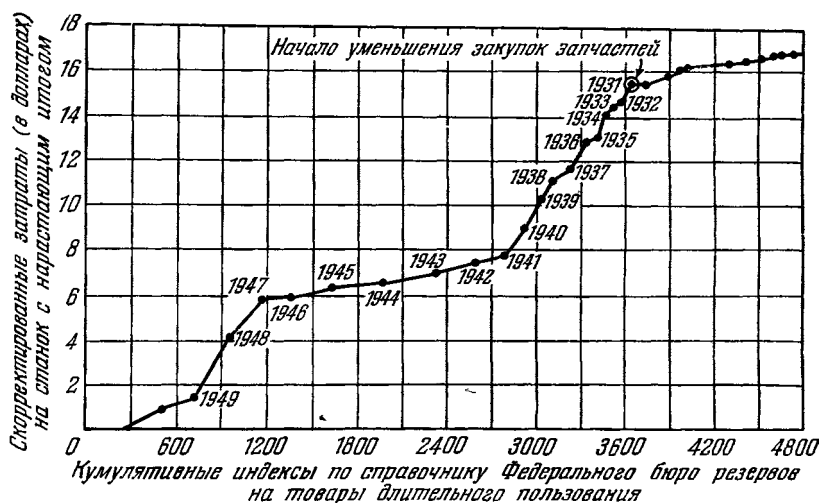


Рис. 2.8. Скорректированные затраты (в долларах) на ремонт станков по годам закупки.

резервов на товары длительного пользования. Годы, предшествовавшие точке спада в заказах на запчасти, были выражены через показатели, взятые из этого справочника. Полученные данные представлены в таблице 2.4.

Таблица 2.4

Год размещения заказа на запчасти	1952	1951	1950	1949	1948	1944	1940
Кумулятивные индексы по справочнику ФБР на товары длительного пользования, указывающие на время уменьшения заказов на запчасти в прошлом	3648	3641	3263	3232	2824	2368	2100

Эта таблица нашла два неожиданных, но существенных применения. Прежде всего, она показывает все возрастающую производительность продукции фирмы, а также дает возможность ее измерить. Это имело немаловажное значение, особенно в пересмотре сделок, связанных с военными поставками. Во-вторых, наблюдаемая тенденция в соединении с предсказаниями справочника ФБР на текущий год могла быть использована для составления списка закупок станков по годам, который мог бы лечь

в основу стратегии упорядочения сбыта запчастей. Такой список фактически содержит указание на то, кем, когда и какие станки, как можно ожидать, будут куплены в замену устаревших. Оба эти результата частично восполнили тот недостаток, что данный анализ нельзя было использовать для оценки (определимой надежности и точности) годового оборота сбыта станков для замены.

Был исследован ряд других задач прогнозирования, включая и задачу, которая показала тесную взаимосвязь между заказами на запчасти и продажей станков. Эти исследования дали возможность заложить более прочную базу для производства более крупных партий деталей в условиях неопределенности.

ПЛАНИРОВАНИЕ ПРОИЗВОДСТВА

Последний этап исследований потребовал разработки процедуры, с помощью которой можно было бы произвести фактическое планирование производства. Было нереалистично предполагать, что работники отдела

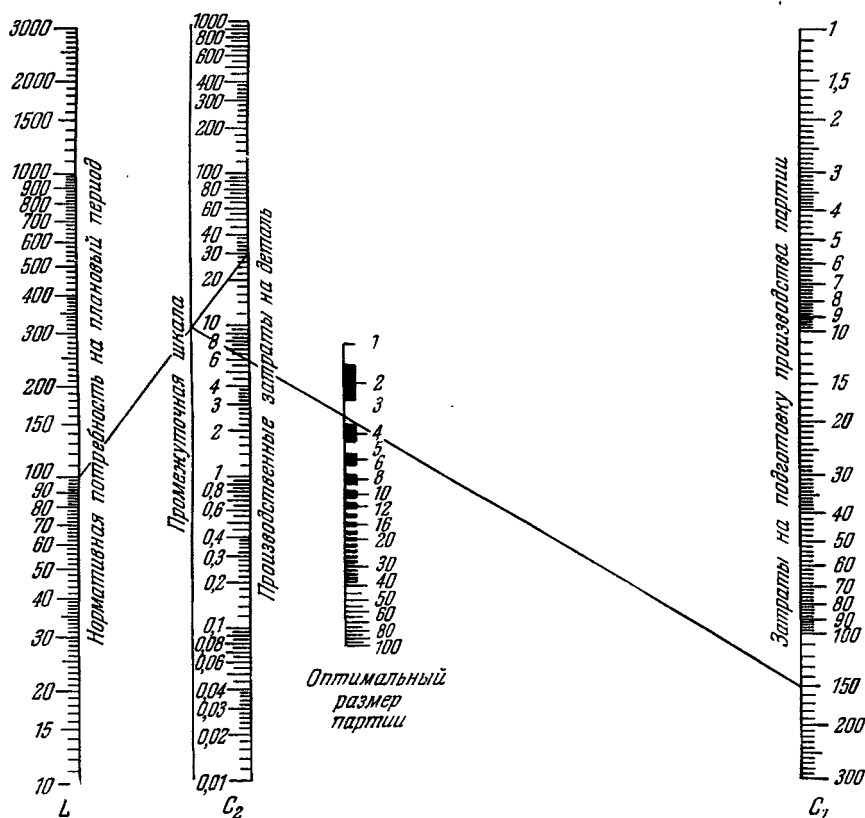


Рис. 2.9. Номограмма для определения оптимального размера партии при заданной месячной потребности в деталях $\sqrt{199C_1/LC_2}$. Пример: $L = 100$, $C_1 = \$150$, $C_2 = \$30$. При этих условиях оптимальный размер партии равен 3.

планирования будут непосредственно использовать уравнение оптимального размера партий. Поэтому был применен графический прием (номограмма),

чтобы дать сотрудникам отдела планирования возможность определять экономические размеры партий без всяких подсчетов, простым проведением двух линий (см. рис. 2.9).

Для использования номограммы необходимо располагать значениями расходов на подготовку производства, непосредственно производственных затрат, а также данными о месячных потребностях в деталях и удобной системой их регистрации. Представлялось естественным использовать для этого карту учета запасов. Форма карты пересматривалась применительно к аппаратуре IBM. Требуемую информацию оказалось возможным нанести на новую карту. Это, конечно, подсказало возможность выполнения подсчетов оптимальных размеров партий на аппаратуре IBM, что и было сделано легко, экономично и довольно быстро. Нограмма была тем не менее сохранена для проверки производства тех деталей, которые представляли особую сложность.

ЗАКЛЮЧЕНИЕ

В только что рассмотренном примере оптимизация была первоначально достигнута только в очень ограниченной области. Однако эта процедура оптимального планирования производства не могла быть применена на практике, пока не были решены многочисленные побочные задачи, связанные с ней. Влияние предложенного принципа планирования было прослежено фактически в функциях почти всех отдельных подразделений фирмы. Была затронута почти каждая сторона производства: планирование производства, изготовление изделий, процедуры, связанные с хранением, операции сборки (при исследовании проблемы уменьшения дефицита). Кроме того, в процессе исследования были так или иначе затронуты финансовые операции, технические проблемы, вопросы снабжения и сбыта, и все эти стороны деятельности фирмы были скоординированы в общий план. Попытка ввести в производство экономические размеры партий без этого подхода к системе как единому целому либо была бы обречена на провал еще в самом начале, либо привела бы к губительным результатам для фирмы.

Глава 3

КОМПЛЕКСНЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ КОНТРОЛЯ КАЧЕСТВА*

ВВЕДЕНИЕ

Одним из новых принципов исследования операций является комплексный подход к задаче. Новые, более эффективные решения задач находятся только тогда, когда сама задача рассматривается в новом свете и когда к ней применяются новые методы анализа. Комплексный подход обогащает исследование операций новизной принципов и методов решения. Каждый член операционной группы подходит к задаче со своим багажом опыта и образования, располагая разнообразными аналитическими средствами, которые уже зарекомендовали себя в соответствующей области знаний.

* Лоринг Г. Миттен (Loring G. Mitten), профессор кафедры организации производства университета штата Огайо.

Из приводимого ниже примера станет совершенно очевидным, что эффективные решения были получены благодаря последовательному применению комплексной операционной группой научных принципов исследования.

ЗАДАЧА

Около четырех с половиной лет назад крупная промышленная фирма обратилась в университет штата Огайо с просьбой выполнить исследование факторов, влияющих на показатели визуального контроля качества. Фирма была заинтересована в том, чтобы найти пути повышения качества и снижения затрат на визуальный контроль одной из деталей, которую она производила для последующего использования в своем конечном изделии. Задача контроля требует стопроцентного осмотра изделия со сравнительно большой скоростью с целью обнаружения изъянов и дефектов, некоторые из которых настолько незначительны, что едва видны. Огромное количество производимых и контролируемых ежегодно деталей (более 2 млрд. штук) и важность качества этих деталей как для репутации фирмы, так и для удовлетворения запросов покупателя и гарантии надежности — все это вместе взятое свидетельствует о значительном объеме и сложности поставленной задачи.

Контракт на проведение исследования был заключен, и фирма выделила начальника ОТК в качестве руководителя проекта. Университет организовал исследовательскую группу, в которую вошли оптик-механик, психолог и автор статьи — статистик и специалист по организации производства. Было оговорено, что исследование должно быть направлено на обнаружение тех факторов, которые имеют существенное влияние на показатели визуального контроля качества. Впоследствии была также предпринята попытка установить некоторую меру оценки количественных и качественных показателей работы контролеров ОТК.

ПЕРВЫЙ ЭТАП

В течение года в университетских лабораториях велись исследования различных аспектов задачи. Оптик-механик изучал корреляцию между показателями визуального контроля и различного типа и силы освещения, мерой остроты зрения контролеров и характером движений глаз. Психолог изучал проблемы стимулирования, факторы удовлетворенности работой и проблемы усталости. Специалист по организации производства занимался исследованием влияния различного типа конструкций контрольного оборудования и статистическим анализом экспериментальных данных.

Был применен широкий и разнообразный комплекс научных методов, начиная от физиологической оптики до тестов на определение степени профессиональной пригодности. Обмен мнениями между представителями трех дисциплин и взаимное использование специфических для каждой дисциплины методов во многом способствовали успеху совместных исследований. Одна из наиболее сложных задач, с которой пришлось столкнуться, заключалась в определении удовлетворительной меры для оценки показателей визуального контроля. Было установлено, что принятые фирмой оценки в целом неудовлетворительны. Еще до завершения экспериментов были разработаны восемь новых методов оценки показателей визуального контроля. Отсутствие рациональных и надежных оценок вызвало серьезные затруднения в предыдущих исследованиях проблемы контроля качества, проводившихся ранее силами самой фирмы или другими организациями.

Исследования показали, что существует ряд факторов, влияющих на показатели визуального контроля, наиболее существенным из которых является отношение контролера к своей работе. Далее был сделан вывод, что при оптимальных условиях производительность труда контролеров может быть увеличена на 300%—400% при значительном улучшении качества самой выполняемой работы.

Имея на руках эти данные, исследовательская группа и руководитель проекта представили личный доклад высшей администрации фирмы. Фирме был дан совет предпринять значительно расширенную программу исследований, для того чтобы отыскать средства использования полученных результатов и практические меры по улучшению визуального контроля. Руководители фирмы, люди очень консервативные, сначала отнеслись ко всему исследованию чрезвычайно скептически, но к концу совещания от их скептицизма не осталось и следа, и они сами предложили продолжить исследования в еще более широком масштабе, чем рекомендовала группа.

ВТОРОЙ ЭТАП

В качестве второго этапа исследований было решено тщательно изучить условия работы отдела технического контроля на предприятии и процедуры, составляющие содержание его работы. В результате был получен ряд интересных наблюдений, два из которых имеют особенно важное значение.

Обнаружение дефектов. Поскольку многие дефекты, которые контролеры ОТК должны были обнаруживать на поверхности детали, чрезвычайно малы, фирма снабдила каждую контрольную установку увеличительным стеклом. Логика этой меры совершенно ясна: когда требуется рассмотреть что-нибудь очень мелкое, нужно взять увеличительное стекло. Оптику-механику, входившему в состав группы, удалось, однако, доказать, что логический анализ столь сложных проблем, как эта, может оказаться чрезвычайно тонким. Он обнаружил, что каждый контролер в отделе применял увеличительное стекло не для увеличения, а как корректирующую линзу. Другими словами, линзы использовались не для того, чтобы объект выглядел крупнее, а для того, чтобы не напрягать зрение. После исследования задачи оптик-механик сумел сконструировать новую линзу, которая сочетала эффект увеличения и оптической коррекции. Одновременно новая линза снижала напряжение зрения, вызванное необходимостью изменения глубины фокуса глаза при переходе от одной поверхности к другой. В результате показатели визуального контроля были значительно улучшены.

Отношение контролеров к работе. Вышеизложенное служит наглядным примером того вклада, который может внести ученый в решение операционных задач. Факты, которые для оптика-механика оказались такими очевидными, возможно, никогда не обратили бы на себя внимание людей, повседневно занимающихся вопросами контроля качества.

Другой интересный аспект связан с изучением отношения контролеров к работе. Сразу же после одобрения предложения о расширении программы исследований группа обратилась к администрации за разрешением провести обследование отношения персонала отдела технического контроля к выполняемой им работе. Администрация заявила, что ее интересуют только результаты работы, а не отношение к ней. Поскольку предварительные данные исследования указывали на то, что отношение контролеров к работе, является, возможно, одним из основных факторов, определяющих качество

визуального контроля, группа настойчиво добивалась удовлетворения своей просьбы. Администрация указывала, что она не намеревается предпринимать подобное исследование, поскольку она уверена, что работники отдела используют предлагаемое обследование как повод для требования увеличения заработной платы. Таким образом, стало ясно, что нежелание администрации было вызвано не ее безразличным отношением к настроению работников, а уверенностью в том, что она точно знает это настроение («мало платят»). Наконец, администрация уступила просьбам группы, и группа начала задуманное обследование.

Здесь следует отметить, что определение отношения к работе — дело чрезвычайно сложное и капризное и может быть поручено только очень компетентным специалистам. Задача в данном случае усложнялась тем, что контролерами ОТК были только женщины. Перед операционной группой стояла прежде всего задача завоевать доверие браковщиц и определить возможно более объективно их настроение, которое администрация считала негативным. Задача была решена ко всеобщему удовлетворению, когда задание по опросу работниц было поручено молодому (и красивому) психологу. (Это был молодой человек такого типа, при одном виде которого девушки сразу начинают шушукаться: «Какой симпатичный!») Не стоит говорить, что, поручив опрос такому человеку, мы довольно легко завоевали доверие работниц.

Анализ результатов опроса позволил выяснить некоторые очень интересные факты. Прежде всего, из 150 опрошенных браковщиц только три вообще заговорили о заработной плате; две из них добровольно сообщили, что они считают оплату вполне достаточной, и только одна пожаловалась на низкую заработную плату. В то же время 136 из 150 заявили, что стулья для работы неудобны. Эти результаты для администрации оказались совершенно неожиданными. Когда руководители пришли в себя после этого приятного сюрприза, они немедленно выставили разнообразные образцы стульев, указав, что каждая браковщица может выбрать для себя любой стул, который она посчитает наиболее удобным. Через неделю весь отдел был оборудован новыми стульями, что обошлось фирме в 10 тыс. долларов. Само собой разумеется, что между администрацией и работницами отдела появилось чувство взаимопонимания, которое никогда раньше не наблюдалось за всю историю фирмы.

В процессе выполнения обследования группа, к сожалению, столкнулась с серьезной проблемой. Браковщицы считали, что новые, более удобные стулья были установлены отчасти благодаря стараниям молодого психолога, проводившего опрос. В их глазах он стал героем. Сохранить научную объективность, когда становишься идолом ста пятидесяти женщин, — это испытание свыше человеческих сил для большинства людей. Поэтому для его же собственного блага мы были вынуждены на год отослать молодого ученого назад в университетскую лабораторию психологии. Там ему поручили наблюдения за поведением крыс в лабиринте, чтобы помочь ему снова обрести понимание перспективы и способность к научной абстракции.

Кроме жалоб на неудобство стульев, исследование отношения контролеров к работе выявило также ряд других обстоятельств, которые на первый взгляд казались тривиальными, но на самом деле служили причиной значительной раздраженности и неудовлетворенности работниц. Это может служить примером того, как научные методы могут в руках опытного эксперта дать фактические данные, на которые администрация может опираться при принятии своих решений.

Обучение. Следующий этап нашего исследования требовал создания «опытного предприятия», на котором мы могли бы проводить наши эксперименты в условиях, приближающихся к реальным. Группа наняла и провела обучение случайно отобранных 12 девушек. Хотя программа обучения контролеров ОТК не была выделена как отдельный объект исследования в этом проекте, необходимость обучения 12 новых контролеров позволила пересмотреть принятые методы обучения. Используя принципы психологической теории обучения, группе удалось разработать новую программу подготовки, которая была проверена на половине вновь набранных контролеров. Результаты показали, что новая система требует в два раза меньше времени, чем система, в то время принятая на предприятии. Новая система была применена на практике, и новые контролеры достигли нормативной производительности и начали обеспечивать себе необходимый заработок в два раза быстрее, чем прежде.

Применение новой системы привело к одному еще даже более важному результату. Договор между фирмой и профсоюзом требует установления определенного испытательного срока для всех, вновь принятых на работу. Во время этого периода фирма имеет право уволить работника по любой причине, но по истечении испытательного срока на работника распространяется действие коллективного договора, и он может быть уволен только при условиях, предусмотренных договором. Период обучения по старой системе превышал испытательный срок, а кроме того, известно, что качество работы во время обучения не дает возможности достаточно правильно оценить производственные показатели будущего работника. Таким образом, фирма почти не имела возможности заранее отсеять посредственных и плохих контролеров до окончания испытательного срока. Новая система обучения позволила полностью исправить это положение; испытательного срока теперь было вполне достаточно для того, чтобы закончить обучение и дать надежную оценку качества работы контролера в нормальных производственных условиях. В результате фирма смогла через некоторое время поднять общий уровень квалификации контролеров ОТК.

Характер движений глаз и освещенность. Другое исследование, проведенное на «опытном предприятии», было связано с изучением характера движения глаз. Совершенно очевидно, что если глаза контролера не направлены на то место поверхности детали, где имеется дефект, то дефект не будет обнаружен, а если это будет часто повторяться, то качество контроля будет очень низким. Задача заключалась, таким образом, в том, чтобы установить, куда именно направлен взгляд контролера. Один из научных сотрудников оптико-механического отделения при университете штата Огайо помог решить эту задачу, разработав новый фотоаппарат, который позволил определить с точностью до $1/32$ дюйма ($\sim 0,75$ мм), куда направлен взгляд контролера при осмотре детали. Анализ фотографий движения глаз показал, что обычная браковщица фактически видит менее 100% поверхности осматриваемых деталей. Это означало, что если она обнаружит и отбракует все детали, которые, по ее наблюдению, являются дефектными, то качество контроля будет все равно далеко не безупречным.

Применив знания возможностей и ограничений человеческого глаза, удалось сконструировать новое контрольное устройство, обучить контролеров рациональному движению глаз, и таким образом эффективно устранить прежние трудности.

На основе приложения законов оптики к решаемой задаче был разработан новый усовершенствованный светильник, который обеспечивал необ-

ходимую освещенность глаз контролера лучами нужной части спектра, что повысило качество работы.

Переменная скорость операций. Исследования операционной группы показали, что трудность задачи контроля меняется в зависимости от процента дефектных деталей в партии. Группа внесла предложение заменить обычное контрольное устройство с фиксированной скоростью устройством с регулируемым приводом, скорость которого контролер может сам менять, регулируя скорость осмотра в соответствии с количеством дефектных деталей в партии и со своими индивидуальными возможностями. Вначале это предложение не было одобрено представителями администрации, которые высказали опасение, что если предоставить самому контролеру возможность регулировать скорость операций, то это может привести к снижению производительности. Но в конце концов такое устройство было смонтировано и испытано. В результате оказалось, что общая производительность не уменьшилась, а увеличилась и при этом значительно улучшилось качество контроля.

Исследование стимулирования. Пожалуй, самым интересным аспектом для исследования явилась проблема стимулирования. Все прочие результаты исследования имели вспомогательный характер. Другими словами, они позволяли улучшить технические параметры контрольных операций до такой степени, что браковщица при желании могла работать с гораздо большей скоростью и точностью. Задача заключалась в том, чтобы пробудить в ней такое желание.

Решение задачи, очевидно, логично было бы искать во введении некоторой системы материальной заинтересованности. Такой подход и был предложен администрацией, и соответствующая система была введена и изучена. Было достигнуто существенное увеличение производительности, но результат обследования индивидуальных возможностей работниц убедил группу в том, что браковщицы могут достигнуть еще более высоких показателей. Логика, лежащая в основе обычных систем материального поощрения, проста и прямолинейна: большая оплата за большой труд. Однако логический анализ задачи, так же как и в случае увеличительных стекол, оказался несколько более тонким.

Анализ накопленного опыта и результаты опроса браковщиц «опытного предприятия» и штатных работниц показали, что лишь для незначительной части зарплата служила основным средством существования. Большинство браковщиц были либо молодые девушки, только что окончившие школу, которые жили с родителями и работали до тех пор, пока не выходили замуж, либо замужние женщины, которые работали, чтобы несколько увеличить доход семьи в дополнение к заработку мужа. Тщательно изучив систему предпочтений на основании экспериментальных данных, группа пришла к выводу, что браковщицы гораздо более заинтересованы в дополнительном свободном времени, чем в дополнительных деньгах. Для того чтобы проверить эти выводы, была введена система поощрения свободным временем. Была установлена норма еженедельной выработки, и, когда браковщица выполняла свою норму, она могла оставшуюся часть недели не выходить на работу, получая при этом, естественно, полную недельную зарплату. При такой системе производительность необычайно резко возросла, а некоторые браковщицы умудрялись выполнять пятидневную норму за два — два с половиной дня. Это дало увеличение производительности почти в три раза по сравнению с обычной на данном предприятии.

Исследование того, как используют браковщицы свое свободное время, вскрыло новый мощный фактор стимулирования. Так, например, одна браковщица стремилась получить больше дополнительного времени, чтобы дольше оставаться дома со своими детьми. Оказалось также, что у старшей и самой опытной браковщицы, которая была на голову выше всех, муж работает в ночную смену. Стало ясно, что дальнейшее исследование этого источника стимулирования в данном случае не требуется. Короче говоря, было установлено, что, вопреки принятому мнению, деньги не являлись самым важным стимулом для данной группы работниц. В их системе предпочтений свободное время играло более важную роль, чем деньги, и они были готовы работать более напряженно, чтобы получить его. Тщательное научное исследование данной группы позволило обнаружить этот факт и использовать его для повышения производственных показателей.

Хотя новая система и оказалась весьма эффективной с точки зрения повышения производительности, при ее введении возникла другая задача. Прежняя система контроля качества не обеспечивала сохранения заданного качества готовой продукции. Поэтому операционная группа занялась научной разработкой новой системы качества контроля. Используя сложный статистический аппарат, группа разработала общий принцип, в соответствии с которым браковщица средней квалификации могла получить наибольшее количество свободного времени, работая с такими количественными и качественными показателями, при которых минимизируются общие затраты фирмы. Группа, таким образом, сумела добиться уникального результата — такого положения, когда наибольшая выгода для работника согласуется с наибольшей выгодой и для фирмы.

ЗАКЛЮЧЕНИЕ

Описание этого примера показывает, как комплексная научная группа может внести нечто новое в анализ и решение организационно-управленческих задач. Рассматривая множество операций с различных точек зрения, группа выявила необыкновенно широкий набор управляемых переменных, которые в дальнейшем были использованы для значительного улучшения показателей контроля качества. «Стандартное» изучение контрольных операций не обнаружило бы такого разнообразия факторов, влияющих на решение задачи.

Контроль качества — это процесс, с помощью которого обнаруживаются уже совершенные ошибки. Более важная задача — предотвращение ошибок. Но работа операционной группы была ограничена администрацией проблемой контроля качества. По мнению группы, операционное исследование задач предотвращения брака значительно сузило бы задачи контроля качества и привело бы к гораздо большим выгодам с точки зрения снижения затрат и повышения качества продукции. Фактически операционная группа была вынуждена искать субоптимальные решения. Однако субоптимизация является недостатком только тогда, когда ученые имеют возможность найти оптимальное решение и не используют ее. Условия, поставленные администрацией в данном случае, были очень жесткими, и проведенное исследование может служить примером эффективного использования тех ограниченных возможностей, которые были либо предоставлены операционной группе, либо вскрыты в результате ее собственных усилий.

Одной из наиболее важных особенностей научного подхода к решению практических задач является умение поставить точную цель исследования. Однако в настоящее время при решении многих практических задач этим зачастую пренебрегают. Например, какую цель преследует Великобритания в своей экспортной политике? Для чего, в конечном счете, готовят университеты студентов? Конечно, на такие вопросы трудно дать ответ. Но обычно можно дать решение, близкое к искомому, что всегда заслуживает внимания.

К таким задачам надо подходить двояко. С одной стороны, существуют общие методы выяснения конечной цели. С другой стороны, надо пытаться определить, что именно достигается в каждом конкретном случае. Только объединяя широкий философский подход с тщательным анализом фактов, можно добиться успеха в определении верного направления в проведении исследований [71].

Существует старая поговорка, что правильно поставить задачу — значит наполовину решить ее. Это всем известно. Однако не все знают, как добиться правильной постановки задачи. Следует признать, что правильная постановка задачи является сама по себе сложной задачей. Когда задача включает систему операций — правительственных, военных, промышленных или коммерческих — то редко можно сразу дать полную и точную постановку задачи. Поэтому ученые склоняются к мысли, что задача редко бывает на самом деле такой, какой она представляется сначала. В качестве иллюстрации можно сослаться на первый пример, приведенный в главе 2. Первоначально ставилась задача определения количества производимой продукции и затрат на обеспечение ее сбыта по каждому из двух видов изделий. Окончательная же формулировка задачи свелась к повышению эффективности производства только одного из этих изделий.

Обычно в первоначальной постановке задачи, которую руководство ставит перед группой исследования операций, скорее даются симптомы, чем диагноз. Обязанностью группы исследования операций является определение точного диагноза (т. е. постановки) задачи, для чего надо привлечь людей, интересы которых затронуты в этой задаче. Чтобы поставить такой диагноз, нужно хорошо знать работу организации и ее систему управления. Но такие сведения редко можно получить от одного лица или одной группы: обычно организация бывает слишком сложной, чтобы отдельные ее сотрудники могли быть в курсе всех операций этой организации. Поэтому группа исследования операций независимо от того, входят в ее состав сотрудники изучаемой организации или нет, должна начать с систематического изучения организации, ее работы и способов управления ее работой.

В четвертой главе «Анализ организации» рассматриваются некоторые методы, которые оказываются полезными при изучении организации. Приводятся сведения, более подробные, чем необходимо только для предварительного анализа организации. Это объясняется двумя причинами.

1. При первоначальном анализе организации может оказаться, что эффективные меры, разработанные группой исследования операций, нельзя успешно применить, так как или (а) информационная система организации не может обеспечить информацию, необходимую для принятия решения, или (б) управление изучаемой системой

разбито на децентрализованные участки. В любом случае следовало бы перестроить организацию, чтобы можно было рассчитывать на эффективные результаты при применении разработанных предложений. Такие случаи встречаются довольно часто. В главе 4 приводится соответствующий пример. Когда приходится сталкиваться с такими трудностями, то сначала следует решать соответствующие информационные задачи или задачи управления. В этих случаях читателю могут помочь подробные сведения, приводимые в главе 4.

2. При внедрении результатов, разработанных в группе исследования операций, почти всегда оказываются необходимыми некоторые изменения в процессах управления и передачи информации. Поэтому подробный анализ этих процессов в организации является очень важным. Следует напомнить, что в приведенном в главе 2 примере задачи хранения и дефицита деталей решались путем изменения процедуры обработки заказов на запасные части и централизации контроля над складскими операциями. На этой стадии исследования был необходим очень тщательный анализ систем связи и управления.

В главе 5 рассматриваются вопросы постановки задачи. Авторами изложена наилучшая, с их точки зрения, процедура постановки задачи. На практике редко встречаются такие случаи, когда можно полностью воспользоваться предлагаемой процедурой. Но она может служить основой, на которой исследователи будут разрабатывать свои методы постановки задач.

Одна из операций, необходимых при постановке задачи называется *взвешиванием целей*. Этому вопросу посвящена отдельная глава (глава 6). В этой главе приведено несколько примеров, в которых метод оценки целей находит самостоятельное применение.

Глава 4

АНАЛИЗ ОРГАНИЗАЦИИ

ОСНОВНЫЕ ПОЛОЖЕНИЯ

Работы над проблемами организации начались в 30-е — 40-е годы; они проводились в группах, состоявших из физиологов, инженеров-электриков, математиков и ученых-социологов. Было обнаружено, что многие организации имеют сходные черты. Например, многие расстройства нервной системы человека подобны повреждениям электрической системы управления оружием. Схемы нервной системы человека (рис. 4.1), изображенные биологами и физиологами, даже выглядят аналогично схемам электрических цепей.

Вскоре группы ученых, работавших в Кембридже, Массачусетском технологическом институте и в других местах, выяснили возможность разработки обобщенной теории организации и управления, которая опирается на различные научные дисциплины. Профессор Норберт Винер в 1948 году подытожил работу этих смешанных групп, включавших ученых различных специальностей. В своей книге «Кибернетика, или управление и связь в животном и машине» [74] он подчеркивал, что связь (или передача информации) и управление являются основными процессами, характеризующими любую организацию. Профессор Винер использовал информацию в качестве общего понятия, имея в виду любой символ или сигнал, который организация может использовать в своей деятельности. Под информацией может пониматься электрический импульс, химическая реакция или письменное сообщение, т. е. любое указание, которым руководствуется организация в своей работе*.

Таким образом, в кибернетике отмечается, что все организации, а именно: (а) организации, состоящие из клеток в организме; (б) организации, состоящие из машин в автоматических производствах или электрических системах связи, и (в) организации людей в социальных группах — подчиняются общим законам информации и управления в своей деятельности.

Часто можно анализировать промышленные или военные организации, даже очень сложные, с помощью тех же самых понятий информации и управления. Такой анализ дает возможность создавать информационную модель** (или модель управления) организации.

НЕКОТОРЫЕ ОБЩИЕ ЗАМЕЧАНИЯ ОБ ИНФОРМАЦИОННОЙ МОДЕЛИ

Информационная модель не является математической моделью; она не используется для точного предсказания результата или для расчетов. Обычно она представляется в форме схемы. Такая схема помогает собрать воедино сведения об организациях. Кроме того, она дает возможность

* Приведенное здесь определение понятия «информация» не является точным. Под информацией понимаются обычно сведения, уменьшающие неопределенность ситуации у получателя информации. При этом электрический импульс сам по себе не является информацией, но может быть использован в качестве сигнала, несущего информацию, если наличие этого импульса (или его отсутствие), в соответствии с принятой системой кодирования ситуаций, уменьшает неопределенность исхода какого-либо эксперимента. Химическая реакция также не является информацией, но может служить сигналом или являться объектом информации, если получателю сообщаются сведения о ее протекании. Письменное сообщение может содержать информацию даже тогда, когда организация не руководствуется им в своей работе. (Прим. ред.)

** К. В. Дейтч [31], см. также обзор развития кибернетических моделей у Дейтча [28].

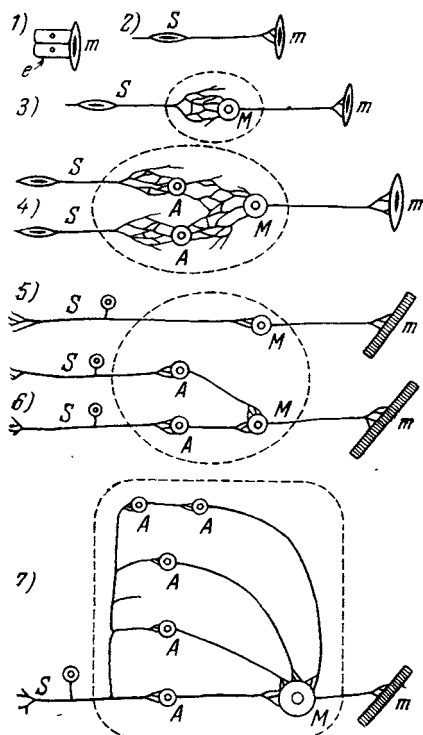


Рис. 4.1. Эволюция нервной системы по Бейлису [12, стр. 468]. Схемы центральной нервной системы, изображаемые физиологами, выглядят, как электрические схемы, применяемые инженерами-электриками. В общем случае любую организацию можно описать в виде таких взаимосвязанных частей.

Схема эволюции центральной нервной системы.

Пунктирными линиями показаны границы нервных центров. *S* — чувствительный (сенсорный) нейрон, *A* — ассоциативный нейрон, *M* — мотонейрон, *e* — эпителиальная клетка, *m* — мышечная клетка. 1. Губка. 2. Анемон. 3. Простейшая форма земляного червя. 4. Прослойка ассоциативных нейронов у земляного червя. 5. Исклчительно простая рефлекторная дуга у позвоночных. Обычно наблюдается в случае коленного рефлекса. 6. Обычная система позвоночных. Клетки чувствительных нейронов расположены в дорзальном отделе ганглиев спинного мозга, а не в органах рецепции, за исключением органа обоняния. 7. Добавление высших центров, состоящих только из ассоциативных нейронов, некоторые из которых заторможены. Если они есть, то образуют более длинные параллельные или переменные связи между рецепторными и эффекторными органами. Эти связи можно проследить на рис. 4.2.

приступить к решению организационных задач, отсортировать ценную информацию от несущественной, выбрать аналогии среди различных видов организаций и взять в качестве рабочей гипотезы одно из решений организационных задач. Такие предположения и указания часто бывают необходимы группе исследования операций, особенно в начале разработки нового проекта.

Поскольку информационные модели имеют такое важное практическое значение, мы будем уделять больше внимания их использованию, нежели обсуждению их теоретических особенностей. В связи с этим настоящая глава делится на 3 части:

1. Упрощенное теоретическое обсуждение информационных моделей.

2. Краткое описание практических способов построения информационных моделей.

3. Пути использования построенной информационной модели. (В конце главы приводится библиография по рассматриваемым вопросам.)

Информационную модель можно представить себе в виде рыболовной сети, паутины или нервной сети, по отдельным путям которой передается информация. В этой главе дается простая схема информационной модели как сети связи, приводятся различные характеристики организации, представленной в виде такой модели, потоки информации в модели, а также изменения этих величин во времени.

ХАРАКТЕРИСТИКИ ИНФОРМАЦИОННОЙ МОДЕЛИ

Подробно модель описывается в главе 7. Однако здесь важно отметить, что модель является миниатюрным или компактным представлением оригинала. Обычно в модели отражены наиболее интересные стороны оригинала, причем структуры модели и оригинала в этом отношении аналогичны. Кроме того, может быть приведен набор различных правил, разъясняющих работу модели.

Структура модели и интересующие нас характеристики модели изменяются

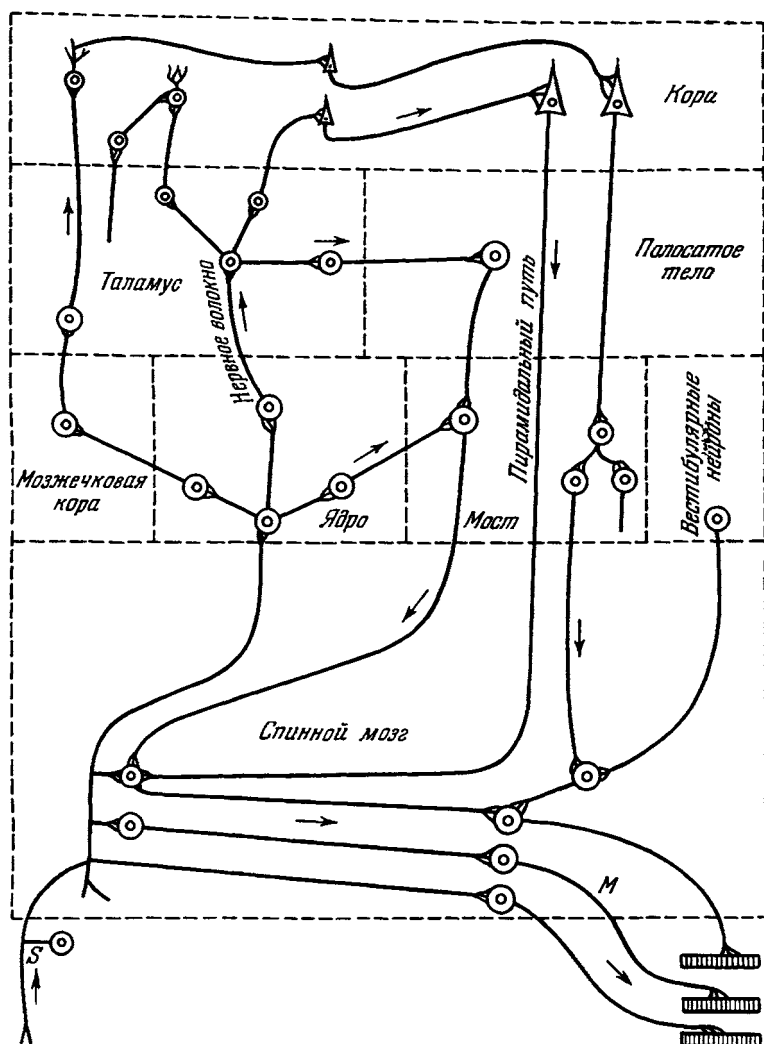


Схема центральной нервной системы млекопитающего по фон Монахову и Мотту. Показана развитая система ассоциативных нейронов, дающих параллельные или переменные связи между первичными сенсорными нейронами (S) и выходными общими трактами (M).

Рис. 4.2. Центральная нервная система млекопитающего по фон Монахову и Мотту [12, стр. 478]. Эта схема является развитием схем, показанных на рис. 4.1. Следует обратить внимание на увеличение сложности взаимосвязей, что объясняется более высоким уровнем нервной системы.

в соответствии с изменениями оригинала. Например, если дорога, ведущая из одного города в другой, закрыта или заброшена, то ее можно убрать из последующих изданий карты дорог [59].

В соответствии с этими положениями строится информационная модель. Чтобы модель была определена, должны быть известны: 1) Информационная сеть, существующая в данный момент времени (отображает интересующие нас пункты и их связи). 2) Рабочие процессы в сети (правила работы модели). 3) Изменения сети и процессов в ней со временем.

Например, физиолог может представить нервную систему в виде кружков и взаимосвязанных линий (как на рис. 4.1). Увеличение сложности организации нервной системы потребует увеличения количества нервных центров или изменения связей между ними (см. рис. 4.2). Аналогично происходит построение информационной модели.

СХЕМА ИНФОРМАЦИОННОЙ МОДЕЛИ

Организацию можно представить как группу элементов (подразделений в компании, работающих блоков в машине, людей в социальной группе), которые связаны между собой с помощью информации, т. е. писем, телефонных звонков, материальных потоков, личных бесед и т. д.

Если приведена схема, на которой показано, как осуществляется связь между различными элементами организации (например, показано движение нарядов на материалы внутри производственной организации, а также связь организации с внешним миром (например, приведена карта мест, в которые разосланы рекламы компании, и карта мест, из которых получены заказы), то можно представить, хотя бы частично, работу данной организации.

Информационная схема на бумаге выглядит, как карта дорог или как схема, приведенная на рис. 4.1. На рис. 2.2—2.5 показаны схемы потоков документов и материалов компании, поставляющей запасные части для машин. Линии представляют собой передачу различных документов или информации. Точки (или квадраты) представляют собой места, в которых информация используется, перерабатывается или запасается. Такая схема дает наглядное представление о том, насколько сложной является изучаемая организация. По схеме видно, как связаны между собой части организации. Поэтому при изучении организации надо прежде всего определить структуру информационной сети. Такая структура отображена на схеме информационной модели.

ВНУТРЕННИЕ ПРОЦЕССЫ В ОРГАНИЗАЦИИ: КАК ОНА УПРАВЛЯЕТСЯ

Работа организации — компании, частей машины, функциональных элементов человеческого организма — протекает в соответствии с информационной моделью, однако важно отметить следующую особенность: элементы организации работают совместно для достижения (или поддержания) внешней цели (или соответствующей ей подцели внутри организации) *.

* Здесь используется понятие цели, введенное Винером. Подробнее см. [57]. Приведем цитату из этой статьи. Если разделить поведение на активное и пассивное, то «активное поведение можно подразделить на два класса: бесцельное (или случайное) и целесообразное. Термин «целесообразное» означает, что действие или поведение можно рассматривать как направленное на достижение цели, т. е. конечного состояния, при котором объект, поведение которого изучается, достигает определенного соответствия во времени и в пространстве с другим объектом или событием. Бесцельным поведением считается поведение, которое не направлено на достижение цели». Далее в статье приводится важное ограничение в этом определении: «... мы ограничили

Для обсуждения информационных моделей (в упрощенном виде) цель можно определить как рабочий стандарт, используемый организацией в данный момент. Целью является определенное состояние, которого стремятся достичь в данный момент времени. Например, мастеру в цехе дается производственное задание на неделю; бухгалтерия определяет стандартные цены и т. д. Это сравнительно простые цели. Перед организацией могут стоять также более сложные цели или целый комплекс простых и сложных целей. Исходя из простоты или сложности целей и методов их достижения, можно классифицировать организации по их способности использовать информацию и решать свои задачи.

БЛОК ПРОСТОГО ПРЕОБРАЗОВАНИЯ

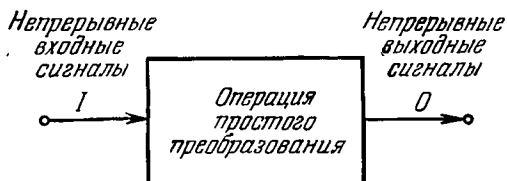


Рис. 4.3. Блок простого преобразования. Непрерывное функционирование возможно при наличии непрерывных входных сигналов. Собственная цель отсутствует. Пример: зубчатая передача.

Простейшая организация получает непрерывные указания от внешнего источника. У нее нет своих собственных целей, ей нужно все время указывать, что делать, ее нельзя оставить одну. Такие организации можно сравнить с простыми механическими или электрическими передаточными блоками (например, зубчатая передача, усилители и т. д.), представленными схематически на рис. 4.3. В таких блоках можно выделить три главные операции:

(1) прием, (2) переработка или преобразование и (3) выходное воздействие. В промышленности к таким моделям можно отнести, например, процесс превращения заказов в товары.



Рис. 4.4. Блок простой сортировки.

Два выхода соответствуют одному входу. Правила сортировки (или принятия решения) заложены в самом блоке. В блоке сортировки выполняются простые операции поиска и распознавания, свойственные более сложным процессам.

ПРОСТАЯ СИСТЕМА СОРТИРОВКИ

Другой простейшей организацией является сортировка, например сортировка фруктов или просеивание гравия. Принципы сортировки заложены в блоке при его создании; в таком блоке также необходим непрерывный поток информации от внешнего источника. Один вход (например, груз

гравия) может соответствовать двум или более различным выходам (например, различные размеры гравия). Простая организация такого типа представлена на рис. 4.4. Схема аналогична схеме рис. 4.3, но несколько сложнее. Наиболее известным практическим примером сортировки является операция сортировки на почте.

понятие телеологического поведения, используя данное определение только для целесообразных реакций, которые контролируются ошибкой реакции, т. е. разницей между состоянием объекта в любой момент времени и окончательным состоянием, которое считается целью. Таким образом, телеологическое поведение отождествляется с поведением, управляемым отрицательной обратной связью, и тем самым это определение выигрывает в точности».

Хотя в этой главе не обсуждается целесообразное поведение по сравнению с бесцельным поведением (или в философском смысле детерминизм сравнительно со случайностью), *но этот* вопрос является фундаментальным в развитии кибернетики. Для лучшего понимания этой главы следует обратиться к оригиналу [57]. Можно порекомендовать также ряд работ по этому вопросу: [22, 34, 58 и 69].

Следует отметить, что в блоке сортировки принимается решение, критерий которого заложен в самом блоке. Так, при сортировке гравия необходимо иметь ячейки различных размеров для просеивания.

ПРОСТОЙ РЕГУЛЯТОР

Простейшим типом самоуправляемой организации является организация, которая может управлять своей работой в зависимости от внешней цели. Такой блок получает задание и должен его выполнить. Примером чисто механического устройства такого рода является регулятор в паровой машине (рис. 4.5), который регулирует скорость двигателя при изменении нагрузки. Регулятору задается определенная скорость, которую необходимо поддерживать неизменной.

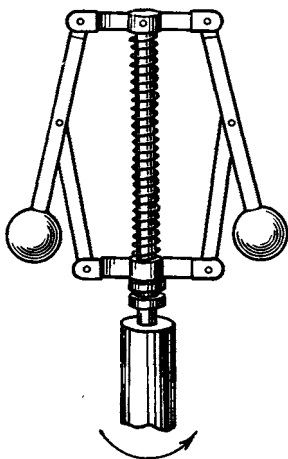


Рис. 4.5. Простое механическое управляющее устройство (регулятор), впервые исследованное Клерком Максвеллом. Регулятор поддерживает скорость парового двигателя неизменной при изменении нагрузки.

В общем случае можно сказать, что организация управляет своей деятельностью, если она сравнивает свою действительную работу с поставленной целью, выявляет ошибки и старается их устранить.

ОБРАТНАЯ СВЯЗЬ, НЕОБХОДИМАЯ ПРИ УПРАВЛЕНИИ

Чтобы организация могла определить, отвечает ли ее деятельность поставленной перед ней цели, необходимо ввести обратную связь с выхода на вход, чтобы сравнивать часть выходного сигнала с уставкой на входе. Если обратная связь уменьшает рассогласование, то она называется отрицательной обратной связью, потому что она противодействует* деятельности организации. Регулятор паровой машины является устройством с отрицательной обратной связью; в деловых операциях отрицательная обратная связь проявляется, например, в постоянном сравнении текущих расходов с запланированными ассигнованиями.

Точно так же можно с помощью отрицательной обратной связи объяснить выражение «придерживаться эталонов». В тех организациях, в которых определены такие эталоны, существует возможность сравнивать финансовое и социальное положение организации с эталонами, выявлять отклонения, которые организация стремится уменьшить.

Чтобы понять действие отрицательной обратной связи, рассмотрим пример из электротехники. На рис. 4.6 показана простая схема с обратной связью, используемая в следящих системах. Следящие системы используются, например, для приведения в движение антенны радара, чтобы ее направление соответствовало положению, заданному на удаленном пульте управления, независимо от сопротивления ветра.

* Это замечание неправильно. Отрицательная обратная связь не противодействует, а, как правило, содействует деятельности организации, поскольку она стабилизирует ее работу в условиях изменяющихся внешних воздействий и повышает ее динамическую устойчивость. Стабилизирующая обратная связь называется отрицательной оттого, что в установившемся режиме сигнал обратной связи направлен противоположно управляющему воздействию и сигнал на входе системы определяется их разностью. (Прим. ред.)

Уставка задается на пульте управления A , сигнал с которого приводит в действие двигатель или приводной механизм B , который воздействует на антенну C . Действительное положение антенны, которое может отличаться от заданного, например, из-за сопротивления ветра, подается в виде сигнала обратной связи на вход, и рассогласование между уставкой и выходным сигналом поступает в виде сигнала ошибки на приводной механизм, который действует в сторону уменьшения ошибки.

Работу этой схемы можно описать следующим выражением (рис. 4.7):

$$E_2 = E_1 \left(\frac{K}{1 - (-b)K} \right),$$

где E_1 — входной сигнал, или уставка, E_2 — выходной сигнал схемы, K — коэффициент усиления или механический коэффициент передачи и $(-b)$ — коэффициент обратной связи, соответствующий доле выходного сигнала, подаваемой на вход. В общем случае, чем больше сигнал отрицательной обратной связи, тем больше будет уменьшаться ошибка.

Вместо отрицательной обратной связи можно ввести положительную обратную связь $(+b)$. Тогда ошибка будет увеличиваться, в системе возникнут колебания и т. д. Инженер, разрабатывающий систему управления, должен интересоваться такими вопросами, как уменьшение ошибки, возникновение колебаний, устойчивость системы. В данной главе нет возможности серьезно рассмотреть характеристики обратной связи, однако следует указать, что для использования обратной связи в информационных моделях необходимо познакомиться с приведенной в библиографии литературы по этому вопросу.

Из рис. 4.6 видно, что в системе управления образуется замкнутый контур воздействий. Этот контур состоит из цепи обратной связи и приводного механизма. Такие системы можно назвать «целенаправленными» системами или «приданными» системами, учитывая их замкнутость через цепь обратной связи.

В схемах информационных моделей следует выявлять наличие таких замкнутых контуров, чтобы определить имеющиеся в изучаемых организациях обратные связи и управляющие воздействия. Операционисты должны знать, какие процессы управляются, какие — нет, какова эффективность

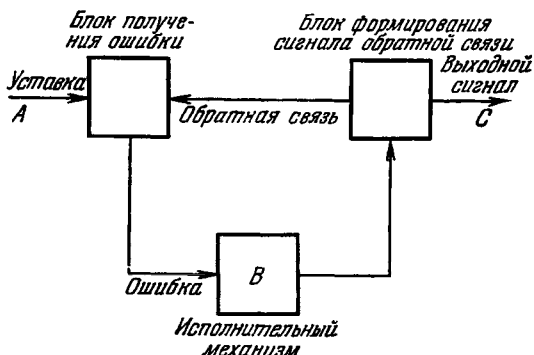


Рис. 4.6. Принципиальная схема отрицательной обратной связи.

Простейшая организация, которая обладает способностью к саморегулированию; важно отметить наличие замкнутого контура. Уставка (цель) задается на блоке A , затем эта уставка поддерживается с помощью цепи обратной связи. Аналогично работает регулятор парового двигателя.

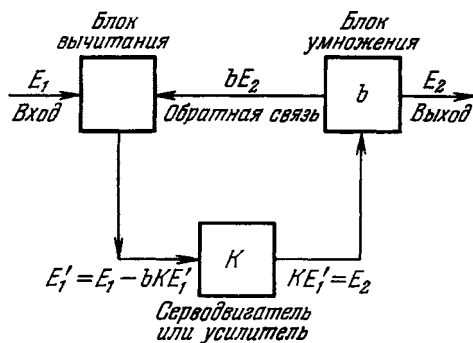


Рис. 4.7. Простая система с отрицательной обратной связью; приведены математические соотношения между величинами, описывающие работу системы.

цепей обратной связи, каков характер обратной связи, каковы граничные значения коэффициентов отрицательной (положительной) обратной связи, при которых обратная связь может быть полезной или вредной. Системы планирования, например, следует анализировать с точки зрения устойчивости, времени запаздывания, контроля при помощи обратных связей.

СОРТИРОВКА С ОБРАТНОЙ СВЯЗЬЮ

Обратную связь можно ввести как в систему простого механического преобразования (подобно регулятору скорости в паровой машине), так и в систему простой сортировки. При этом выходы с блока сортировки сравниваются со стандартными величинами, чтобы определить, правильно ли работает блок сортировки. Результаты такого сравнения дают возможность

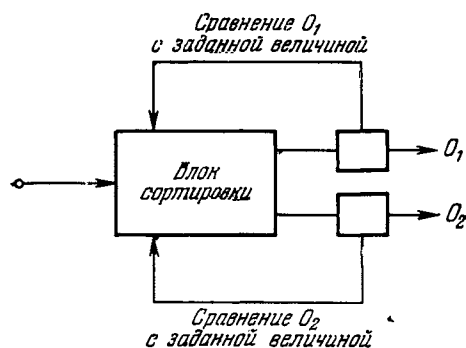


Рис. 4.8. Блок простой сортировки с обратной связью. Фактический выход сравнивается с заданным, заложенным в блок.

улучшать характеристики системы — ее устойчивость и качество сортировки. Блок-схема системы показана на рис. 4.8. Примером системы сортировки с обратной связью является, например, система проверки качества готовой продукции в промышленности, где происходит сравнение со стандартными образцами и сортировка изделий на годные и бракованные.

СОЧЕТАНИЕ ПРЕОБРАЗУЮЩЕГО И СОРТИРУЮЩЕГО БЛОКОВ

Для получения более сложной и более подвижной структуры организации можно комбинировать различным образом блоки преобразования

и сортировки (с обратной связью или без нее). Внутри организации возможны различные комбинации отделов и групп. Однако определить наиболее удачную комбинацию не всегда возможно, так как число таких комбинаций и вариантов внутри крупной организации очень велико. Кроме того, объединение различных частей может привести к получению характеристик, совершенно отличных от исходных свойств этих частей, особенно в промышленных и общественных организациях. Когда профессора Винера его коллеги-социологи убеждали распространить математическую теорию кибернетики на область взаимоотношений людей в организации, он долго колебался, так как прекрасно сознавал, насколько трудно получить точные результаты ввиду быстрых изменений в социальных организациях, наличия только кратковременной статистики и невозможности проведения абсолютно объективных наблюдений.

Приведем выдержку из [74, стр. 191]. В социальной науке «мы вынуждены оперировать с малыми объемами выборок; к тому же мы не можем быть уверены, что значительная часть наших наблюдений не является нашим вымыслом. Так, исследование работы фондовой биржи, вероятно, даст совершенно неверную картину. Мы слишком легко настраиваемся на наблюдаемые нами явления, чтобы быть беспристрастными наблюдателями. Короче говоря, какими бы ни были наши исследования социальных наук — статистическими или динамическими, а должны проводиться и те и другие, — они никогда не будут достаточно хорошими, за исключением отдельных

случаев, и никогда не дадут нам достаточного количества важной информации, которую можно проверить, что характерно для естественных наук. Однако мы не можем пренебрегать такими исследованиями, и в то же время не следует ожидать от них слишком многого. Хотим мы этого или нет, все же многое придется оставить на долю ненаучного описательного метода профессионального историка».

Если же исследователь знает обо всех этих трудностях и стремится хотя бы к некоторым усовершенствованиям (что часто и имеет место), то для такого исследователя будет полезно ознакомиться с описанием сложных организаций, состоящих из простых элементов.

СИСТЕМА С АВТОМАТИЧЕСКИМ ИЗМЕНЕНИЕМ ЦЕЛИ

Если организация может действовать различным образом, причем существуют правила, на основании которых она может выбрать, как ей поступать при изменении внешних условий (т. е. может предвидеть наилучший способ действий при изменении условий), то такая организация

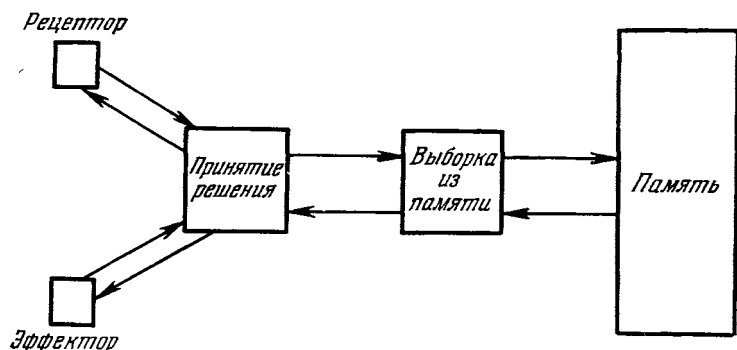


Рис. 4.9. Схема с обратной связью и памятью.

Добавление памяти и более сложных цепей обратной связи дает организации возможность лучше управлять своей деятельностью. В систему закладываются различные возможные варианты действий в зависимости от изменения внешних условий (обнаруживаются с помощью рецептора). Примером является автоматическое переключение на телефонной станции.

может управлять своей деятельностью более эффективно, чем простая система с обратной связью. В такой организации необходима обратная связь второго порядка, к тому же требуется память для запоминания возможных вариантов.

Примером такой организации, которая может менять свои цели, чтобы выполнять различные действия, является центральная телефонная станция. Непосредственной целью телефонной станции является поиск определенного номера, набранного абонентом. За день может быть набрано много таких номеров; станция должна быть готова получить различные номера и для каждого из них автоматически выбрать нужную последовательность операций (на рис. 4.9 показана упрощенная схема такой системы, которая фактически является более сложной системой сортировки). Другим примером изменения цели является погоня кота за крысой. Кот не повторяет всех положений крысы в каждый момент времени, но руководствуется своей памятью о поведении других крыс в прошлом.

Если организация может управлять своими действиями, особенно если она может изменять свои цели, то такую организацию называют *автономной*

организацией. Автономность организации с автоматическим изменением цели заключается в ее памяти и способности вспоминать запомненное. Чем лучше память и чем быстрее происходит выборка из памяти, тем более автономной является организация.

Накопление информации, на основании которой организация может выбрать план действий, является процессом обучения. Обучение может привести к перестройке внутренних связей организации или информационной сети. Процесс обучения организации меняется со временем. Например, сети в телефонной центральной станции могут быть расширены для включения «номеров» других абонентов за счет перестройки связей на станции. Увеличение объема памяти обычно влечет за собой усложнение связей в информационной сети. В физиологии это означает, что вводятся новые связи между нервными клетками. В библиотеке увеличение объема памяти означает увеличение связей между справочными индексами в каталогах.

После того, как мы выяснили, что представляют собой связь и управление в системе с автоматическим изменением цели, можно поставить следующие вопросы: как же изменяются эти процессы во времени? Как развиваются внутренние информационные каналы в организации? Как они отмирают? Как поддерживаются? Где расположена память организации? Какой вид информации помещается в память? Каким образом информация хранится? Какой вид информации выбирается из памяти? Каково содержание памяти, как оно меняется? Обучается ли организация чему-нибудь? Правильно или неправильно происходит утрата (стирание) информации? Что может организация предсказать, основываясь на своей памяти?

Когда рассматривается деятельность организации с памятью, необходимо учитывать ценность информации, поступающей в память и извлекаемой из памяти. В различных ситуациях различные способы действий имеют разную ценность, поэтому для понимания работы системы исследователю важно знать ценность или приоритеты отдельных вариантов решения задач. Для примера сошлемся вновь на телефонную станцию. Когда получены сразу десять телефонных вызовов, то необходимо решить, на какой вызов ответить в первую очередь.

СИСТЕМА С СОЗНАТЕЛЬНЫМ ИЗМЕНЕНИЕМ ЦЕЛИ

Если организация может собрать информацию, запомнить ее, а затем переработать содержимое памяти с целью выработки нового образа действий, то такая организация достигла нового уровня автономии. Механизм, в задачу которого входит рассмотрение различных целей и способов действия, можно назвать сознанием (разумом) организации *. В таких системах с обратной связью третьего порядка происходит сознательный (обдуманый) процесс принятия решения. Деятельность такой организации больше всего приближается к нашему представлению о настоящей промышленной или общественной организации (см. рис. 4.10).

Чтобы показать конкретно, что же такое сознание, представим себе человека, который, сидя в мягком кресле, размышляет, чем бы ему заняться — закончить исследование прежде, чем продать свое изобретение, или же по просьбе жены заменить в умывальнике неисправный кран. Он решил доставить удовольствие жене в ущерб своему бумажнику. В процессе такого принятия решения участвует система с сознательным изменением цели, т. е. сознание.

* В частности, см. работы Дейтча о росте и обучении (например, [31, гл. 8]).

Сознательное обучение может быть выборочным, при котором из большого объема внешней информации выбирается такая информация, которая необходима для выживания организации или для других важных целей.

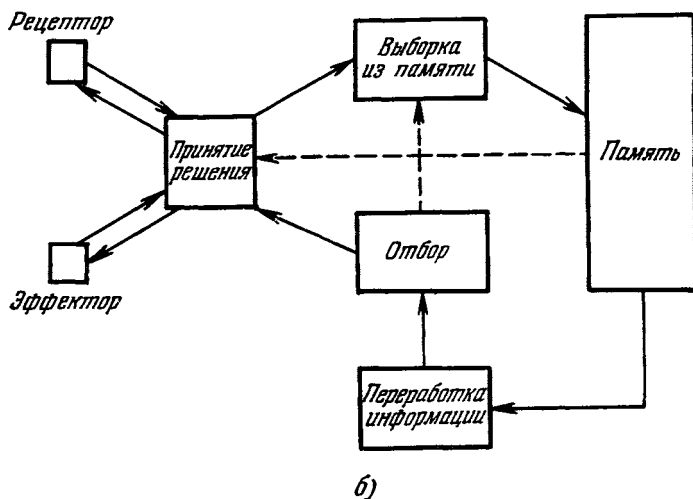
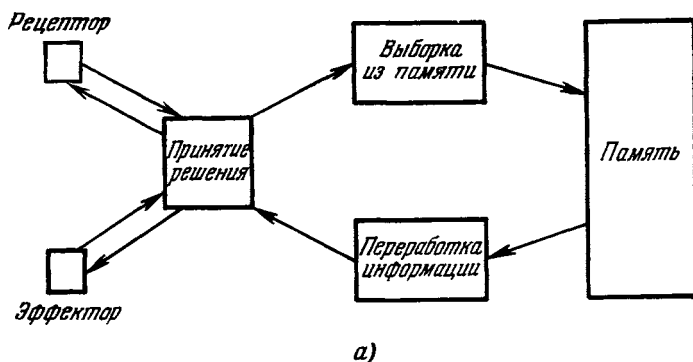


Рис. 4.10. а) Усовершенствованная схема с памятью

Если информацию в памяти можно переработать (на машине или в самой организации) с целью формирования новых стратегий, то такая организация становится более гибкой и автономной. В такой системе возможно простое предсказание.

б) Дальнейшее усовершенствование схемы с памятью: введение сознания.

Если можно комбинировать различную информацию в памяти, а затем из этих комбинаций отбирать некоторые для дальнейшей оценки формирования новых комбинаций и т. д., то организация достигает еще большего уровня универсальности или автономии. Пунктирные линии означают сравнение деятельности организации в настоящее время с прошлой деятельностью и с возможной будущей деятельностью (предсказание второго и третьего порядка). Во многих организациях такие сравнения почти не производятся.

Сознание может изменить направление внимания организации: заставить организацию сосредоточиться на каких-то одних явлениях и не обращать внимания на другие. На основе сознания с учетом входной информации происходит формирование одной последовательности действий (стратегии) и отказ от другой. Сознание может также анализировать структуру организации, исследовать ее память, выявить несоответствие между различными действиями и целями, к которым они направлены. Здесь перечислены только

несколько функций сознания, которое можно считать управляющим центром третьего порядка.

Организация, имеющая такой центр, может управлять своим собственным ростом. Способность распознавать ценность полученной организацией информации или ценность различных комбинаций информации, содержащейся в памяти, дает организации возможность вводить различные новшества. Эти свойства очень важны для многих организаций, поэтому специалист по исследованию операций заинтересован в том, чтобы иметь четкое представление о таких организациях (т. е. знать, что делают и чего не делают руководители организации).

Для вычислительной техники также очень важно получение таких схем с сознательным изменением цели. Например, для того чтобы вычислительные машины устраняли свои неисправности, необходимо, прежде всего, чтобы машина могла понять (осознать), какие нарушения произошли внутри нее. «Сознание» такой машины действовало бы примерно следующим образом: управляющий центр (сознание) определяет, какие элементы организации (например, блоки или лампы) вышли из строя или умышленно удалены. Затем такой центр управляет заменой неисправных или устаревших элементов новыми или усовершенствованными. Для вычислительных машин это недостижимо, однако в промышленных организациях такие операции проводятся ежедневно!

Можно было бы ожидать, что сознание укажет все нарушения, которые можно найти в управляющих органах, действующих в организациях, например: заблуждения, неправильные указания, неправильное толкование сообщений, непонимание новых возможностей, нечеткое определение рабочих целей и другие. Эти ошибки будут рассматриваться во второй половине этой главы. Однако здесь уместно привести один пример.

Представим себе вычислительную машину, которая может исправлять свои неисправности. Для руководства ремонтом имеются управляющие схемы (сознание). Но если нарушения произойдут в самих управляющих схемах и поступит неправильный приказ о ремонте исправной машины, то машина может выйти из строя. Результат был бы таким же, как если бы пьяный техник на телефонной станции случайно распыл какие-то соединения реле. В этом случае будет фактически невозможно обнаружить все новые неисправности. Таким образом, ненадежность современных электронных элементов и схем препятствует созданию в настоящее время вычислительных машин, обладающих «сознанием», т. е. способных исправлять свои неисправности. Аналогично промышленные организации могут выйти из строя, если руководители (элементы этой организации) не выполняют своих обязанностей (ненадежны).

Группа исследования операций должна извлечь как можно больше пользы из теорий сложных организаций, изучаемых кибернетикой, с учетом памяти и сознания. Одной из задач комплексной группы исследования операций является увеличение памяти организации с помощью сбора различной информации о работе организации. Кроме того, эта группа должна помогать руководителям (которые представляют сознание организации) в разработке и оценке различных направлений деятельности организации.

КОМБИНИРОВАННАЯ ИНФОРМАЦИОННАЯ МОДЕЛЬ

На рис. 4.11 показана комбинированная модель организации, в которой объединены различные схемы связи и управления в организациях. Эта схема была предложена К. В. Дейтчем в качестве общей информационной

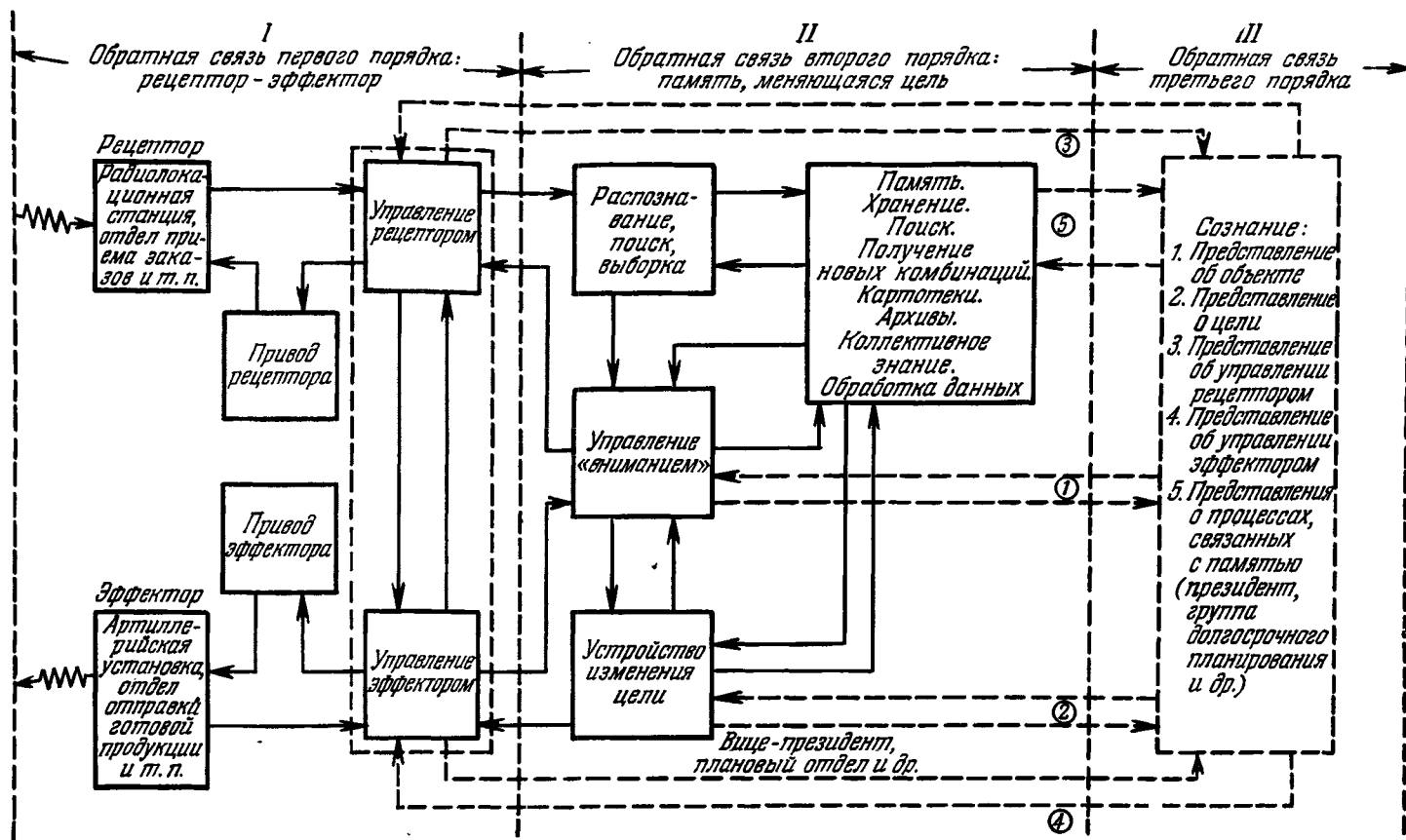


Рис. 4.11. Блок-схема системы управления с обратной связью третьего порядка.

Упрощенная схема из Дейтча [30]. Объяснения см. в тексте. Эта схема состоит из более простых блоков, показанных на рис. 4.3—4.10.

модели, которую можно использовать для описания сложных организаций *. Для определенности будем считать, что это блок-схема входного радарного устройства в системе управляемого орудия с блоком памяти.

В области *I* рис. 4.11 расположена схема простой системы с обратной связью и фиксированной целью. Схема разделяется на блок рецептора (воспринимающее устройство) и блок эффектора (исполнительное устройство), или, иначе, на радарное оборудование по обнаружению самолетов и механизмы наведения орудия и ведения огня. Когда с помощью рецептора самолет обнаружен, то эффекторное устройство изменяет положение исполнительных механизмов в соответствии с положением самолета или цели, а затем следит за положением самолета как можно точнее.

В области *II* расположены память и блок изменения цели, с помощью которых можно предсказать будущее положение самолета, чтобы не следовать все время за самолетом, а упреждать его движения и тем самым обеспечить большее число попаданий. В этой области расположено устройство с автоматическим изменением цели, потому что в устройство заложены правила, на основании которых происходит изменение цели. Работа системы управления орудием меняется (самим устройством) в зависимости от типа самолета, погоды, предполагаемых качеств летчика и т. д.

В области *III* расположен блок сознательного изменения цели, или сознание. Это только эскиз такого блока, приведенный для того, чтобы читатель мог проследить развитие системы от простейших схем типа рецептор — эффектор до сложных схем с обратными связями типа блока сознания. Этот блок обозначен пунктирной линией, так как в современных вычислительных машинах такие схемы отсутствуют.

Для сравнения на рис. 4.11 приведена аналогия с промышленными организациями. Так, область *I* соответствует зависимости «производство — заказы», т. е. полученные заказы выполняются на производственном оборудовании. Область *II* представляет собой управленческий аппарат организации, архивный отдел, а также вице-президента, который обладает полуавтоматическими или тактическими возможностями изменения цели. И наконец, область *III* соответствует операциям долгосрочного планирования, которое выполняет президент компании или правление директоров.

Цель приведенного описания некоторых характеристик внутренних связей и систем управления в организациях заключается в том, чтобы показать, насколько тщательно надо воспроизводить схему организации для выявления важных аналогий. В информационных сетях общепринято изображать схемы входных и выходных устройств, а также схемы переработки информации, показанные соответственно в левой и правой частях.

Подытожим все сказанное об информационных моделях

ЗАКЛЮЧЕНИЕ. ХАРАКТЕРИСТИКИ ИНФОРМАЦИОННОЙ МОДЕЛИ

С помощью информационной модели можно получить следующие сведения.

1. Схему информационной сети организации.
2. Представление о процессах управления, связанных с обеспечением стабилизации определенной цели или направленных на достижение новой цели.

* Рис. 4.11 взят в упрощенном виде из книги Дейтча [30]. Можно доказать применимость схемы Дейтча для любого уровня социальной организации, включая и одного человека.

3. Данные о процессах управления изменением цели в более сложных организациях. Кроме того, модель отображает процессы нововведений, роста, обучения, дает представление о задачах памяти, сознания и понятие об автономии.

Во всех этих вопросах каждый операционист всегда интересуется видом или содержанием информации, которая передается и получается.

Таким образом, информационная модель представляет собой сложную схему, подобную схеме рис. 4.11 (внутренние связи которой могут меняться со временем), дающую сведения о процессах связи и управления, полученные из различных дисциплин. С помощью такой схемы эти данные могут быть объединены.

КАК ПОСТРОИТЬ СХЕМУ МОДЕЛИ

При создании информационной модели прежде всего надо начертить схему потоков информации. Предлагались различные методы отображения потоков информации или связей между отдельными людьми в группах или между более крупными элементами или отделами в организации. Мы покажем некоторые из этих методов. Прежде всего, операционист больше интересуется источниками получения и использованием информационных потоков, а не материальным воплощением связи.

Сбор информации (опросы, беседы). Цель анализа информационной сети — установить, кто с кем связан и в какой степени. С этой целью задаются вопросы: от кого вы получаете приказы? С кем вы разговариваете чаще всего?

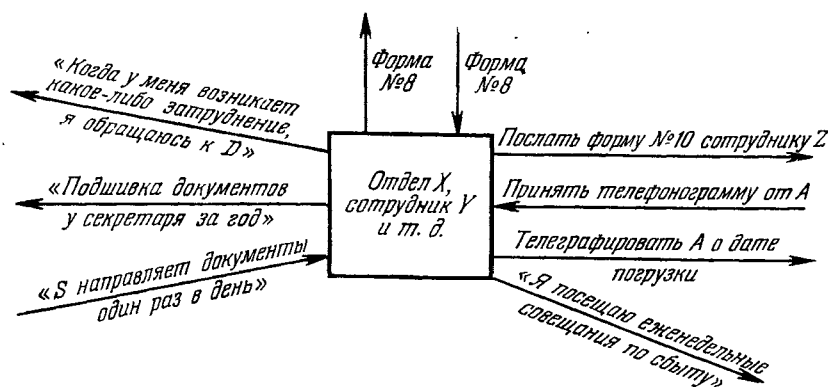


Рис. 4.12. Предлагаемый метод регистрации данных, полученных путем опроса сотрудника организации, выборки из документов или сообщений, а также из непосредственного наблюдения взаимосвязей в организации.

Подробности см. в тексте.

С кем вы советуетесь, когда принимаете решения в своей работе? Откуда приходят документы, с которыми вы работаете? Куда они передаются?

Очень важным является вопрос о направлении потоков информации; поэтому следует внимательно следить за источником сообщений и пунктом назначения. Часто сведения о конкретных видах связи более эффективны, чем более обширная информация общего характера [52]. Например, полезны сведения о потоках определенных заказов, например на материалы, документов по продаже и т. д. (Некоторые практические сведения о сборе информации см. у Маркоссона [46].) После изучения связей каждого члена

организации или какого-либо элемента организации следует подготовить перечень форм, писем или других важных видов информации, с которыми связан этот член организации. Затем для каждого члена организации, отдела и других элементов организации составляется схема, аналогичная показанной на рис. 4.12.

Непосредственное наблюдение. Когда изучаются очень маленькие или очень большие организации, исследователи могут предпочесть беседам непосредственное наблюдение. Опросы могут нарушить информационный процесс в малых группах *, а в больших группах с их помощью иногда нельзя обнаружить некоторые виды связи. Кроме того, наблюдения являются хорошим средством проверки правильности ответов, полученных при опросах.

Профессор Оскар Моргенштерн из Принстонского университета приводит пример большого склада военно-морского флота США, работу которого он изучал в Бруклине (штат Нью-Йорк) **. На складе хранилось от 1 до 2 миллионов изделий (почти в десять раз больше, чем на складах компаний «Мэйси» и «Дженерал моторс»). Целью Моргенштерна было определить, как эти товары прибывают на склад и как контролируются запасы. Из распросов выяснилось, что большинство товаров прибывает поездом, и Моргенштерн начал изучение информационной сети, соответствующей железнодорожным операциям. Однако, проехав несколько дней на маневровом паровозе по сортировочной станции, он установил, что, хотя громоздкие грузы, которые больше бросаются в глаза, действительно прибывают по железной дороге, более мелкие грузы, которых гораздо больше и с которыми больше хлопот, доставляются автотранспортом. Поэтому Моргенштерн переключился на изучение операций, связанных с автотранспортом. Ясно, что без непосредственного практического наблюдения эти операции могли бы пройти незамеченными или им не было бы уделено должного внимания. Аналогично непосредственное изучение документов, переходящих из отдела в отдел, в больших компаниях может быть полезным для построения информационной сети.

Измерения. Подобно измерениям токов и напряжений в электрических цепях, в социальных группах были проведены эксперименты *** по количественному измерению «взаимодействия» в таких группах, для чего были созданы специальные средства. Целью таких измерений по-прежнему является ответ на вопрос: кто с кем общается? Данные исследования взаимодействия можно представить графически подобно данным, полученным из опросов (рис. 4.12).

Выборка. В любом из вышеперечисленных методов сбора данных можно использовать способ выборки, чтобы на основе небольшого числа данных построить схему информационной сети; кроме того, иногда можно пользоваться этим методом для повышения точности анализа. Например, можно

* Формы наблюдения, которые отображены на рис. 4.12, приняты в *социометрии и анализе взаимодействия*. Обзорная статья по этому вопросу приведена в книге Ягоды, Дейтча и Кука [40, стр. 562—585]. См. также работы Чэпла [19], Бейлса [7, 8] и Бавеласа [11].

** Informal Seminar Discussions, Industrial Engineering, 312, Columbia University, New York, Spring, 1953.

*** См., например, Бейлс [8], Чэпл [7], Кристи, Люс и Мэйси [20]. Дейтч [33] перечисляет 14 видов измерений, которые можно проделать для описания организации. Многие из этих измерений в той или иной форме приведены в данной главе.

было бы в течение месяца на коммутаторе компании определять количество телефонных звонков между отделами компании. Точно так же, чтобы получить представление о потоках письменной информации между отделами компании, совсем не обязательно подсчитывать все проходящие документы, для этого достаточно взять выборку из этих документов. (Подробности относительно этой процедуры см. в главе 20.)

Разница между постоянными и непостоянными (случайными) связями. Постоянные связи (потоки информации) характерны для обычной работы организации. Такие связи легко выявить с помощью опросов или методов выборки. Однако представляют большой интерес случайные «важные» связи. Например, с помощью вопроса: «К кому вы обращаетесь в критических ситуациях?» — можно обнаружить редко используемые, но важные каналы связи.

ОКОНЧАТЕЛЬНАЯ СХЕМА ИНФОРМАЦИОННОЙ СЕТИ

Обычно схему информационной сети получают за несколько этапов. На основании собранных данных получают ряд цифр, которые затем соответственно обрабатываются. Кроме того, необходимо иметь схемы движения документов в целом по всей организации. Схемы, отображающие отдельные участки организации, располагаются на общем чертеже в наиболее подходящем месте. Затем проводятся линии, связывающие источники и пункты приема информации. В результате получается схема, аналогичная рис. 2.1—2.5 или, в более общем виде, подобная схеме рис. 4.12.

Один или два человека могут построить схему информационной сети (включая сбор данных) за период от 2 недель до 3 месяцев в зависимости от размера и сложности изучаемой организации.

ПРОВЕРКА СХЕМЫ

После построения схемы следует проверить ее с помощью сотрудников изучаемой организации. Для этой цели могут оказаться полезными следующие вопросы: как вы считаете, правильна ли эта схема? Не замечаете ли вы каких-нибудь явных ошибок или пропущенных связей? Такие вопросы часто помогают выявить ошибки или пропуски, которые попали в окончательный вариант схемы. Однако к таким исправлениям следует относиться осторожно и проверять, не являются ли замеченные «ошибки» следствием неправильного представления самого сотрудника.

СХЕМА ИНФОРМАЦИОННОЙ СЕТИ В КАЧЕСТВЕ НАГЛЯДНОГО ПОСОБИЯ ПРИ РЕШЕНИИ РАЗЛИЧНЫХ ЗАДАЧ ОРГАНИЗАЦИИ

При обсуждении стоящих перед организацией задач необходимо достичь соглашения между группой исследования операций и членами организации. При этом может помочь схема информационной сети, на которую можно ссылаться при обсуждениях и для объяснений. Обсуждать различные задачи гораздо удобнее, имея перед собой схему, на которой можно что-то указывать, чертить и исправлять. Часто разные люди воспринимают одни и те же словесные объяснения по-разному. С помощью схемы можно избежать такого неоднозначного понимания.

КАК ИСПОЛЬЗОВАТЬ СХЕМУ ИНФОРМАЦИОННОЙ СЕТИ?

Чтобы понять методы построения схем информационных сетей, следует ознакомиться с использованием этих схем. Поэтому рассмотрим, как операторы применяют эти схемы и какие ошибки можно выявить с их помощью.

ВЫБОР ЗАДАЧ И УСТАНОВЛЕНИЕ СВЯЗИ МЕЖДУ НИМИ

Построение схемы информационной сети часто приводит исследователей к пониманию специфических задач, с которыми сталкивается организация. Исследователи могут решать эти задачи с помощью приведенных в этой книге специальных методов (модели запасов, теория очередей, модели поиска и т. д.), а также связывать эти задачи между собой, прослеживая по информационной сети взаимосвязи между различными процессами в организации.

ИЗУЧЕНИЕ СРЕДСТВ СВЯЗИ

В этой главе рассматривается нормальное состояние организации, т. е. исключаются аварийные ситуации. Никакие явные отклонения от нормы не учитываются. И все же мы хотим знать, работает ли организация в соответствии со своими возможностями. Какие улучшения можно сделать? Какие важные задачи можно поставить? Каковы слабые места в системе, которые могут вызвать нарушения в работе системы в аварийном режиме? Исследователи часто пренебрегают изучением каналов связи при отсутствии аварийных режимов, однако часто даже простой анализ средств связи дает возможность внести значительные улучшения в работу системы при минимальных затратах.

Приведем пример простой коммерческой операции — политики сбыта аппаратов «Флексорайтер» (Flexowriter). Это модифицированный телетайпный аппарат, который работает, как обычная пишущая машинка, и одновременно пробивает перфоленту.

Предположим, что компания получает заказ на какое-то издание. Заказ можно отпечатать на этом аппарате (с адресом заказчика, величиной заказа и другими данными). Печать сопровождается автоматической пробивкой перфоленты. Затем эту ленту можно использовать по всему предприятию для получения счетов, накладных на складе, документов по погрузке и других. Таким образом, все эти документы будут получены автоматически, в то время как раньше они заполнялись вручную. Следовательно, имеем значительную экономию во времени и в рабочей силе.

Такого рода усовершенствования относятся к средствам связи. Для того чтобы продемонстрировать это обстоятельство, фирма, заинтересованная в продаже аппаратов, может представить две схемы информационной сети организации: схему без использования телетайпного аппарата и новую схему с использованием аппарата и значительным сокращением объема канцелярской работы. Сравнение этих двух схем позволяет рассчитать ожидаемое сокращение расходов*.

Обе вышеупомянутые проблемы относятся к области исследования операций, однако поскольку специальные методы решения таких задач приведены более подробно в остальных главах книги, то в данной главе рассматриваются обобщенные проблемы связи в организации.

* Анализ системы связи еще более интересно провести с учетом использования крупных вычислительных машин в процессе обработки и выдачи документов [21].

НЕКОТОРЫЕ ЗАМЕЧАНИЯ О НЕИСПРАВНОСТЯХ СИСТЕМЫ

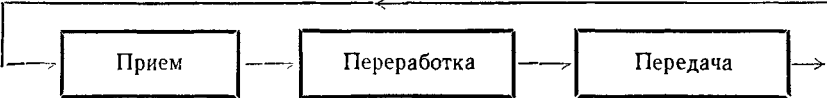
Из таблицы 4.1 можно получить общее представление о неправильной работе информационной системы. Показаны различные варианты трех операций в организации: (1) прием, (2) переработка и (3) передача информации. Перечислены также наиболее вероятные исполнители этих операций и некоторые общие ошибки, связанные с каждой из операций [43].

Задача заключается в том, чтобы обнаружить, какие нарушения возможны в информационных системах, определить, как можно описать эти нарушения на языке информационных систем, и найти общие ошибки для трех процессов. Для справок можно пользоваться таблицей 4.1.

Далее возникает задача определения причин нарушений и создания оптимальной организации для выполнения поставленной задачи. Хотя оптимальная для данной цели организационная или информационная структура неизвестна, однако в настоящее время в этой области проводятся

Таблица 4.1

Схема неисправностей в системе

			
Виды операций	1. Прием 2. Наблюдение 3. Надзор 4. Выборка 5. Экспериментирование методом проб и ошибок 6. Экспериментальная установка 7. Стимул	1. Внутренняя переработка информации 2. Обдумывание 3. Установление соответствия 4. Планирование эксперимента 5. Выдвижение гипотезы или плана 6. Построение окончательного плана 7. Организация—структура	1. Передача 2. Деятельность 3. Исполнение 4. Проведение эксперимента 5. Проверка гипотезы 6. Создание новой установки 7. Получение результата
Функциональные элементы, выполняющие операции	1. Шпионы 2. Репортеры 3. Антенна, детектор 4. Исследование спроса и предложения 5. Входные перфоленды или перфокарты 6. Радиолокационная установка	1. Штабные офицеры разведотдела и др. 2. Редакторы 3. Внутренние радиосхемы 4. Сотрудники, обрабатывающие заказы, плановики, руководители 5. Внутренние схемы вычислительного устройства, память, сортировка и др. 6. Вычислительное устройство орудия	1. Младший командный состав 2. Издатели 3. Громкоговоритель 4. Рабочие, мастера 5. Буквопечатающие аппараты, перфораторы, табуляторы и др. 6. Привод устройства управления огнем

Продолжение табл. 4.1

Обычные ошибки	1. Неправильно ориентированное внимание 2. Неправильно собирается информация. Формы громоздкие и т. д. 3. Информация собирается слишком медленно или слишком быстро 4. Прием обладает слишком малой или слишком большой чувствительностью 5. Приемник нестабилен 6. Нет обратной связи в приемнике 7. Внешние помехи 8. Приемник искажает содержание информации	1. Ошибки при распознавании 2. Неправильное хранение информации, недостаточная скорость запоминания, поиска и т. д. 3. Информация обрабатывается слишком быстро или слишком медленно 4. Каналы связи неисправны, перегружены, конкурируют между собой 5. Искажения содержания информации при обработке 6. Сознательный центр организации неисправен, недостаточен эффективен и т. д. 7. Элементы организации (системы) конфликтуют между собой, связаны циклически и т. д. 8. Цели неправильно поставлены, плохо определены, противоречивы и т. д. 9. Цель жесткая, ее нельзя изменить в зависимости от входящей информации 10. Плохое «обучение» 11. Отдельные задачи перекрываются, дублируются и т. д. 12. Блок изменения цели не влияет на процессы в организации	1. Эффектор (исполнительное устройство, передатчик) неправильно ориентирован 2. Передатчик нестабилен 3. Выход передатчика не соответствует окружающим условиям 4. Нет обратной связи в передатчике 5. Реакция эффектора слишком медленная или слишком быстрая 6. Эффектор обладает недостаточной или чрезмерной мощностью 7. Внешние помехи
----------------	--	--	--

широкие исследования. Обычно для определения наилучшей структуры информационной сети для определенной задачи отбирают несколько стандартных структур и сравнивают их с помощью пробных испытаний. Аналогично повреждения, обнаруженные в стандартных информационных сетях, помогают объяснить причины многих нарушений в организациях, ориентированных на выполнение определенной задачи или цели. Процесс экспери-

ментальной проверки структуры или выявления причины неисправностей легче всего проходит на низших уровнях обратных связей организаций, поскольку здесь цели четко определены, циклы непрерывно повторяются и т. д.; например, производственное задание выдается на неделю, и 1000 одинаковых изделий выпускается ежедневно.

Пример. Чтобы продемонстрировать роль анализа организации, рассмотрим случай, когда основные недостатки организации можно выявить с помощью такого анализа.

Рассмотрим компанию, выпускающую мощные двигатели, которые являются основными элементами автомашин, моторных лодок и различных промышленных установок. Компания является ведущей в этой области. Она существует приблизительно 35 лет и насчитывает около 3000 работников.

Чтобы выдержать конкуренцию, компания была вынуждена сократить сроки поставок. В результате персоналу, наблюдающему за ходом работ, и руководству пришлось потратить много времени для ускорения работ в различных подразделениях, включая сборочную линию. В конце концов положение ухудшилось настолько, что значительная часть деталей стала «дефицитной», т. е. необходимо было отдавать предпочтение то одной, то другой детали, чтобы уложиться в сроки поставки. В результате выполнение заказов на двигатели в заданные сроки было поставлено под угрозу.

Руководство полагало, что если бы запасы деталей были достаточны и хорошо укомплектованы, то основная причина задержек и нехватки деталей была бы устранена. Для изучения этой задачи с помощью внешней организации по исследованию операций была создана группа исследования операций для данной компании.

Эта группа начала с анализа организации. Операционисты проследили процесс передачи информации и материалов, начиная от получения заказов через отдел сбыта вплоть до поставок готовой продукции потребителю. В результате обработки большого количества собранной информации была построена графическая модель процессов связи и управления (см. рис. 4.13).

Затем определили среднее время, потраченное на каждой стадии процесса. Отсюда рассчитали общее время всего процесса. Это время оказалось на 50% больше даже обычных сроков поставки. Чтобы уложиться даже в эти сроки, требовались значительные затраты денежных и материальных ресурсов. На одни только операции планирования и обработки документов уходила значительная часть всего времени. Следовательно, эти операции необходимо было изменить, чтобы не срывать сроков поставки. Однако в результате первоначального анализа организации не было получено достаточного количества информации для разработки новой схемы процессов планирования и передачи информации. Поэтому был проведен более детальный анализ всего процесса в обратном порядке. В результате были построены более подробные графические модели для каждой стадии процесса. При этом были обнаружены причины ряда задержек, наиболее существенными из которых являются следующие.

1. Неполная информация о заказах.
2. Территориальное разделение отделов, в которых принимаются зависящие друг от друга решения.
3. Децентрализация взаимосвязанных процессов планирования.
4. Неполные и устаревшие документы на складе.
5. Дублирование работы.

После этого в группе исследования операций была построена модель организации, в которой весь процесс заканчивается за заданное время

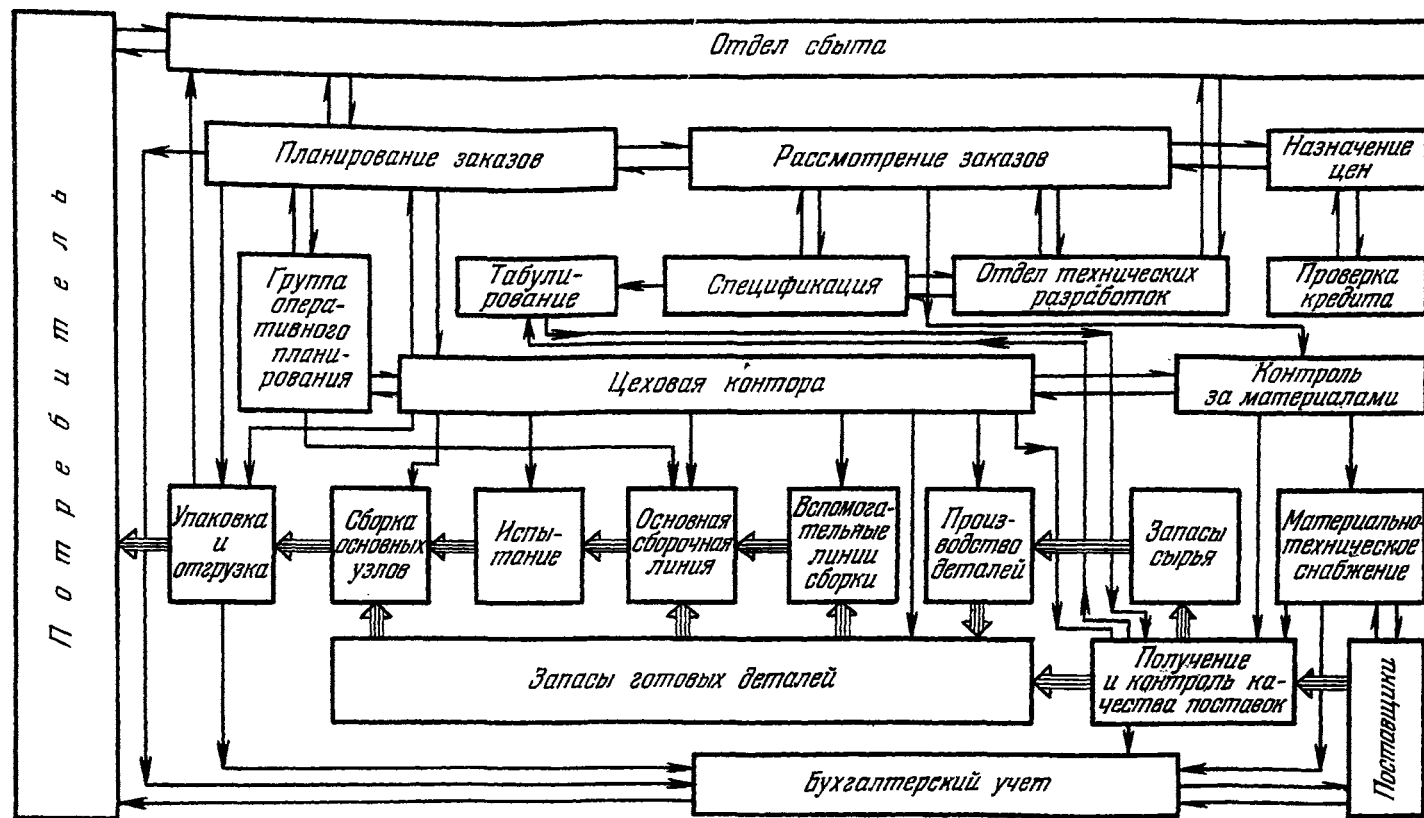


Рис. 4.13. Схема информационных и материальных потоков в организации.

(см. рис. 4.14). Это была идеализированная модель, потому что в ней не учитывались различные практические задачи, возникающие при изменениях в организации. Но эта модель могла служить в качестве эталона для разработки практической модели. Однако для того, чтобы построить практическую модель, потребовалось проделать значительную работу.

Существенной особенностью новой модели была разработка эффективной системы оборота складских документов. В старой системе документы отставали от соответствующих операций на период до 40 дней. Разработанная ранее система с ежедневной отправкой документов по почте при использовании перфорационных машин была заменена новой, в которой задержка документов составляла не более одного дня. Было также предложено объединить ряд работ технического отдела и отдела снабжения. Из отдела сбыта была затребована более полная информация в виде подробных форм по заказам. Группа спецификации была переведена в другое помещение, чтобы устранить или по крайней мере уменьшить дублирование работы. Например, работа по составлению спецификации материалов была передана из чертежного отдела в группу спецификации. Раньше же эта работа отнимала у чертежника четверть его рабочего времени. А в помощь работникам группы спецификации направлялись инженеры-производственники для составления перечня по оборудованию, тем самым было устранено дублирование еще одной работы. Было решено также упразднить группу оперативного планирования, причем составление ежедневных планов было передано в отдел производственного планирования. Операции по сборке начали планировать только при наличии всех необходимых деталей, а данные о деталях должны были поступать с новых складских перфокарт.

Были пересмотрены методы составления и использования накладных на материалы и суммарных перечней материалов, в результате чего значительно сократилось количество документов, поступающих на линию сборки. В окончательном варианте было намечено направлять в технологический отдел и в отдел контроля качества машинный документ на материалы, в котором указываются все основные элементы, узлы и другие сведения. Такой документ также может служить для контролеров в качестве полного контрольного перечня, что дает возможность ликвидировать существовавший до реорганизации контрольный документ.

Когда готовилась эта книга, реорганизация только начиналась. Но в то же время разрабатывались новые методы принятия решений в процессе планирования. Велись работы по определению величины партии деталей в производстве. Уже были выяснены вопросы, связанные с промежуточным (буферным) складированием деталей в процессе обработки. При этом значительно уменьшалось число срочных заказов на производство деталей, а также время переналадки оборудования. В результате уплотнения рабочего графика оборудования появилось дополнительное рабочее время, которое можно выгодно использовать, поскольку значительно возрос спрос на эту продукцию.

Полная реорганизация и применение новых правил принятия решения в производственном планировании должны дать значительную ежегодную экономию. Так, по самым скромным подсчетам, только мероприятия по изменению величины партии деталей дадут экономию в пределах \$250 000. Для компании также очень важно, что обслуживание потребителя улучшилось и будет дальше улучшаться. Это связано не только с использованием новой системы принятия решений, но вызвано тем, что после реорганизации длительность процесса обработки сократится и будет продолжаться сокращаться, так что эффективность новой системы будет возрастать.

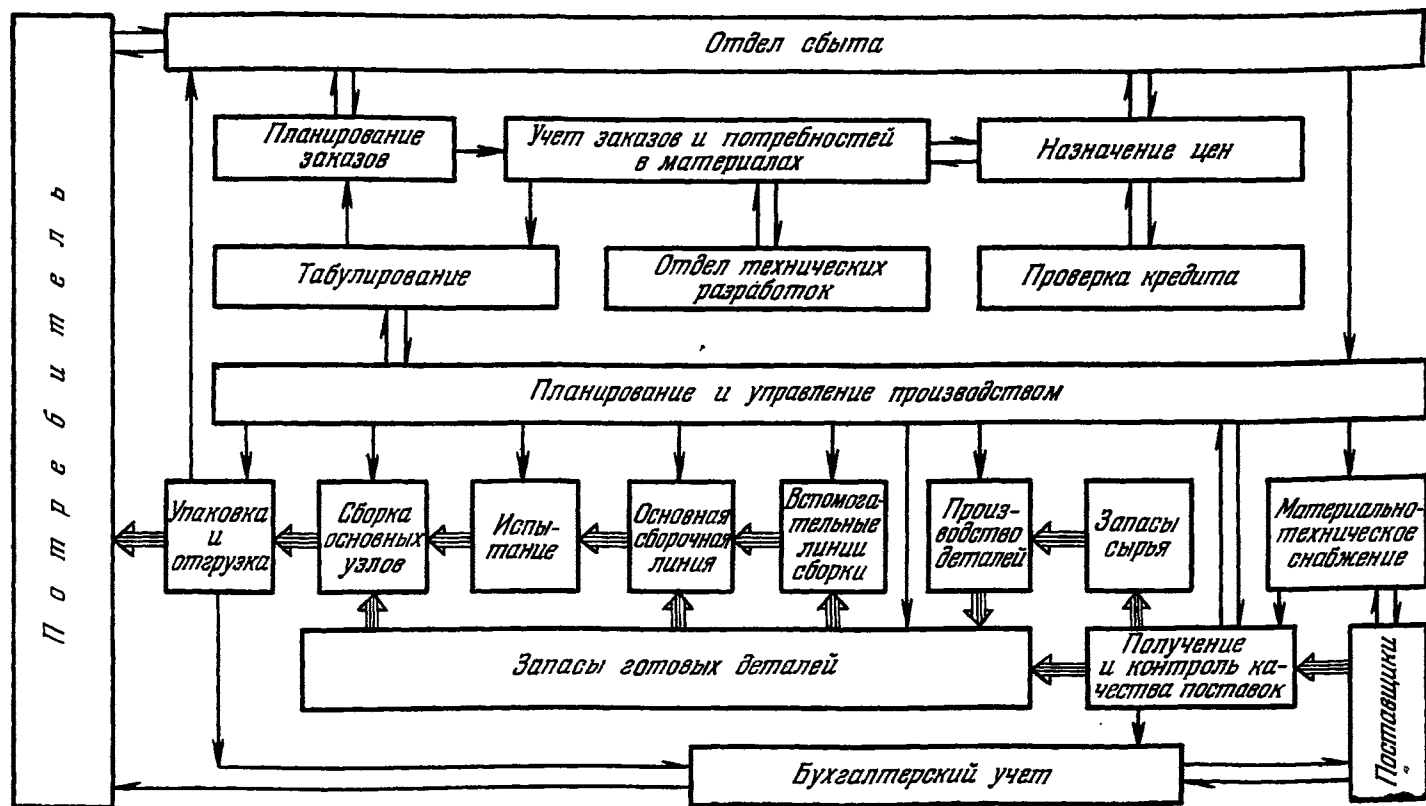


Рис. 4.14. Схема информационных и материальных потоков в организации.

ЗАКЛЮЧЕНИЕ

В данной главе рассматривались вопросы, связанные со структурой информационных моделей, их построением и использованием, а также приведены некоторые задачи, возникающие при работе с такими моделями. Отмеченные методы разработки моделей и способы их использования не являются исчерпывающими и только предлагаются для начала исследования задачи. Более подробные сведения можно найти в библиографии к данной главе. Соответствующие ссылки даются в тексте и в примечаниях.

Содержание этой главы можно подытожить в следующих выводах по построению и использованию информационных моделей.

1. Начертите схему информационной сети. Используйте один из предложенных методов или придумайте новый. Чем яснее исследователь представляет себе, чего он хочет, тем более четкой будет схема, которую он получит. Задайте вопросы для проверки.

2. Проверьте схему. Правильна ли эта схема? Нашли ли отображение наиболее важные виды связей? Достаточно ли подробна схема? Не слишком ли она подробна?

3. Исследуйте схему. Попытайтесь обнаружить путем исследования схемы очевидные ошибки (разрыв связи, лишние связи, резкие изменения потоков информации и другие). По аналогии или непосредственно из полученных данных попытайтесь обнаружить наряду с очевидными ошибками также и менее явные ошибки.

4. Если обнаружены специальные задачи (такие, как управление запасами, задачи очередей, поиска и другие), то используйте приемы, разработанные в соответствующих главах этой книги. Проследите влияние каждой задачи на другие элементы организации, чтобы убедиться, что задача действительно является стандартной.

5. Если обнаружены нестандартные задачи (для которых нет стандартных специализированных моделей), используйте схему информационной модели, чтобы выявить неизвестные величины, которые следует изучить. Для этих неизвестных постройте подходящую модель.

6. В частности, если обнаруженные задачи связаны в основном с дефектами системы связи или управления, то используйте таблицу 4.1, чтобы определить соответствующие неизвестные для дальнейшего изучения. Поищите аналогии в других дисциплинах. Сравните схему информационной сети со стандартными схемами управления или схемами других организаций. Еще раз задайте более конкретные вопросы. Снова исследуйте схему, чтобы попытаться найти решение обнаруженных задач.

7. Сравните предложенные решения (полученные на предыдущих стадиях) с первоначальной схемой информационной сети, чтобы (а) проследить влияние решения частной задачи на остальные элементы организации, (б) рассчитать предполагаемую экономию или (в) предложить другие решения.

8. На каждом из этих этапов смелее пользуйтесь интуицией, привлекайте аналогию, обдумывайте возможные решения. Записывайте эти размышления и пробы. Проверяйте предложенные решения или с помощью эксперимента, или путем разработки новых специализированных моделей. Когда получено удовлетворительное решение, запишите процедуру его получения. Это позволит позднее проверить решение или использовать информацию об этих решениях в дальнейших исследованиях.

Л И Т Е Р А Т У Р А

1. A d r i a n E. D., The Basic of Sensation; the Action of the Sense Organs, W. W. Norton & Co., New York, 1928.
2. A d r i a n E. D., The Physical Background of Perception, Clarendon Press, Oxford, England, 1947.
3. A l d e r s o n and S e s s i o n s C o., Communication Analysis and Organization Planning, Cost and Profit Outlook, VII, no. 4 (Apr. 1954). (Alderson and Sessions Company, 1401 Walnut St., Philadelphia, Pa.).
4. A l l e e W. C., Animal Aggregations, University of Chicago Press, Chicago, 1931.
5. A l l e e W. C., The Social Life of Animals, W. W. Norton & Co., New York, 1951.
6. B a c k K. N., The Exertion of Influence Through Social Communication, in Festinger et al., Theory and Experiment in Social Research, Research Center for Group Dynamics, University of Michigan, Ann. Arbor, 1950.
7. B a l e s R. F., Interaction Process Analysis, Addison-Wesley Press, Cambridge, Mass., 1950.
8. B a l e s R. F., The Interaction Recorder, Hum. Relat., 1, 456—463 (1948).
9. B a r r e t t E., Cybernetics as Applied to a Study of Normal and Abnormal Adjustment Mechanisms, J. Psychol., 30, 11—31 (1950).
10. B a r r e t t E. and P o s t G., Introduction to Some Principles of Applied Cybernetics, J. Psychol., 30, 3—10 (1950).
11. B a v e l a s A., Communication Patterns in Task Oriented Groups, J. Acoust. Soc. Amer., 22, 725—730 (1950).
12. B a y l i s s W. M., Principles of General Physiology, Longmans, Green and Co., New York, 4th ed., 1927.
13. Bibliography on Cybernetics, Electrical Engineering Department, Massachusetts Institute of Technology, Cambridge, 1950.
14. B r o n f e n b r e n n e r V., The Graphic Representation of Sociometric Data, Sociometry, 7, 283—289 (1944).
15. C a n n o n W. B., Wisdom of the Body, W. W. Norton & Co., New York, 1939.
16. C h a p p l e E. D., Anthropological Engineering: Its Use to Administrators, Appl. Anthropol. (Jan.—Mar. 1943); reprinted in S. D. Hoslett (ed.), Human Factors in Management, Park College Press, Parkville, Mo., 1946.
17. C h a p p l e E. D., The Interaction Chronograph: Its Evaluation and Present Application, Personnel, 25, 295—307 (1949).
18. C h a p p l e E. D., How a Consultant Deals with Organizational Problems (unpublished), Industrial Engineering Seminar 312, Columbia University, New York, Mar. 9, 1953.
19. C h a p p l e E. D., Measuring Human Relations: An Introduction to the Study of the Interaction of Individuals, Genet. Psychol., 22, no. 1 (Feb. 1940).
20. C h r i s t i e L. S., L u c e R. D. and M a c y J., Jr., Communication and Learning in Task-oriented Groups, Technical Report No. 231, Research Laboratory for Electronics, Massachusetts Institute of Technology, Cambridge, May 13, 1952.
21. C h u r c h m a n C. W., Introduction, in Proceedings of the Conference on Operations Research in Production and Inventory Control, Case Institute of Technology, Cleveland, p. 8, Jan. 20—22, 1954.
22. C h u r c h m a n C. W. and A c k o f f R. L., Purposive Behavior and Cybernetics, Social Forces, 29, 32—39 (1950).
23. D e m p s e y E. W., Homeostasis, in S. S. Stevens (ed.), Handbook of Experimental Psychology, John Wiley & Sons, New York, pp. 209—235, 1951.
24. D e u t s c h K. W., Communication in Self-governing Organizations (unpublished), Industrial Engineering File No. 1589, Columbia University, New York, Spring, 1951.
25. D e u t s c h K. W., Communication Theory and Social Science, Amer. J. Orthopsychiat., 22, 469—483 (1952).
26. D e u t s c h K. W., Higher Education and the Unity of Knowledge, in Lyman Bryson et al. (eds.), Goals for American Education, Conference on Science, Philosophy and Religion in their Relation to the Democratic Way of Life, Inc., New York, pp. 55—139, 1950.
27. D e u t s c h K. W., Innovation, Entrepreneurship, and the Learning Process, in H. H. Cole and R. R. Wohl (eds.), Change and the Entrepreneur, Harvard University Press, Cambridge, 1949.
28. D e u t s c h K. W., Mechanism, Organism, and Society, Phil. Sci., 230—252 (July 1951).
29. D e u t s c h K. W., Mechanism, Teleology, and Mind, Phil. Phenom. Res., 12, 185—222 (1951).
30. D e u t s c h K. W., Models of Communication and Education (unpublished), Industrial Engineering Seminar 312, Columbia University, New York, Mar. 13, 1951.

31. Deutsch K. W., *Nationalism and Social Communication*, Technology Press, Cambridge, and John Wiley & Sons, New York, 1953.
32. Deutsch K. W., *On Communication Models in the Social Sciences*, Publ. Opinion Quart., 16, 356—380 (1952).
33. Deutsch K. W., *Political Community at the International Level*, Doubleday Short Studies in Political Science, Doubleday & Co., New York, 1954.
34. Frank L. K., *Teleological Mechanisms*, Ann. N. Y. Acad. Sci., 50, 182—278 (1948).
35. Hatfield H. S., *The Inventor and His World*, Penguin Books (Pelican Edition), New York, 1948.
36. Hebb D. O., *The Organization of Behavior — a Neuropsychological Theory*, John Wiley & Sons, New York, 1949.
37. Hertz D. B., *Theory and Practice of Industrial Research*, McGraw-Hill Book Co., New York, 1950.
38. Homans G. C., *A Conceptual Scheme for the Study of Social Communication*, Amer. Sociological Rev. (Feb. 1947).
39. Horsfall A. B. and Arensberg C. M., *Teamwork and Productivity in a Shoe Factory*, Hum. Organiz. (Winter 1949).
40. Jahoda M., Deutsch M. and Cook S. W., *Research Methods in Social Relations*, Part II, The Dryden Press, New York, 1952.
41. Jenkins D., *Feedback and Group Self-Evaluation*, J. Social Issues, 4, 2, 50—60 (1948).
42. Lashley K. S., *The Problem of Serial Order in Behavior*, in Lloyd A. Jeffress (ed.), *Cerebral Mechanisms in Behavior*, John Wiley & Sons, New York, pp. 112—130, 1951.
43. Lasswell H. D., *The Structure and Function of Communication in Society*, in Lyman Bryson et al. (eds.), *The Communication of Ideas*, Harper and Brothers, New York, pp. 37—51, 1948.
44. Lasswell H. D., Smith B. L. and Casey R. D., *Propaganda, Communication, and Public Opinion*, Princeton University Press, Princeton, 1946.
45. Lazarsfeld P. F., *Mathematical Thinking in the Social Sciences*, Free Press, Glencoe, Ill., 1954.
46. Marcosson Isaac F., *Adventures in Interviewing*, John Lane, The Bodley Head, Ltd., London, 1920.
47. Mason S. J., *On the Logic of Feedback*, Technical Report No. 153, Research Laboratory for Electronics, Massachusetts Institute of Technology, Cambridge, 1953.
48. Mayo E., *The Social Problems of an Industrial Civilization*, Harvard Graduate School of Business Administration, Boston, 1945.
49. Meyer H., *On the Heuristic Value of Scientific Models*, Phil. Sci., 18, no. 4 (Oct. 1951).
50. Miller G. A., *Language and Communication* McGraw-Hill Book Co., 1951.
51. Morgenstern O., *Prolegomena to a Theory of Organization*, Project RAND, RM-734, Santa Monica, 1951.
52. Nejedlski L., *Communication in Practical Affairs*, in L. Bryson et al. (ed.), *The Communication of Ideas*, Harper and Brothers, New York, 1947.
53. Norton-Taylor D., *Why Don't Businessmen Read Books*, Fortune, 116—117 (May 1954).
54. Penfield W. and Rasmussen T., *The Cerebral Cortex of Man*, The Macmillan Co., New York, 1950.
55. Pitts W. and McCulloch W. S., *How We Know Universals, the Preception of Auditory and Visual Forms*, Bull. Math. Biophys., 9, 124—147 (1947).
56. Rapoport A. and Shimmel A., *Mathematical Biophysics, Cybernetics, and General Semantics, Etc., A Review of General Semantics*, 6, 145—159 (1949).
57. Rosenblueth A., Wiener N. and Bigelow W., *Behavior, Purpose, and Teleology*, Phil. Sci., 10, 18—24 (1943).
58. Rosenblueth A. and Wiener N., *Purposeful and Non-Purposeful Behavior*, Phil. Sci., 17, no. 4, 318—326 (Oct. 1950).
59. Rosenblueth A. and Wiener N., *The Role of Models in Science*, Phil. Sci., 12, 316—322 (1945).
60. Ruesch J. and Bateson G., *Communication: The Social Matrix of Psychiatry*, W. W. Norton & Co., New York, 1951.
61. Ruesch J. and Bateson G., *Structure and Process in Social Relation*, Psychiatry, 12, 105—124 (1949).
62. Schneilra T. C., *The Levels Concept in the Study of Social Organization in Animals*, in Rohrer and Muzafer (eds.), *Social Psychiatry at the Crossroads*, Harper and Brothers, New York, 1951.
63. Shannon C. E. and Weaver W., *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949.

64. Simon H. A., Administrative Behavior, The Macmillan Co., 1947.
65. Simon H. A., Modern Organization Theories, Advanced Mgmt., 15, 2—4 (Oct. 1950).
66. Simon H. A., On the Application of Servomechanism Theory in the Study of Production Control, Project RAND, P234, Santa Monica, Aug. 15, 1951.
67. Simon H. A., A Study of Decision-Making Processes in Administrative Organization, The Macmillan Co., New York, 1947.
68. Stumpers F. L., A Bibliography of Information Theory, Communication Theory, and Cybernetics, Trans. I.R.E., PGIT-2 (Nov. 1953).
69. Taylor R., Purposeful and Non-Purposeful Behavior, Phil. Sci., 17, no. 4 (1950).
70. Trimmer J. D., Instrumentation and Cybernetics, Sci. Monthly, 69, 328—331 (1949).
71. Waddington C. H., The Scientific Attitude, Penguin Books, London, pp. 122—124, 1941.
72. Weber Max, The Theory of Social and Economic Organization, W. Hodge, London, 1947.
73. From Max Weber: Essays in Sociology, translated, edited, and with an introduction by H. H. Gerth and C. Wright Mills, Oxford University Press, New York, 1946.
74. Wiener N., Cybernetics. Control and Communication in the Animal and the Machine, John Wiley & Sons, New York, 1948.
75. Wiener N., The Human Use of Human Beings: Cybernetics and Society, Houghton Mifflin Co., Boston, 1950.
76. Wiener N., Speech, Language and Learning, J. Acoust. Soc. Amer., 22, 696—697 (1950).
77. Wiener N., Time, Communication, and the Nervous System, in Teleological Mechanisms, Ann. N. Y. Acad. Sci., 50, 197—219 (1948).
78. Whyte W. H., Is Anybody Listening?, Simon and Schuster, New York, 1952.

Глава 5

ПОСТАНОВКА ЗАДАЧИ

ВВЕДЕНИЕ

Исследование должно начинаться с постановки задачи, но обычно переходят к следующей стадии исследовательского процесса, не получив точной постановки задачи. Процесс постановки задачи обычно бывает непрерывным. После окончания этапа первоначальной постановки задачи переходят к исследовательскому процессу, однако по ходу процесса постановка задачи непрерывно меняется и уточняется. Этот процесс продолжается вплоть до получения решения. Другими словами, до самого конца исследований неизвестно, правильно ли поставлена задача, а иногда этот вопрос так и остается открытым.

Стремление скорее начать исследования часто приводит к сокращению времени и сил, выделенных для постановки задачи, что впоследствии может привести к большим потерям. Группа исследования операций должна применять некоторые стандартные процедуры постановки задачи, на выполнение которых должно быть выделено определенное время.

В этой главе рассматривается идеализированный процесс постановки задачи. Этот процесс является наилучшим из существующих в настоящее время при условии, что нет никаких ограничений. На практике редко встречаются такие идеализированные условия. Однако, чтобы разработать практически применимый процесс, необходимо иметь представление о наилучшем возможном процессе. Кроме того, представление о таком идеализированном процессе дает нам некоторую цель, к которой мы стремимся и тем самым совершенствуем свои исследования.

Идеализированный процесс является довольно длительным, требует значительного времени и изобретательности. Поэтому может возникнуть вопрос, зачем нужен этот процесс, особенно если учесть, что во многих исследованиях этот этап специально не выделяется. Дело в том, что исследование операций предполагает исследование экономики операций. Тем более необходимо проводить исследование экономики своих собственных операций. Кроме того, обычно разница между идеализированным и реальным процессами вызвана неправильным представлением об экономической стороне дела. Под «практичным» обычно понимают процесс, который проводится небольшими силами с минимальными затратами времени и средств. Представление об «идеальном» связано с большими затратами времени и средств. Поэтому различие между идеальным и практичным (реальным) не сводится к различию между «белым и черным», а определяется чисто экономическими соображениями. Практичный подход обычно дешевле с точки зрения более близких тактических целей, но обходится дороже с позиций более постоянных стратегических целей. Идеальный подход, наоборот, обеспечивает минимизацию затрат времени и средств по отношению к общим стратегическим целям, но сопряжен с большими расходами с точки зрения решения тактических задач.

Правильное экономическое решение в такой ситуации означает выбор удачного компромисса между затратами, связанными с достижениями тактических и стратегических целей. Следовательно, рациональная экономическая политика не сводится только к практичному или только к идеальному решению, но предполагает точное знание идеального решения, чтобы в каждой конкретной ситуации знать, насколько близко следует придерживаться этого решения. Только зная все возможности исследования, можно определить наилучший в экономическом смысле вариант решения в каждом конкретном случае.

В данной главе особое внимание уделяется идеальной процедуре постановки задачи, поскольку эту процедуру часто недооценивают. Практические соображения будут рассмотрены только в связи с наиболее сложными шагами этой процедуры.

ПЕРИОД ОРИЕНТАЦИИ

Первый период исследования, связанного с постановкой задачи, называется периодом ориентации. Этот период может продолжаться от одного до нескольких месяцев. Он может оканчиваться письменным или устным отчетом перед заказчиками. В этом отчете иногда даются временные и стоимостные оценки, хотя обычно они не являются очень точными.

Помимо выполнения возможности постановки задачи во время периода ориентации выполняются еще две важные функции. В примере, приведенном в предыдущей главе, показывается, как этот этап исследования помогает «внешней» группе исследования операций оценить задачу и организацию, прежде чем заключать договоры на исследование конкретной задачи. Аналогичная возможность предоставляется и заказчикам. В конце периода ориентации определяются условия, в которых будут проводиться исследования, а также принимаются административные меры по выполнению этих условий. Обычно в период ориентации достигается взаимопонимание между группой исследования операций и заказчиками. Такой период нужен и в тех случаях, когда исследовательская группа составлена только из сотрудников данной компании, поскольку в каждой новой задаче исследования операций встречаются новые операционные и административные

задачи. Этап ознакомления необходим и тогда, когда исследование носит «научный», а не «прикладной» характер и «заказчиками» являются ученые или научные организации (учреждения, университеты, правительство и т. д.).

ЭЛЕМЕНТЫ ЗАДАЧИ

Прежде чем мы сможем сформулировать задачу, мы должны представить себе, что же такое задача, иными словами, каковы элементы задачи.

Во-первых, перед определенным лицом или перед группой людей должна стоять задача. Этот человек или группа не удовлетворены положением дел в какой-то области и хотят найти решение, чтобы изменить это положение. По этой причине назовем этого человека (или группу людей) лицом, принимающим решение. В тех случаях, когда речь идет об управлении действиями организованной системы людей, машин или смешанной системы, лицо, принимающее решение, можно назвать лицом, определяющим политику, или руководителем. Таким образом, первым элементом задачи является лицо (группа), принимающее решение.

Во-вторых, для существования задачи необходимо, чтобы человек, принимающий решение, желал изменить положение, т. е. должны быть некоторые цели, которых надо достичь. Цели — второй элемент задачи.

В-третьих, задача связана с какими-то условиями или средой, в которой учитывается наличие или отсутствие различных ресурсов. Такой средой в задачах исследования операций обычно является организованная система, включающая людей и машины. Система или среда является третьим элементом задачи.

И наконец, необходима возможность выбора между различными (по крайней мере двумя) способами действия, или политиками. Недовольство может существовать и без возможности выбора, но задача не существует без этого. Задача всегда включает вопрос: что делать? Но этот вопрос становится задачей только в том случае, если существуют различные пути достижения цели. Различные способы действия (стратегии) являются четвертым элементом задачи.

Теперь рассмотрим, что нужно знать о каждом из этих четырех элементов (лицо, принимающее решение, цели, система, различные способы действия), чтобы поставить исследовательскую задачу.

Лицо (группа), принимающее решение. Сначала надо определить, кто принимает решение. В задачах исследования операций надо определить, кто обладает полномочиями для принятия, отмены или изменения определенных стратегий, связанных с управлением организацией и системой. В некоторых системах такими полномочиями обладает группа людей. В любом случае следует понять, как лица, обладающие властью, принимают решения, что особенно важно для выработки рекомендаций для руководства в ходе проекта.

Необходимо определить структуру группы лиц, принимающих решения. Принимают ли они решения сообща или в определенной последовательности? Большинство голосов? Если нет, то кто обладает правом налагать вето и кто имеет решающий голос? Является ли процесс принятия решения формальным или нет?

Приведем ряд вопросов, которые могут помочь в изучении процесса принятия решения в задачах:

1. Кто дает рекомендации по изменению политики организации?
2. Чье утверждение необходимо и как оно выражается?

3. Что представляет собой окончательное утверждение (большинство голосов совета, утверждение наиболее авторитетного лица после ряда других резолюций и т. д.)?

4. Обладает ли кто-нибудь абсолютным правом налагать вето? Если нет, то как можно отклонить какое-то предложение?

5. Кто несет ответственность за проведение решений в жизнь, если они утверждены?

6. Кто должен оценивать предпринятые действия?

Схемы организации не дают ответа на эти вопросы, но с их помощью можно определить, у кого можно получить интересующие сведения.

Цели, стоящие перед лицом, принимающим решение. Непосредственные опросы лиц, принимающих решения, редко дают возможность выяснить все цели, связанные с задачей. С этого можно начать, но редко удастся получить достаточную информацию для полного выявления всех целей. Укажем один особенно эффективный способ определения истинных целей. Исследователь пытается составить список всех возможных результатов (целей) проекта. На этом этапе решения задачи список не обязательно должен быть точным, полным или даже реальным. Затем следует предложить заказчику возможное решение и спросить у него, будет ли он удовлетворен, если все перечисленные результаты будут достигнуты. При этом часто оказывается, что предложенные исследователями рекомендации заказчика не устраивают. Выяснение причин отказа поможет выявить новые цели.

Приведем пример. Заказчик хотел определить, где целесообразно разместить новое предприятие. Непосредственный опрос позволил выяснить очевидные цели заказчика (например, минимизацию транспортных издержек, затраты на рабочую силу и т. д.). Но недостаточное изучение целей, хотя бы с помощью вышеописанного приема, может привести к плохим результатам. Проведенные исследования позволили определить наиболее выгодный пункт для строительства предприятия. Но заказчик отказался от этого пункта, потому что он не хотел иметь дела с лидером местного профсоюза. Экономические соображения оказались на втором плане по сравнению с личной антипатией. В результате большая работа была проделана впустую.

При изучении целей мы должны рассматривать не только цели, которые надо достигнуть, но также уже достигнутые цели, которые надо сохранить. Так, заказчик стремится поддержать по крайней мере определенный уровень занятости или не хочет занимать дополнительных сумм или расширять производство. Другими словами, могут быть цели, которые надо достигнуть, и цели, которые надо сохранить.

Например, цели, которые надо достигнуть, можно сформулировать следующим образом: (а) уменьшить производственные расходы, (б) улучшить обслуживание потребителя, (в) увеличить сбыт. Цели, которые надо сохранить: (а) поддержать стабильную занятость, (б) сохранить ведущую роль в своей области производства, (в) сохранить хорошую репутацию.

В задаче, приведенной в главе 2 (управление производством и запасами), рассматривается только одна цель, которую надо достигнуть: уменьшить общую стоимость производства и хранения. В то же время надо сохранить ряд условий (целей): (а) не увеличивать площадь складов, (б) сохранить кредитоспособность компании, (в) сохранить существующую скорость сборки станков.

Система. Большинство организованных систем включает в себя следующие элементы: руководители; исполнители, которые выполняют их указания; инструменты и материалы, используемые при этом; внешние организации, на которых отражается деятельность данной организации; социальные условия, в которых действует система. В частности, в коммерческих и промышленных системах существуют следующие элементы: (1) *руководство*, которое управляет (2) *рабочими*, управляющими (3) *машинами*, которые превращают (4) *материалы* в готовую продукцию или услуги для (5) *потребителей*, которых стараются перехватить (6) *конкуренты*; (7) *правительство* и *общество*. Можно привести аналогичный пример и для военных систем.

Однако руководство, рабочие, машины и материалы образуют систему только благодаря *организации*. В организации существует разделение труда между группами, каждая из которых участвует в операциях, направленных на достижение общей цели. Чтобы понять организацию и систему, надо проследить последовательность операций, которые выполняют отдельные подразделения организации, и выяснить, каким образом ими руководят для достижения общих целей организации. В предыдущей главе была описана подобная процедура построения структуры организации.

Опыт показывает, что почти в каждом случае такой анализ много дает руководству и персоналу самой организации, а не только внешней группе исследования операций. Во многих случаях такой анализ дает дополнительные результаты, которые оправдывают затраты времени на исследование. Например, такой анализ помогает обучать новых сотрудников, а также создавать программы подготовки руководителей. Схемы организации используются руководством аналогично тому, как в военном штабе используют военные карты. И действительно, схемы информационных и материальных потоков можно рассматривать как карты изучаемой системы. При анализе организации можно обнаружить дублирование отдельных работ и недостатки управления. В некоторых случаях в результате такого анализа можно обнаружить другие пути исследования, независимые от работы группы исследования операций.

Не все вышеупомянутые элементы активно участвуют в работе организации. Например, конкуренты, правительство и общество обычно налагают определенные ограничения на систему, в то время как руководители принимают непосредственное участие в процессе принятия решения. Во многих задачах не требуется подробных сведений о том, каковы эти ограничения. Например, в главе 2 дается анализ производственного процесса, поэтому там роль конкурентов, правительства и общества не рассматривается. Если же встает вопрос о политике цен, то при анализе системы надо учитывать эти элементы системы.

Цели и «решающих» участников системы. Поскольку в систему входят различные «решающие» участники, следует определить их интересы (цели). Так как в предложенных вариантах решения задач могут затрагиваться интересы тех, кто проводит решения в жизнь, а также потребителей и конкурентов, то для полного понимания задачи важно выяснить эти интересы. Кроме того, успех практического использования любого предложенного решения зависит от того, как его примут различные заинтересованные стороны. В тех же случаях, когда изучение их интересов показывает, что их реакция, по-видимому, будет отрицательной, можно рекомендовать проведение предварительной воспитательной разъяснительной кампании. Если же даже разъяснительная кампания не поможет, то решение, которое неприемлемо для некоторых участников, следует отверг-

нуть. Таким образом, интересы сторон накладывают ограничения на возможные решения. Однако в некоторых случаях понимание этих интересов может расширить область возможных решений.

Для выяснения целей неосновных участников можно использовать вышеописанные приемы, связанные с целями лиц, принимающих решение. Но обычно в этом случае гораздо труднее практически реализовать указанные приемы. Поэтому часто приходится использовать другие методы или полагаться на общие соображения.

В главе 2 приведен пример изучения вопроса о размещении заказов на запчасти. Был создан комитет, в который входили лица, заинтересованные в данном вопросе. Их цели удалось выяснить, исходя из возражений против предлагаемых изменений. Естественно, каждый из них пытался сохранить свое положение, а возражения в основном были продиктованы страхом, что они могут оказаться под угрозой увольнения. Было также установлено, что одной из главных их целей было улучшение работы компании при условии, что им будет оказано доверие и поддержка. После того, как удалось это выяснить, можно было с уверенностью принимать их предложения, дав им гарантию спокойной работы и сохранения прежнего положения.

Различные способы действия (альтернативные стратегии). Различные способы действия (возможные варианты решения) обычно выясняются уже на ранних стадиях в процессе постановки задачи. Однако вначале список альтернатив не является полным. Исследователи должны стремиться получить возможно более полный список, желательно включать даже варианты, которые не имеют практического значения.

Чтобы быть уверенным в полноте списка, лучше всего проанализировать саму систему. Рассмотрим, например, компанию, которая обеспечивает защиту от грабителей. По соглашению между потребителем и компанией устанавливается система защиты, которая связана с центральным пультом компании. Компания содержит группы вооруженной охраны, которые отправляются на место по получении сигнала тревоги от абонента. При такой системе защиты компания получала много фальшивых вызовов. Эти вызовы обходятся очень дорого и потому представляют значительную часть расходов.

Компания пригласила операционную группу для изучения возможности сокращения этих расходов. Главной целью компании было увеличить чистую прибыль, хотя конкретная цель была первоначально поставлена как уменьшение расходов по вызовам. Какие еще возможности существуют у компании? Может ли она обеспечивать защиту иначе? Анализ системы позволил выявить некоторые возможности, которые вначале не были очевидными.

На входе системы производится продажа потребителю услуг (т. е. обеспечения защиты). Здесь существуют две возможности:

1. Выбрать только тех потребителей, у которых система защиты дает меньшее количество ложных вызовов.

2. Назначать потребителям плату за обслуживание пропорционально количеству ложных вызовов, которые можно ожидать от выбранной ими системы защиты.

После заключения контракта проектируется и устанавливается система защиты. Здесь существует много вариантов, связанных с типом элементов системы и способом их соединения. Эти варианты можно разделить на две группы, одна из которых предполагает использование стандартных элементов и другая — новых элементов, которые спроектированы специально с целью уменьшить ложные вызовы.

Когда система установлена, то ложные вызовы могут происходить не только из-за неисправной работы оборудования, но также из-за злоупотреблений или неправильного использования оборудования абонентом. Выход может состоять в лучшем разъяснении абонентам правил использования оборудования и в наборе абонентов, которые лучше используют оборудование.

Поскольку даже хорошая установка со временем, по мере ее использования, может стать непригодной, компания должна следить за ее состоянием, ремонтировать ее, а иногда заменять. При этом следует рассмотреть различные варианты, связанные с периодичностью и характером осмотра, критериями для замены и ремонта оборудования и т. д.

И наконец, существуют различные способы обслуживания вызовов, полученных на центральном пульте компании. Здесь можно учесть численность и местонахождение отрядов охраны, систему связи с ними, возможные действия, предпринятые ими, и типы рапортов, которые они посылают в центральную контору компании.

Чтобы получить перечень различных способов действия такой, как в приведенном примере, можно воспользоваться следующими вопросами. Для каждой части системы следует определить, будут ли (1) изменения состава персонала влиять на эффективность системы, понимая под этим интересы заказчика, (2) изменения в операциях влиять на эффективность системы, (3) изменения материалов и машин влиять на эффективность системы, (4) изменения окружающей среды влиять на эффективность системы.

В тех случаях, когда ответ будет утвердительным, можно исследовать различные возможности улучшения работы системы.

В некоторых случаях альтернативы представлены в следующем виде: делать A или не делать A ; например, производить новый продукт или нет, продавать новому покупателю или нет. Иногда же есть много разных вариантов; например, производство n продуктов в партии, где n может принимать различные значения. На этом этапе решения задачи важно определить число различных вариантов или хотя бы их возможный диапазон.

Разработка новых стратегий. В некоторых случаях ни один из существующих возможных вариантов решения задачи нас не устраивает. Например, может оказаться, что владельца автомобильного завода не устраивают все существующие типы рессор с точки зрения поставленных целей. Или при производстве бумаги все существующие способы контроля качества представляются слишком дорогими. В таких случаях нужны новые стратегии; но поскольку их нет, надо их разработать. Такие задачи можно назвать *задачами разработки*. А задачи, которые предполагают выбор одного из существующих вариантов, назовем *задачами оценки*. Задачи разработки и оценки являются крайними возможными типами задач, каждая же из конкретных задач в зависимости от степени разработанности методов ее решения расположена между этими крайними типами задач. Хотя здесь рассматриваются в основном задачи оценки, следует привести пример задачи разработки.

Предположим, что разрабатывается более эффективное оружие против определенной цели. Группа исследования операций изучает существующие виды оружия и путем анализа пытается выявить те характеристики, от которых зависит эффективность оружия. С помощью такого анализа вырабатывается представление о новом оружии, улучшенные характеристики которого можно оценить путем экстраполяции. Затем проектировщики изучают требования, выдвинутые группой исследования операций, и обнаруживают, например, что их можно удовлетворить, только если использовать определен-

ный чувствительный механизм внутри оружия, а это потребует уменьшения количества взрывчатых материалов. После проведения такого исследования операционная группа может приступить к изучению способов доставки взрывчатого вещества ближе к цели, чтобы компенсировать уменьшение его разрушительной силы. На каждом этапе группа исследования операций оценивает новый предложенный вариант и сравнивает его с существующими.

Из приведенных примеров видно, что участие группы исследования операций в разработке новых задач имеет большое значение для промышленности, особенно при разработке новых продуктов и процессов.

Контрстратегии. До сих пор, обсуждая различные стратегии, мы рассматривали только возможные стратегии стороны, принимающей решения. Однако следует также определить класс стратегий, принимаемых другими участниками системы (теми, кто проводит решения в жизнь, потребителями, конкурентами и публикой), т. е. следует определить класс возможных *контрстратегий*. Под контрстратегией мы будем понимать стратегию, которую может принять любой из указанных участников, причем такую, которая может уменьшить эффективность стратегий, реализуемых участником, принимающим решение. Эти контрстратегии очень важно учитывать, потому что выигрыш, который могут дать стратегии лица, принимающего решения, может обернуться проигрышем при наличии такой контрстратегии. Например, снижение цен может сначала вызвать повышенный спрос и увеличение прибылей. Однако если конкурент предпримет еще большее снижение цен, то это может вызвать войну цен или заставит владельца установить убыточный уровень цен.

Во многих случаях можно выяснить возможные контрстратегии путем непосредственного опроса участников относительно их реакции на определенную стратегию. Так можно поступать с сотрудниками предприятия и потребителями. Но к конкурентам такой подход неприменим. К счастью, обычно существует история конкурентной борьбы. С ее помощью можно предугадать возможные стратегии конкурента в новой обстановке.

Здесь, кстати, можно отметить, что контрстратегии принадлежат обычно к сфере предсказаний; не только участники системы, но и сама природа также может противодействовать людям, и если не предусмотреть это, то выбранная стратегия может оказаться губительной. Более подробно об этом говорится в главе 18.

Определение элементов задачи можно считать законченным после приведенного описания различных стратегий и контрстратегий. Теперь группа исследования операций может приступить ко второму этапу постановки задачи — превращению задачи, стоящей перед лицом, принимающим решение, в исследовательскую задачу.

ИССЛЕДОВАТЕЛЬСКАЯ ЗАДАЧА

Чтобы получить исследовательскую задачу, необходимо провести следующие операции.

1. Уточнить перечень целей, определенных на первом этапе постановки задачи.
2. Уточнить перечень различных стратегий достижения целей.
3. Определить меру (критерий) эффективности.

Прежде чем перейти к детальному рассмотрению каждого пункта, следует сделать несколько замечаний относительно логики процесса принятия решения. Рассмотрим упрощенную абстрактную задачу. Допустим, что существуют только две цели O_1 и O_2 ; возможны только две стратегии

C_1 и C_2 . Предположим теперь, что мы оценили эффективность каждой стратегии по отношению к достижению каждой цели (по шкале от 0 до 1). Результаты запишем в виде матрицы:

	O_1	O_2
C_1	0,8	0,4
C_2	0,2	0,6

Какую стратегию выбрать? Было бы ошибкой ответить, что следует выбирать C_1 или C_2 . На этот вопрос нельзя ответить, не зная относительной значимости этих целей. Если O_1 гораздо важнее O_2 , то следует выбрать C_1 ; но если O_2 гораздо важнее, чем O_1 , то надо выбрать C_2 . Как же можно сделать критерий выбора точным? Если бы мы могли измерить относительную значимость целей O_1 или O_2 , то можно было бы найти такой критерий. Например, предположим, что можно измерить относительную значимость по шкале от 0 до 1. При этом получили величины: 0,3 для O_1 и 0,7 для O_2 . Тогда можно взвесить эффективность каждой стратегии по отношению к каждой цели:

	O_1	O_2	Сумма
C_1	$0,3 \cdot 0,8 = 0,24$	$0,7 \cdot 0,4 = 0,28$	0,52
C_2	$0,3 \cdot 0,2 = 0,06$	$0,7 \cdot 0,6 = 0,42$	0,48

Сумму взвешенных эффективностей (эффективностей, умноженных на относительную значимость) для определенной стратегии назовем *общей эффективностью* этой стратегии. Эта величина и должна быть основной для выбора определенной стратегии.

Но чтобы определить эти величины, нужно уметь измерять значимость целей. Однако такие измерения выполнить очень трудно. Поэтому многие исследователи склонны выбирать наиболее важную цель и рекомендовать ту стратегию, которая наиболее эффективна для достижения этой цели. Как можно видеть из приведенного примера, такой метод ведет к неправильному решению задачи. Отказ от рассмотрения менее важных целей приводит не просто к упрощению задачи, а часто к значительным ошибкам.

В следующей главе приводится метод получения приблизительных оценок значимости целей. Дается соответствующий пример.

Здесь описывается элементарная логика процесса принятия решения. Если же учитываются контрстратегии и неопределенность, то нужна более сложная логика (см. главы 17 и 18).

УТОЧНЕНИЕ ПЕРЕЧНЯ ЦЕЛЕЙ И СТРАТЕГИЙ

Сложность исследования обычно зависит от числа целей и стратегий, которые следует учитывать. Поэтому желательно, прежде чем исследовать перечень различных целей и стратегий, по возможности укрупнить и упростить его. И хотя систематической процедуры для этого не существу-

ет, однако можно указать несколько принципов, установленных из практических соображений. Ниже рассматриваются некоторые из этих принципов.

Перечень целей. Для составления перечня целей надо укрупнить и упростить первоначальный перечень, полученный на первом этапе постановки задачи. Процедуру составления такого перечня можно разбить на 3 шага. Первый шаг заключается в том, чтобы исследовать, не является ли какая-либо цель только подцелью для достижения другой цели в перечне. Если да, то такую цель можно вычеркнуть из списка. Например, предположим, что одна из целей — увеличить ежегодную чистую прибыль компании, а другая — уменьшить производственные расходы. Очевидно, что уменьшение производственных расходов необходимо для увеличения чистой прибыли. Если это так, то уменьшение производственных расходов можно рассматривать как подцель, средство, а не цель.

Предположим, что приведены следующие две цели: увеличить чистую прибыль и увеличить сбыт. Вторая цель может совпадать с первой, но возможен вариант, когда вторая цель совпадает с первой только для некоторых стратегий, но не для всех. Например, цены можно снизить до такого уровня, что прибыли начинают падать, хотя сбыт увеличивается. Тогда компания может заявить, что она не заинтересована в увеличении сбыта, если это не влечет за собой увеличения прибылей. В этом случае вторая цель является средством для достижения первой цели, поэтому ее можно вычеркнуть из списка. Но возможен и такой случай, когда компания заинтересована в увеличении сбыта даже за счет уменьшения прибылей (в определенных пределах). В таких случаях эта цель является самостоятельной и остается в списке.

Второй шаг при составлении перечня целей заключается в исследовании каждой цели относительно различных стратегий, чтобы определить, возможно ли достижение каждой из целей с помощью представленных стратегий. Если нельзя достигнуть цели, то ее надо вычеркнуть из списка. Например, предположим, что одна из целей определяется как «повышение качества продукции», а различные стратегии включают только определение размера партии деталей. Если размер партии не влияет на качество, то поставленную цель нельзя достигнуть с помощью имеющихся средств, и ее следует вычеркнуть из списка.

Третий шаг необходим для объединения целей различных участников, если эти цели в основном совпадают. Например, и предприниматель и сотрудники могут быть заинтересованы в устойчивой занятости, либо предприниматель и потребитель могут быть заинтересованы в низких ценах и высоком качестве.

Перечень стратегий. Необходимо исследовать список возможных стратегий (вариантов решения), находящихся в распоряжении лица, принимающего решение, с тем, чтобы выяснить, нельзя ли некоторые варианты исключить из дальнейшего рассмотрения. В некоторых случаях предшествующие исследования показывают неприемлемость определенных стратегий. Иногда одна или несколько стратегий оказываются невозможными, так как при их использовании нарушаются некоторые поставленные ограничения. Например, решается задача о размещении предприятия, причем известна строго фиксированная максимальная величина земельного налога; тогда всякое решение, предполагающее более высокий налог, исключается из рассмотрения.

В некоторых случаях ограничения по времени или по средствам не дают возможности рассмотреть все варианты решения. Поэтому некоторые

варианты следует отбросить на том основании, что при имеющихся средствах и времени они не приведут к успешным результатам. Например, ставится задача сокращения издержек производства и хранения товаров. Большинство предложенных вариантов решения этой задачи связано с изменением производства. Но один вариант учитывает возможность изменения политики распределения товаров. Однако было установлено, что, даже если этот вариант приведет к хорошим результатам, все равно его нельзя реализовать за время, установленное компанией для завершения проекта и использования его результатов. Заключение новых контрактов и аннулирование старых с многочисленными потребителями было бы очень длительной процедурой, в то время как изменить внутреннюю политику компании можно сравнительно быстро.

Во всех случаях, когда какая-либо стратегия вычеркивается из списка, нужно зафиксировать причину, на основании которой это делается. Это позволит другим исследователям пересмотреть эти причины и устранить возможные ошибки.

ОПРЕДЕЛЕНИЕ МЕРЫ ОБЩЕЙ ЭФФЕКТИВНОСТИ

Как отмечалось выше, для измерения общей эффективности нужно определить (1) значимость целей и (2) эффективность стратегии. В этом параграфе мы рассмотрим, каким образом определяются эти элементы и устанавливается мера общей эффективности для задачи.

Выбор метода зависит от характера целей. Особенно важно знать, является ли цель количественной или качественной. Например, увеличить чистую прибыль — количественная цель (ее можно выразить в долларовой шкале), сохранить ведущее положение семьи в делах компании — является качественной целью (если ее не сформулировать иначе). Качественные цели обычно носят психологический или социальный характер, их часто называют неосозаемыми, потому что очень трудно измерить степень достижения этих целей. Поэтому желательно выражать цели в количественном виде. Ниже будет показано, что невозможно решать задачи, у которых все цели носят качественный характер.

Для задач, учитывающих качественные цели, гораздо труднее определить меру общей эффективности. Однако это не значит, что нельзя превратить качественные цели в количественные. История науки дает много примеров, когда свойство, рассматривавшееся в одну эпоху как чисто качественное, в другую эпоху оценивалось в количественном выражении. Одно время такие свойства, как «красный», «твердый», «образованный», «общительный», считались чисто качественными. Теперь мы знаем о них больше. Нет ни логических, ни методологических препятствий (хотя могут быть практические трудности) выразить количественно такие понятия, как «добрая воля», «моральное состояние», «ответственность».

Чтобы лучше понять, как измерять общую эффективность, рассмотрим задачи, связанные с выбором одной из двух стратегий (C_1 и C_2) при наличии двух целей: (1) увеличение чистой прибыли и (2) уменьшение среднего времени обслуживания. Допустим, нам известно, что (а) если применить стратегию C_1 , то чистая прибыль увеличится на \$1000 за год и среднее время обслуживания сократится на 2 дня, и (б) если применить стратегию C_2 , то чистая прибыль увеличится на \$2000 за год, а среднее время обслуживания сократится на 1 день. Как же выбрать одну из этих стратегий? Для этого надо сложить эффективность двух целей по каждой из стратегий. Но это возможно только в том случае, если удастся определить соотношение

между единицами увеличения чистой прибыли и сокращения среднего времени обслуживания. Например, если сокращение времени обслуживания на один день стоит \$500 в год, то можно измерить общую эффективность каждой стратегии. Общая эффективность C_1 равна $\$1000 + 2(\$500) = \$2000$. Общая эффективность C_2 равна $\$2000 + 1(\$500) = \$2500$. Следует выбрать C_2 .

Таким образом, первая задача по измерению общей эффективности свелась к определению однородной меры эффективностей для различных целей. Для этого необходимо «взвешивание» единиц, в которых измеряются различные цели, т. е. нужно определить относительные цены этих единиц. Такую относительную цену назовем *весом* единицы для данной цели.

Если мы действительно пользуемся выбранной стратегией в течение нескольких лет, то нельзя ожидать, что результат будет все время одинаков, ибо чистая прибыль и время обслуживания будут меняться. Поэтому будет неправильно говорить, например, что стратегия C_1 дает увеличение чистой прибыли на \$1000 за год. За несколько лет мы получим различные величины. Нужно как-то отразить это обстоятельство.

Отсюда вторая задача по измерению общей эффективности — найти способ отображения колебаний эффективностей (для каждой стратегии по отношению к каждой цели). Для этого необходимо построить функцию эффективности.

Если объединить веса и функции эффективности для каждой стратегии, то получим функцию общей эффективности. Следовательно, последняя задача заключается в том, чтобы получить критерий, с помощью которого можно выбрать одну из функций, т. е. определить, какая функция общей эффективности является наилучшей.

Ниже рассматривается каждая из этих задач, сначала для множества количественных целей, затем для множества, содержащего одну качественную и остальные количественные цели, и, наконец, для множества качественных целей.

Общая эффективность для количественных целей. В задачах, содержащих две или больше количественных целей, определение процедуры измерения общей эффективности является очень сложным. Для этого нужна различная информация, которую часто трудно получить. Поэтому не всегда можно воспользоваться рекомендуемой процедурой. Но точная формулировка этой процедуры послужит основой, с помощью которой можно попытаться найти наилучшее решение практических задач.

Покажем, какие необходимы операции для измерения общей эффективности. Одновременно приведем примеры различных практических ограничений. Но прежде чем подробно разбирать эти операции, перечислим их.

1. Определить меру эффективности по отношению к каждой цели.
2. Если меры эффективности, полученные в первом шаге, различны, то привести их к одной общей или стандартной мере эффективности.
3. Для каждой стратегии по каждой цели определить вероятность достижения каждого возможного уровня эффективности. При этом получим функцию эффективности каждой стратегии для каждой цели.
4. Для каждой стратегии сложить функции эффективности, чтобы получить функцию общей эффективности каждой стратегии по всем целям.
5. Выразить цель процесса принятия решения в виде максимизации или минимизации ожидаемого дохода, прибыли (положительного дохода) или потерь (отрицательного дохода).

6. Построить функцию дохода для каждой стратегии. Функция дохода выражает ожидаемый результат (результат, умноженный на вероятность его получения) в виде прибылей и потерь.

Если все цели в перечне выражены в количественном виде (например, «увеличить чистую прибыль», «увеличить сбыт», «уменьшить среднее время

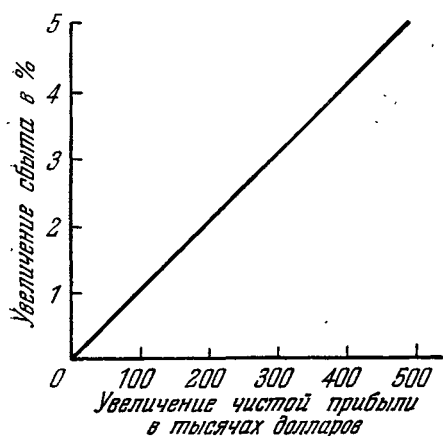


Рис. 5.1. Зависимость увеличения сбыта от чистой прибыли.

обслуживания»), то можно использовать следующую процедуру для «взвешивания» целей.

1. Определить единицы, в которых измеряется степень достижения каждой цели. Например, вышеперечисленные цели можно выразить следующим образом: (а) увеличить чистую прибыль на x долларов, (б) увеличить сбыт на y процентов, (в) уменьшить среднее время обслуживания на z дней.

2. Выбрать две цели: наиболее важную и следующую за ней по важности. Построить график, на котором наиболее важная цель «изображена» в соответствующих единицах по оси абсцисс, а другая цель — по оси ординат. Затем выбрать несколько значений по оси абсцисс, определить экви-

валентные величины по оси ординат и нанести на чертеж полученные точки. Соединить эти точки между собой, в результате чего получим прямую или кривую. Предположим, например, что увеличение сбыта на один процент

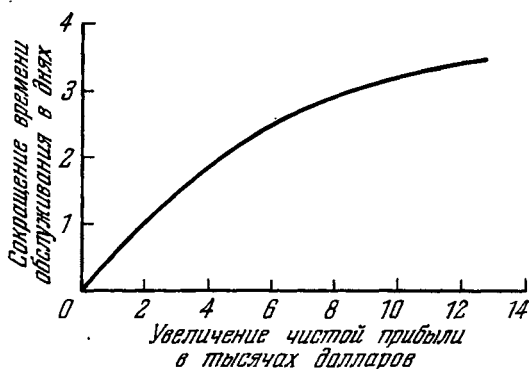


Рис. 5.2. Зависимость времени обслуживания от чистой прибыли.

соответствует \$100 000 чистой прибыли. Тогда получим прямую, показанную на рис. 5.1. С помощью этого графика можно каждой величине увеличения сбыта в процентах ставить в соответствие доллары.

Может оказаться, что уменьшение времени обслуживания становится более важной целью, чем увеличение сбыта. Тогда получаем график, показанный на рис. 5.2.

3. Повторить шаг 2 для других целей, сравнивая единицы, в которых выражены цели, расположенные по оси ординат, с единицами наиболее важной цели.

Во всех приводимых графиках наиболее важная для данной задачи цель является стандартной (т. е. всегда изображается по оси абсцисс). Это объясняется тем, что в процессе «взвешивания» всегда наблюдаются некоторые ошибки. Чтобы уменьшить эти ошибки, надо все сравнения приводить относительно одной и той же величины.

Иногда приводятся возражения против подобной методики, основанные на том, что ошибка преобразования может быть настолько велика, что имеет смысл просто исключить из рассмотрения все цели, кроме наиболее важной. Но это необоснованное утверждение. Ошибка преобразования обязательно будет меньше, чем ошибка, полученная при исключении неосновных целей. Такое исключение эквивалентно приписыванию определенным стратегиям нулевой эффективности по исключаемым целям. Это вызовет гораздо большую погрешность в расчетах, чем ошибка преобразования.

Во многих случаях соотношение между единицами устанавливается на основе изучения данных, полученных из прошлых задач; в других случаях полагаются на оценки, данные различными специалистами. Рассмотрим сначала случай изучения прошлого опыта.

Взвешивание единиц на основе изучения прошлого опыта. Электрическая компания заинтересована в уменьшении стоимости ремонтных работ, выполняемых для клиентов, но она также заинтересована в сокращении для клиентов времени ожидания своей очереди по ремонту. Можно рассчитать стоимость обслуживания при любой политике управления делами компании. Для каждой политики можно рассчитать также среднее время ожидания на вызов. Теперь необходимо сложить эти две оценки, для чего необходимо преобразование единиц (долларов или времени). Следует выяснить, во сколько обходится для компании минутное ожидание клиента? Как оценивает компания время ожидания клиентов?

Чтобы ответить на эти вопросы, можно изучить предыдущий опыт компании. В течение пяти предшествующих лет компания дважды меняла свою политику, так что можно выяснить, каковы при каждом изменении политики потери или выигрыш стоимости обслуживания, а также изменение среднего времени ожидания клиента. С помощью этих цифр можно было определить, во сколько обошлось компании сокращение или увеличение среднего времени ожидания. Полученные в результате расчетов цифры были достаточно стабильными. Эти результаты представляли собой стоимостную оценку руководством компании времени ожидания клиента и давали возможность перевести временную шкалу в денежную.

Следует отметить, что таким способом нельзя определить оценку времени ожидания клиентом. Но можно определить, сколько платит клиент в среднем за пользование электричеством в течение минуты. Можно использовать и другой метод приближенного расчета, изложенный в следующей главе.

Взвешивание единиц с помощью экспертных оценок. Если нет таких данных для расчетов, как в приведенном примере, то можно использовать оценки специалистов. Например, можно просто прибегнуть к консультации опытного специалиста, сколько долларов стоит единица времени ожидания. Или можно воспользоваться процедурой, рекомендуемой в следующей главе, для чего надо сначала взять различное число единиц по каждой цели, чтобы получить приблизительно сравнимые веса. Затем эти единицы взвешиваются согласно приведенной в следующей главе процедуре.

Результаты объединяются в следующую таблицу:

Взвешиваемые величины	Вес
\$ 1 000 чистой прибыли в год	0,5
1% увеличения сбыта	0,3
1 день сокращения среднего времени обслуживания	0,2

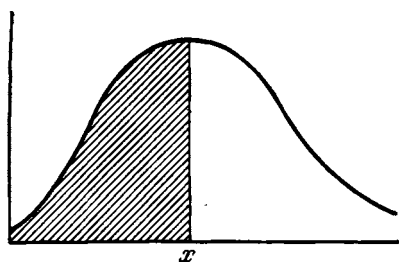
Используя приведенные веса, можно получить эквиваленты различных единиц. Например, увеличение сбыта на 1% эквивалентно

$$(0,3/0,5) \$1000 = \$600.$$

Аналогично можно найти долларовый эквивалент сокращения среднего времени обслуживания на один день:

$$(0,2/0,5) \$1000 = \$400.$$

Построение функции эффективности. Предположим, что получены следующие результаты для трех целей: увеличение сбыта на 0,01% эквивалентно \$1,0 чистой прибыли, сокращение среднего времени обслуживания на 0,10 дня эквивалентно \$1,0 чистой прибыли. Теперь



Стандартная мера эффективности

Рис. 5.3. Функция эффективности.

предположим, что определенная стратегия гарантирует увеличение чистой прибыли на \$1000, увеличение сбыта на 1% и сокращение среднего времени обслуживания на 1 день. Можно подсчитать общую эффективность этой стратегии:

$$\$1000 + (1/0,01) \$1,00 + (1/0,10) \$1,00 =$$

$$= \$1000 + \$100 + \$10 = \$1110.$$

Однако, как уже отмечалось, в этой процедуре есть существенный недостаток. Если выбранная стратегия применяется многократно, то маловероятно, что результаты будут каждый раз одинаковые, так что чистая прибыль, увеличение сбыта и сокращение ожидания будут меняться при сохранении одной и той же стратегии.

Чтобы уметь «управлять» определенной стратегией, следует рассмотреть эту стратегию по отношению к каждой цели отдельно и попытаться оценить вероятность успеха по каждой из этих целей. Результаты можно объединить с целью получения меры общей эффективности данной стратегии.

Прежде всего необходимо построить «функции эффективности», которые показывают вероятность того, что будет достигнута определенная эффективность (измеренная по стандартной шкале) по каждой цели (см. рис. 5.3). Вероятность того, что при данной стратегии будет достигнута эффективность, не большая чем x , выражается отношением заштрихованной площади кривой ко всей площади кривой *.

Мера эффективности может выражаться в деньгах, времени, затратах рабочей силы, процентах выполненной работы и любой комбинации из этих величин.

* Эта кривая называется кривой плотности вероятности; ее можно изображать в виде функции $f_{ij}(x)$, где i означает определенную стратегию C_i , а j означает определенную цель Q_j .

Для каждой стратегии и каждой цели надо построить свою функцию эффективности. Так, для трех стратегий и четырех целей существует $12 (3 \times 4)$ функций эффективности.

После этого следует объединить все функции эффективности для каждой стратегии. Математически такое объединение может быть очень сложным, но с точки зрения логики это простая операция.

Предположим, например, что существуют две цели O_1 и O_2 и одна стратегия S_1 , причем надо определить вероятность получить прибыль в \$10 для обеих целей. Можно показать 11 различных вариантов получения такой прибыли (предполагается, что прибыль выражается в целых числах) (см. таблицу 5.1).

Таблица 5.1

Варианты	Минимально возможная прибыль для цели O_1 , долл	Минимально возможная прибыль для цели O_2 , долл	Минимально возможная общая прибыль, долл	Варианты	Минимально возможная прибыль для цели O_1 , долл	Минимально возможная прибыль для цели O_2 , долл	Минимально возможная общая прибыль, долл
1	0	10	10	7	6	4	10
2	1	9	10	8	7	3	10
3	2	8	10	9	8	2	10
4	3	7	10	10	9	1	10
5	4	6	10	11	10	0	10
6	5	5	10				

Можно подсчитать вероятность каждого из 11 вариантов, для чего надо перемножить соответствующие вероятности для обеих целей. Например, вероятность первого варианта $P(0, 10)$ равна вероятности $P_1(0) \cdot$ получения прибыли \$0 для цели O_1 , умноженной на вероятность $P_2(10)$ получения прибыли \$10 для цели O_2 , т. е.

$$P(0, 10) = P_1(0) P_2(10).$$

Общая вероятность получения прибыли точно в \$10

$$P(0, 10) + P(1, 9) + \dots + P(10, 0) = \\ = P_1(0) P_2(10) + P_1(1) P_2(9) + \dots + P_1(10) P_2(0).$$

Функция общей эффективности ** при этом выражает общую вероятность по отношению к любой стандартной мере эффективности***.

«Оптимальная» функция общей эффективности. Предположим, что с вероятностью 0,5 мы можем получить доход в \$1 и с вероятностью 0,5 можем ничего не получить. Тогда ожидаемый доход выражается

$$0,5 (\$1,00) + 0,5 (0) = \$0,50.$$

* Предполагается, что получение прибылей для одной цели не зависит от получения прибылей для другой цели.

** Вывод этой функции приведен подробно в приложении, в примере 1 в конце этой главы.

*** Рассуждения и примеры, приведенные в данном разделе, относятся только к классу систем, в которых эффективности отдельных целей входят в функцию эффективности аддитивно. Это обстоятельство очень сильно ограничивает применимость предлагаемых процедур. (Прим. ред.)

В общем случае, когда можно получить n различных доходов $(x_1, x_2, \dots, x_n)$, причем для каждого из них задана вероятность $(p_1, p_2, \dots, p_n)$ его получения, ожидаемый доход можно вычислить из выражения *

$$p_1x_1 + p_2x_2 + \dots + p_nx_n = \sum_{i=1}^n p_ix_i.$$

Один из возможных методов выбора оптимальной стратегии заключается в выборе такой стратегии, при которой ожидаемый доход будет наибольшим. Но это не единственный принцип выбора.

Ожидаемый доход может быть отрицательным, т. е. данная стратегия приводит к потерям, а не прибыли. Например, если стандартная шкала выражена в долларах, то функцию дохода можно изобразить в виде кривой, представленной на рис. 5.4.

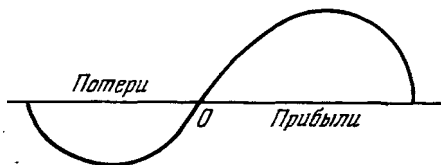


Рис. 5.4. Функция дохода.

Теперь можно вычислить отдельно ожидаемую прибыль и ожидаемые потери. Функции общей эффективности можно сравнивать между собой по прибылям или по потерям. Оптимальную функцию можно определить,

исходя из максимизации ожидаемой прибыли или минимизации ожидаемых потерь. Возможны различные определения «оптимальной» функции, но обычно используются следующие три: максимальный ожидаемый доход, максимальная ожидаемая прибыль, минимальные ожидаемые потери.

Если встает задача выбора стратегии на длительный период времени, то можно выбрать ту стратегию, которая обеспечивает максимальный ожидаемый доход в течение длительного времени, что вполне оправдано, когда допустимы большие, но кратковременные потери.

Но предприниматель не всегда может пойти на то, чтобы допустить большие потери. Например, он может иметь ограниченный капитал. Если он рискнет пойти на потери, превосходящие его капитал, то он может оказаться банкротом. В таком случае он заинтересован в выборе такой стратегии, которая или минимизирует ожидаемые потери, или дает наименьшую вероятность того, что потери превзойдут определенный критический уровень.

Возможен и другой случай, когда предприниматель делает определенную рискованную ставку, т. е. он допускает возможность потери определенной суммы, но при этом цель заключается в том, чтобы получить наибольший возможный выигрыш. В этом случае выбор стратегии определяется максимизацией ожидаемой прибыли. Примером может служить разведка нефти определенной нефтяной компанией. Компания устанавливает бюджет на несколько лет. Производственно-плановый отдел должен вложить эти деньги таким образом, чтобы максимизировать свои шансы на открытие крупных нефтяных месторождений. «Крупное» месторождение можно определить как «приносящее больше чем x долларов прибыли». В этом случае следует выбирать ту стратегию, которая максимизирует вероятность того, что ожидаемая прибыль будет не меньше чем x долларов.

В общем случае, если руководство склонно к осторожной стратегии, то критерий выбора «оптимальной» функции общей эффективности должен включать минимизацию ожидаемых потерь. Если же стратегия сходна с «игрой на повышение», то наиболее подходящей целью является максими-

* Математический вывод для непрерывной функции приведен в приложении, пример 1 в конце этой главы.

зация ожидаемой прибыли. Если же компания заинтересована в проведении успешной долговременной стратегии и обладает необходимыми ресурсами, то следует добиваться максимизации ожидаемого дохода.

Определение, какая стратегия применима, является частью анализа целей руководства. Выбор оптимальной стратегии является метациелью, поскольку этот выбор определяет успех деятельности компании. Эта цель не рассматривается на начальных этапах постановки задачи, но ее используют для нахождения «оптимального» решения.

Практический вариант для количественных целей. На практике основная трудность при использовании описанной процедуры заключается в построении функции эффективности, т. е. в оценке вероятности получения различного дохода относительно каждой цели по каждой стратегии. Обычно гораздо легче и экономичнее оценивать средний доход, поскольку требуется гораздо меньше данных. Поэтому обычно на практике пользуются несколько упрощенной процедурой.

1. (Как и в полной процедуре) определить меру эффективности по каждой цели.
2. (Как и в полной процедуре) привести меры эффективности по разным целям к одной общей мере.
3. Для каждой стратегии определить среднюю эффективность по каждой цели и привести их к общей мере эффективности.
4. Для каждой стратегии сложить средние эффективности по всем целям.
5. Выбрать стратегию, имеющую максимальную среднюю общую эффективность (или минимальную общую отрицательную эффективность — потерю).

Для иллюстрации этой процедуры рассмотрим вновь пример с электрической ремонтной компанией. Возможны две стратегии: C_1 — продолжать работать с тремя гаражами и обслуживать три определенных городских района и C_2 — добавить один гараж, использовать то же число сотрудников, но обслуживать четыре новых района, выбранных отделом обслуживания. Компания стремится достичь двух целей: O_1 — минимизировать эксплуатационные затраты и O_2 — минимизировать время ожидания клиентов. Поставленная задача решается следующим образом.

1. Мера эффективности для O_1 определяется как общая ежегодная стоимость обслуживания. Мерой эффективности для O_2 является среднее время ожидания на вызов.
2. Как уже описывалось, изучение прошлых решений привело к заключению, что минутное время ожидания всех клиентов обходится для компании примерно в \$850.
3. Анализ затрат показал, что стратегия C_2 по сравнению со стратегией C_1 (а) увеличивает стоимость обслуживания в среднем на \$3247 в год и (б) уменьшает время ожидания клиента в среднем на 2,1 минуты в год, что оценивается в \$1832.

Таким образом, C_2 дает в среднем ежегодное увеличение стоимости обслуживания на $(\$3247 - \$1832) = 1415$.

Следовательно, надо рекомендовать стратегию C_1 .

В приложении в конце этой главы дается сжатое изложение различных способов формирования меры общей эффективности по отношению к количественным целям.

Общая эффективность для качественных целей. Если все цели в списке являются качественными, то для их взвешивания можно применить способ приближенного измерения, приведенный

в следующей главе. По этому способу каждой цели назначается относительная оценка v_i , причем сумма всех оценок равна 1.

Как уже указывалось, эффективность качественной цели можно измерить в виде вероятности успешного достижения этой цели. Например, существует цель: «выпустить партию деталей без дефектов». Наличие дефектов означает невыполнение цели.

Следовательно, при производстве любой партии цель или выполняется, или нет. Эффективность такой цели можно просто выразить через вероятность ее достижения.

Однако такая мера эффективности подвержена определенным ограничениям. Предположим, что две стратегии C_1 и C_2 характеризуются одинаковой вероятностью достижения желаемой цели, но одна из них более дорогая. Тогда очевидно, что она менее эффективна, если ставится также цель минимизации стоимости. В таких ситуациях можно действовать двояко.

1. Можно измерять эффективность каждой стратегии при определенных ограничениях по стоимости, времени и т. д. При этом получим следующую формулировку: какова эффективность C_1 и C_2 при условии, что издержки должны быть одинаковы?

2. Минимизацию стоимости можно включить как отдельную цель и оценивать стратегии по этой цели, как в случае количественных целей. В этом случае следует использовать приведенный ниже метод, разработанный для смешанных количественных и качественных целей.

После того как определена эффективность E_{ij} каждой стратегии для каждой цели и назначена нормированная оценка для каждой цели v_j , можно определить общую эффективность каждой стратегии в виде суммы

$$\sum_{j=1}^n E_{ij}v_j.$$

В такой задаче общая эффективность выражается одним числом — суммой взвешенных эффективностей. Непосредственное сравнение этих чисел дает возможность выбрать наиболее эффективную стратегию.

Общая эффективность для случая количественных и качественных целей. В некоторых случаях одна или несколько (но не все) целей являются качественными (например, «сохранить ведущее положение семьи в делах компании» или «завязать нужные связи в обществе»). В таких случаях необходима другая процедура нахождения критерия общей эффективности.

Предположим, у нас есть три количественные цели (увеличение чистой прибыли, увеличение сбыта и сокращение среднего времени обслуживания), а также четвертая цель: «сохранить ведущее положение семьи в делах компании». Веса первых трех целей можно определить согласно приведенной выше методике. Предположим, что в качестве стандартной единицы используются доллары за год. Следовательно, к этой единице надо привести и четвертую цель.

Можно было бы получить оценку четвертой цели в этих единицах с помощью изучения действительного состояния дел, чтобы выяснить, во сколько обходится сохранение ведущего положения в делах компании. Но предположим, что это или невозможно, или практически неосуществимо. В таком случае нужно воспользоваться мнением осведомленных лиц, стремясь при этом получить достаточно надежные сведения. Для этого надо узнать мнение осведомленных лиц по трем вопросам.

1. Во сколько долларов в год «оценивается» ведущее положение семьи?

2. Скольким процентам увеличения сбыта в год соответствует ведущее положение семьи?

3. В скольких днях сокращения среднего времени обслуживания за год можно оценить контрольное положение семьи?

Чтобы оценить надежность ответов, надо перевести каждый из ответов в доллары с помощью известных весов единиц различных целей. Если после такого превращения получаются противоречивые величины, то таким ответам нельзя доверять.

Как уже указывалось, для определения эффективности стратегии относительно качественной цели можно измерить вероятность достижения цели с помощью этой стратегии. Таким образом, одна величина, характеризующая вероятность, может определять эффективность стратегии для качественной цели. Например, если ведущее положение семьи в компании оценивается в \$100 000 в год и принятая стратегия обеспечивает с вероятностью 0,8 сохранение этого положения, то ожидаемый доход равен \$80 000. В данном случае ожидаемый доход выражается одной величиной. Среднее, максимальное и минимальное значения в таких случаях теряют смысл.

Для первых трех целей, как и прежде, можно определить средний ожидаемый доход, максимальные ожидаемые прибыли, минимальные ожидаемые потери и другие величины. Затем можно сложить ожидаемый доход для качественной цели и общий ожидаемый доход относительно количественных целей. Эта процедура носит общий характер и справедлива для любого числа и любой комбинации качественных и количественных целей.

Перечислим основные этапы определения меры общей эффективности для смешанных целей.

1. Определить общий ожидаемый доход для количественных целей.
2. Определить вероятность P_{ij} успеха каждой стратегии C_i для каждой из k' качественных целей.
3. Определить оценку каждой качественной цели v'_j , измерив ее по шкале стандартной эффективности, используемой на шаге 1.
4. Определить общий ожидаемый доход (R_i) для каждой стратегии C_i *.

КТО ОПРЕДЕЛЯЕТ ВЕСА ЦЕЛЕЙ?

При определении весов для различных целей в перечне целей очень важно, чтобы в этой процедуре участвовали различные группы специалистов или представители разных групп. После того, как они произведут

* Нетрудно видеть, что

$$*R_i = \sum_{j=1}^{k'} P_{ij} v'_j + \int_{-\infty}^{\infty} x g_i(x) dx.$$

Аналогично можно определить ожидаемую прибыль:

$$G_i = \sum_{j=1}^{k'} P_{ij} v'_j + \int_0^{\infty} x g(x) dx,$$

где включаются только те цели O_j , для которых $v'_j > 0$, и ожидаемые потери выражаются в виде

$$L_i = \sum_{j=1}^{k'} P_{ij} v'_j + \int_{-\infty}^0 x g(x) dx,$$

где учитываются только те цели, у которых $v'_j < 0$.

оценки, важно суметь правильно объединить их оценки. Для этого существует следующая процедура.

Допустим, v_{ij} представляет собой вес, который j -я группа присвоила i -й цели. Теперь предположим, что можно оценить компетентность каждой группы специалистов, причем компетентность j -й группы оценивается как V_j . Если всего таких групп n , то общий вес i -й цели определяется по формуле

$$V_1 v_{i1} + V_2 v_{i2} + \dots + V_n v_{in} = \sum_{j=1}^n V_j v_{ij}.$$

Например, в оценке участвуют три группы специалистов, которые оцениваются в следующих числах: $V_1 = 0,5$, $V_2 = 0,3$ и $V_3 = 0,2$. Они определили вес цели O_i следующим образом: $v_{i1} = 0,8$, $v_{i2} = 0,3$ и $v_{i3} = 0,6$. Тогда общий вес цели O_i равен

$$0,5 \cdot 0,8 + 0,3 \cdot 0,3 + 0,2 \cdot 0,6 = 0,40 + 0,09 + 0,12 = 0,61.$$

Или предположим, что в задачу этих групп входит также преобразование временных оценок в денежные. Получены оценки: $v_{i1} = \$1000$, $v_{i2} = \$1500$ и $v_{i3} = \$1200$. Общий вес цели O_i равен

$$0,5 \cdot 1000 + 0,3 \cdot 1500 + 0,2 \cdot 1200 = 500 + 450 + 240 = \$1190.$$

В данной процедуре предполагается, что известна оценка компетентности различных групп специалистов. Главная трудность заключается даже не в том, чтобы определить эту оценку — для этого можно применить приведенный в следующей главе метод взвешивания. Главное — найти, кто будет давать такую оценку. И хотя ответ может оказаться неприятным для тех, кто против научного подхода к вопросам оценок, следует отметить, что эти задачи должна решать группа исследования операций. Эта группа не может пройти мимо поставленных задач оценки. Вопрос лишь в том, насколько рационально она будет решать эти задачи. В самом деле, если группа исследования операций попытается «уклониться» от таких задач и примет во внимание только мнение группы заказчиков, то тем самым она припишет этой группе максимальную оценку (1), а оценки всех остальных участников будет считать равными нулю. Поэтому такие оценки следует назначать на основании анализа, обсуждать их, если необходимо, и впоследствии изменять их.

Совершенно очевидно, что следует стремиться к полному и правильному определению весов различных целей. Но хотя это и слишком сложная задача, все же так называемый «практический» выход — опустить эти вопросы — оказывается слишком дорогим. Правда, описанная процедура оценки групп специалистов дает очень приблизительные данные, но вызванные ими ошибки оказываются гораздо меньше, чем ошибки, полученные без учета мнений различных специалистов.

ЗАКЛЮЧЕНИЕ

В заключение дадим краткое изложение основных процедур, связанных с постановкой задачи, которые приведены в этой и предшествующей главах.

А. Проанализировать информационную систему и соответствующие операции, отображенные в этой системе.

1. Найти и проследить связи, соответствующие изучаемым операциям.
2. Выявить все процессы преобразования информации и принятия решений.

3. Определить каждый шаг соответствующих операций.

4. Отбросить все связи и процессы преобразования, которые не влияют на операции (например, реклама в производственных операциях).

5. Сгруппировать операции по органам управления.

6. Построить схему информационной модели, отображающую: (а) органы управления и принимаемые решения, (б) потоки информации между органами управления и потребляемое на передачу информации время, (в) материальные потоки и продолжительность сгруппированных операций.

Б. Поставить задачу руководства.

1. Указать лиц, принимающих решение, и процедуру принятия решения.

2. Определить цели, стоящие перед лицами, принимающими решение.

3. Выявить неосновных участников и указать, как они влияют на решение.

4. Определить цели неосновных участников.

5. Определить различные стратегии, с помощью которых можно принимать решения.

6. Определить контрстратегии со стороны неосновных участников.

В. Поставить исследовательскую задачу.

1. Составить перечень основных целей.

2. Составить перечень основных стратегий.

3. Определить меру общей эффективности: (а) определить меру эффективности по каждой цели, (б) взвесить цели (качественные) или единицы, в которых измеряются цели (количественные), (в) определить критерий оптимального решения в виде функции от суммы взвешенных эффективностей (например, максимум ожидаемого дохода, минимум ожидаемых потерь).

Приложение

РАЗЛИЧНЫЕ СПОСОБЫ ИЗМЕРЕНИЯ ОБЩЕЙ ЭФФЕКТИВНОСТИ ДЛЯ КОЛИЧЕСТВЕННЫХ ЦЕЛЕЙ

Рассмотрим для примера случай с двумя стратегиями C_1 и C_2 и двумя целями O_1 и O_2 . Предположим, что x_{ij} представляет собой некоторую величину эффективности стратегии C_i для цели O_j по шкале X_j .

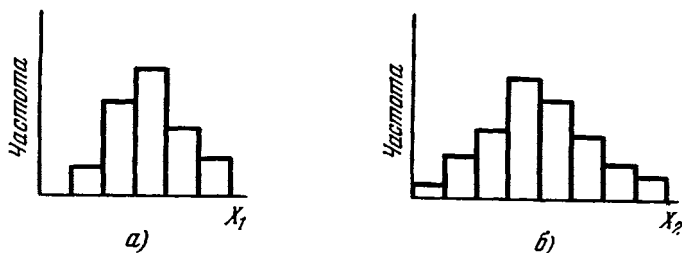


Рис. 5.5. Гистограммы.

а) Эффективность C_1 для O_1 . б) Эффективность C_1 для O_2 .

Описанные процедуры можно распространить на случай m целей и n стратегий.

В а р и а н т 1.

1.1. Применить стратегию C_1 (непосредственно, на основе прошлого опыта или путем моделирования) и определить величины x_{11} и x_{12} , полученные в каждом примере.

1.2. Построить по полученным данным две гистограммы (рис. 5.5).

1.3. Построить плотности вероятности по каждой гистограмме. Они называются функциями эффективности и изображаются через $f(x_{11})$ и $f(x_{12})$ (см. рис. 5.6).

1.4. Выбрать одну шкалу эффективности за стандартную и определить преобразование различных единиц эффективности в единицы стандартной

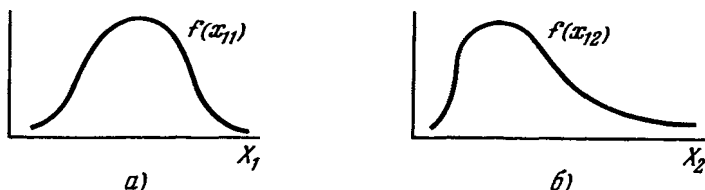


Рис. 5.6. Плотность вероятности.

а) Функция эффективности C_1 для O_1 . б) Функция эффективности C_2 для O_2 .

шкалы. Например, если x_1 является стандартной шкалой, то найти функцию h такую, что $x_{11} = h(x_{12})$.

1.5. Преобразовать функции эффективности для всех целей (кроме той, шкала которой выбрана за стандартную), для чего перевести различные единицы эффективности в единицы стандартной шкалы. В приведенном



Рис. 5.7. Функция общей эффективности.

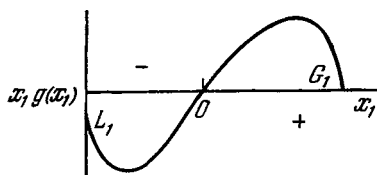


Рис. 5.8. Функция дохода.

примере $f(x'_{12}) = f[h(x_{12})]$, где $f(x'_{12})$ есть преобразованная функция эффективности.

1.6. Определить плотность вероятности для $x_1 = x_{11} + x'_{12}$, которая называется «функцией общей эффективности» и показана на рис. 5.7.

$$g(x_1) = \int_{a=-\infty}^{\infty} f(x_{11}=a) f(x'_{12}=x_1-a) da.$$

1.7. Определить для C_1 :
ожидаемые потери

$$(L_1) = \int_{-\infty}^0 x_1 g(x_1) dx_1 \text{ (отрицательная эффективность),}$$

ожидаемую прибыль

$$(G_1) = \int_0^{\infty} x_1 g(x_1) dx_1 \text{ (положительная эффективность),}$$

ожидаемый доход

$$(R_1) = G_1 + L_1 = \int_{-\infty}^0 x_1 g(x_1) dx_1 + \int_0^{\infty} x_1 g(x_1) dx_1 = \int_{-\infty}^{\infty} x_1 g(x_1) dx_1.$$

Построить «функцию дохода» (см. рис. 5.8).

1.8. Повторить шаги от 1.1 до 1.7 для C_2 .

1.9. Выбрать критерий оптимального решения (например, максимальный ожидаемый доход).

1.10. Выбрать стратегию, удовлетворяющую этому критерию.

В а р и а н т 2.

2.1. Преобразовать X_2 в X_1 , положив $x_{11} = h(x_{12})$.

2.2. Применить стратегию C_1 и для каждой пары значений x_{11} и x_{12} вычислить $x_{11} + h(x_{12})$.

2.3. Используя величины $x_{11} + h(x_{12})$, построить гистограммы, как в шаге 1.2.

2.4. Построить плотность вероятности, как в шаге 1.3. Полученная функция является функцией общей эффективности $g(x_1)$.

2.5. Выполнить шаги от 1.7 до 1.10.

В а р и а н т 3 (нет достаточного количества данных для построения плотности вероятности).

3.1. Выполнить шаг 2.1.

3.2. Выполнить шаг 2.2.

3.3. Вычислить среднее для пары значений (преобразованных и просуммированных). Получим оценку величины средней общей эффективности (СОЭ).

3.4. Использовать в качестве критерия оптимального решения максимум средней величины общей эффективности (МСОЭ).

3.5. Вычислить СОЭ для каждой стратегии и выбрать стратегию с МСОЭ.

ЛИТЕРАТУРА

1. Ackoff Russell L., The Design of Social Research, University of Chicago Press, Chicago, 1953.
2. Bross Irwin D. J., Design for Decision, The Macmillan Co., New York, 1953.
3. Dewey John, How We Think, D. C. Heath & Co., Boston, 1933.
4. Dewey John, Logic: The Theory of Inquiry, Henry Holt & Co., New York, 1938.
5. Wilson E. B., Jr., An Introduction to Scientific Research, McGraw-Hill Book Co., New York, 1952.

Глава 6

ОПРЕДЕЛЕНИЕ ВЕСОВЫХ КОЭФФИЦИЕНТОВ ЦЕЛЕЙ

ВВЕДЕНИЕ

В предыдущей главе было показано, что строго обоснованная процедура решения задач, в которых ставится более одной цели, требует какого-то общего метода, позволяющего приписывать некоторые относительные оценки (веса) рассматриваемым целям. Было также показано, что когда речь идет о целях количественного характера, может оказаться необходимым

определить некоторые веса для различных интервалов вдоль шкал, определяющих цели (например, 1 день задержки в поставке, улучшение на 1% положения на рынке сбыта).

Необходимые относительные веса могут быть заданы в стоимостном выражении, например в долларах, если просто оценить каждую цель некоторым числом долларов. Такая оценка обладает очевидными преимуществами: она легко понятна, «объективна» и общепотребительна. Но трудности, возникающие при использовании денежных шкал, также очевидны. Многие цели невозможно оценить в долларах. Во многих случаях мы оцениваем по-разному две цели, которых можно достигнуть при одной и той же затрате средств. В других случаях чрезвычайно трудно оценить в денежной форме соответствующие цели. Так, например, какова истинная стоимость травмы, жизни, срыва поставки, потери «доброй воли»? Подобные цели в коммерческих операциях были неудачно названы «неуловимыми», что эквивалентно понятию «недоступности». Хотя и верно, что при долларовом методе измерения к таким целям действительно невозможно «подступиться», их можно тем не менее измерить (оценить) некоторыми другими методами, один из которых будет изложен в настоящей главе.

Наконец, даже в случае, когда цель можно оценить в денежном выражении, нередко для определения соответствующей величины приходится преодолевать существенные трудности или тратить большие средства, да и точность ее может быть сомнительной. Так, затраты, связанные с хранением запасов, затраты на обеспечение сбыта товаров и затраты, связанные с распределением товаров в торговой сети, в некоторых отраслях промышленности крайне неопределенны. Вместе с тем они чрезвычайно важны для окончательных решений руководителей. Вообще говоря, не будет преувеличением считать, что операционисты постоянно сталкиваются с задачей определения адекватных оценок затрат и прибылей, проводя исследования в так называемых стимулируемых прибылью организациях.

Тем не менее почти во всех операционных исследованиях используются именно денежные шкалы. Как же удастся преодолеть трудности, связанные с необходимостью приписать определенную цену в долларах различным целям? Обычно для решения этой задачи применяются два метода.

1. Рассмотрим сначала только те цели, в соответствие которым можно поставить точно и объективно некоторую денежную цену. Построим модель операции, для которой функция потерь определяется через одну из таких целей. Найдем с помощью модели алгоритм, обеспечивающий отыскание оптимального решения в смысле заданных целей. Представим этот алгоритм на примере управления фирмой, которое оценивает алгоритм с точки зрения «неуловимых» целей и изменяет его, исходя из опыта и субъективных соображений.

Пример. Рассмотрим задачу оптимизации производства, сформулированную в терминах минимизации общих средних затрат на хранение запасов и затрат на подготовку производства. Модификация решения сводится к изменению найденного плана исходя из возможных колебаний потребностей в рабочей силе. Это означает, что найденный первоначально план может предусматривать наем большого числа рабочих в начале каждого месяца и последующее увольнение к концу месяца. «Неуловимая» цель стабилизации численности штата работников вводится на основании здравого смысла, и план изменяется таким образом, что предусматривается постоянная численность рабочих в течение всего месяца, хотя такое «обобщение» и обуславливает некоторое увеличение общих усредненных затрат на хранение запасов и подготовку производства.

2. Построим модель, в которой «неуловимая» цель фигурирует, скажем, в виде линейного члена. Так, если дефицит рассматривается как недоступная для измерения величина, то в модель можно ввести потери из-за дефицита x изделий как линейный член C_1x . Общие затраты при этом определяются выражением

$$TC = C_1x + \text{прочие затраты.}$$

Но C_1 является неизвестным. Однако можно найти минимум TC по x . Так, например, предположим, что можно взять производную TC по x и что уравнение

$$\frac{d(TC)}{dx} = 0$$

определяет искомый минимум. Тогда имеется возможность вычислить значение x , доставляющее минимум TC , а также вычислить само значение минимума TC . Эти минимальные общие затраты будут теперь функцией только одной величины C_1 . Если утверждать, что фирма действительно применяла оптимальную стратегию, то можно считать реальные затраты оценкой минимума TC . Зная этот минимум, можно разрешить уравнение относительно C_1 . Полученное значение C_1 и есть та цель, которой стремилась достигнуть фирма, независимо от того, была ли эта величина известна или нет, естественно, при условии, что все принятые ранее решения были оптимальными. Недостаток такого метода оценки неуловимых целей заключается в том, что нужно обязательно предположить наличие оптимума (чтобы вообще дать оценку). Однако этот метод обладает и существенным несомненным достоинством, так как, пользуясь им, руководители вынуждены выносить количественные суждения о так называемых неуловимых целях.

ИЗЛОЖЕНИЕ МЕТОДА ОЦЕНКИ ЦЕЛЕЙ

В данной главе излагается один из методов оценки относительной значимости рассматриваемого множества целей *.

Рассматриваемый метод, очевидно, отличается достаточной универсальностью. Если полученные в рамках метода оценки применимы, то можно исключить проблему «неуловимых» целей и приписать оценки всем целям в одинаковых единицах измерения по общей шкале.

Описываемый здесь метод можно использовать также для определения весовых коэффициентов любых результатов независимо от того, трактуются они как цели или нет, и для определения относительных весов предметов или свойств предметов и/или событий. Примеры использования метода для решения этих дополнительных задач будут приведены в конце главы.

Логические основания метода и применяемые в нем процедуры весьма просты. Чтобы пояснить основную идею метода, предположим, что имеется четыре деревянных стержня неравной длины и отсутствует какое-либо средство измерения длины. Предположим, далее, что нужно определить относительную (а не абсолютную) длину указанных четырех стержней. Один из возможных способов решения такой задачи заключается в следующем. Расположим стержни в определенном порядке от самого длинного до самого короткого. Обозначим самый длинный стержень через A , следующий через B , следующий по порядку через C и, наконец, самый короткий через D . Предположим, что A приписывается значение 100% и производится

* Сведения о других методах, применяемых для решения аналогичной задачи, указаны в примечании в конце главы.

индивидуальная оценка длины стержней B , C и D , выраженной в процентах от длины A . Допустим, что получены следующие результаты: $B = 60\%$, $C = 30\%$, $D = 20\%$. Теперь можно приложить стержни B , C и D друг к другу так, чтобы конец одного совпал с началом другого, и сравнить длину A с длиной полученной суммы ($B + C + D$). Если исходные оценки верны, то длина $B + C + D$ равна 110% длины A . Если же такое сравнение показывает на несоответствие, то нужно скорректировать некоторым образом первоначальные оценки. Далее, сравним длину A с длиной $B + C$, ожидая, что длина ($B + C$) равна 90% длины A . Такое сравнение служит еще одной проверкой первоначальных оценок. Наконец, в данном случае сравнивается длина B с длиной ($C + D$), причем если оценки были верны, то длина B должна составлять 60% ($30 + 20$) или 120% длины ($C + D$).

Эта процедура, по существу, состоит в систематической проверке суждений об отношениях оцениваемых величин на базе последовательных сравнений. Описываемый ниже метод в принципе ничем не отличается от изложенной процедуры. Хотя этот метод, очевидно, характеризуется некоторыми ограничениями в смысле его применимости (о чем свидетельствует допущение аддитивности, которое будет изложено далее в явном виде), он все же может быть использован в достаточно широкой области. В излагаемых ниже примерах получаемые результаты определяются субъективными словесными суждениями, однако это «языковое» ограничение не имеет непосредственного отношения к существу самого метода, который можно применить к оценке фактических выборов или к другим проявлениям предпочтений.

Пожалуй, целесообразно предпослать изложению технических деталей метода пример его практического использования. Предположим, что в задаче фигурирует один эксперт (т. е. задачу решает одно лицо) и что возможны четыре различных результата.

Процедура 1.

1. Упорядочим четыре результата по степени их важности. Пусть O_1 есть результат, считающийся наиболее важным, O_2 — следующий по важности результат, далее идет O_3 и, наконец, O_4 .

2. Присвоим значение 1,00 наиболее важному результату и некоторые другие значения остальным результатам, отражающие, как представляется, степень их важности. Так, например, эксперт может приписать значения (оценки) 1,00; 0,80; 0,50 и 0,30 результатам O_1 , O_2 , O_3 и O_4 соответственно. Обозначим эти величины символами v_1 , v_2 , v_3 и v_4 соответственно. Их следует рассматривать как первые оценки «истинных» значений V_1 , V_2 , V_3 и V_4 .

3. Далее проведем сравнение O_1 с (O_2 и O_3 и O_4), т. е. выясним, что выберет эксперт, если ему предоставить возможность «получить» результат O_1 или сумму результатов O_2 , O_3 и O_4 . Предположим, что он утверждает, что O_1 является предпочтительным. Тогда значение оценки v_1 следует изменить так, чтобы выполнялось неравенство

$$v_1 > v_2 + v_3 + v_4.$$

Например, можно принять $v_1 = 2,00$, $v_2 = 0,80$, $v_3 = 0,50$ и $v_4 = 0,30$. Отметим, что первоначальные значения оценок O_2 , O_3 и O_4 остались без изменений.

4. Сравним теперь O_2 с O_3 и O_4 . Предположим, что суммарный результат O_3 и O_4 более предпочтителен. Тогда требуется дальнейшее изменение первоначальных оценок. Например, можно принять

$$v_1 = 2,00; \quad v_2 = 0,70; \quad v_3 = 0,50 \quad \text{и} \quad v_4 = 0,30.$$

В итоге все эти значения не противоречат мнениям эксперта.

5. В таком случае определение оценок закончено. Однако может оказаться целесообразным «нормировать» полученные оценки, разделив каждую из них на величину $\sum v_j$, которая в данном случае равна 3,50. Эти нормированные значения оценок обозначим символами v'_j . Имеем:

$$v'_1 = 2,00/3,50 = 0,57$$

$$v'_2 = 0,70/3,50 = 0,20$$

$$v'_3 = 0,50/3,50 = 0,14$$

$$v'_4 = 0,30/3,50 = 0,09$$

$$\text{Итого } 1,00$$

Д о п у щ е н и я. Прежде чем перейти к формальному изложению метода, проиллюстрированного на изложенном выше примере, пожалуй, стоит указать на некоторые наиболее важные допущения, лежащие в его основе. Прежде всего остановимся на формальных допущениях.

А-1. Каждому результату O_j соответствует действительное неотрицательное число V_j , рассматриваемое как оценка истинной значимости O_j .

А-2. Если результат O_j более важен, чем O_k , то $V_j > V_k$, и если O_j равноценен O_k , то $V_j = V_k$.

А-3. Если оценки V_j и V_k соответствуют результатам O_j и O_k , то оценка $V_j + V_k$ соответствует общему результату O_j и O_k .

Однако в тех случаях, когда результат O_1 логически обуславливает отсутствие результата O_2 , допущение А-3, очевидно, не выполняется. При таких условиях общий результат O_1 -и- O_2 невозможен, а следовательно, нельзя пользоваться оценкой $V_1 + V_2$. Предположим, что O_1 и O_2 определены на некоторой шкале измерений, скажем на шкале годового дохода. Пусть O_1 есть годовой доход точно в 20 тыс. долларов, а O_2 — годовой доход точно в 10 тыс. долларов. Но результат O_1 -и- O_2 невозможен, и его суммарная оценка не имеет никакого смысла. Кроме того, если наличие O_1 влечет O_2 , то O_1 -и- O_2 сводится к O_1 , а этот результат в общем случае не имеет оценки $V_1 + V_2$. Доход не менее 20 тыс. долларов в год означает, что доход в 10 тыс. долларов уже наверняка получен. В этом случае O_1 -и- O_2 сводится к годовому доходу не менее 20 тыс. долларов и очевидно, что оценка не будет иметь значения $V_1 + V_2$. Таким образом, по существу описываемый метод применим только тогда, когда результаты дискретны, непротиворечивы и взаимно независимы.

А-3 есть важнейшее допущение аддитивности, принятое в рассматриваемом методе. Следствия этого допущения таковы:

А-3а. Если результат O_j предпочтительней O_k , а O_k предпочтительней O_l , то общий результат O_j -и- O_k предпочтительней O_l . Это условие не выполняется, если первые две цели (результаты) заключаются, скажем, в следующем: «поесть крабов на обед сегодня», «поесть бифштекс на обед сегодня», а третья — «поесть рыбы на обед сегодня». Обеду, состоящему из крабов и мяса, может вовсе не отдаваться предпочтения по сравнению с обедом, на который подают только рыбу. Однако соответствующим уточнением целей часто можно добиться устранения отмеченного затруднения. Так, в рассмотренном примере достаточно снять ограничение, наложенное условием «сегодня», чтобы преодолеть эту трудность.

А-3б. «Значимость» общего результата O_j -и- O_k эквивалентна значимости общего результата O_k -и- O_j . Это означает, что порядок представления результатов или их группировка не влияют на предпочтения. Например,

предполагается, что человек не делает различия между комбинированной целью «репутация-и-богатство» и целью «богатство-и-репутация».

А-Зв. Если общий результат O_j -и- O_k эквивалентен O_k , то $V_j = 0^*$. Рассматриваемый метод основан также на ряде операционных допущений.

1. Если задан некоторый интервал значений действительного переменного, скажем от 0 до 1, то эксперт может определить первую оценку значимости (ценности) каждого результата по такой шкале, причем каждая оценка содержит некоторую информацию о V_j (как указывалось выше, оценки V_j обозначаются символами v_j).

2. Можно утверждать, что рассматриваемый метод обеспечивает последовательное улучшение оценок V_j . Уже говорилось, что эксперту предлагается два теста, каждый из которых дает информацию о том, насколько важными с его точки зрения являются результаты. В ходе первого теста эксперт назначает *предварительные* оценки величин V_j по шкале, предлагаемой ему операционистом-исследователем. Далее перед экспертом ставят определенные вопросы, относящиеся к различным комбинациям результатов, и высказываемые им предпочтения дают дополнительную информацию о величинах V_j . Так, например, предполагается, что ответы на предложенные вопросы не определяются полностью исходным мнением эксперта при назначении предварительных оценок. Иными словами, если первоначально эксперт определил оценки, допустим, так: $v_1 = 0,7$, $v_2 = 0,5$ и $v_3 = 0,4$, то это мнение не обязательно означает, что во втором тесте он будет утверждать, что результат O_1 менее предпочтителен, чем общий результат O_2 -и- O_3 . Второй набор мнений содержит по крайней мере потенциально информацию, на основании которой можно внести коррекцию в первый набор. Это допущение отчасти практически подтверждается экспериментальными данными, полученными при использовании описываемого метода.

НАДЕЖНОСТЬ И СМЕЩЕНИЕ ОЦЕНОК

Можно определить меру *надежности* оценок. Предварительные исследования показывают, что в контролируемых условиях возможно приближенно установить степень повторяемости оценок. Однако рассматриваемый метод не обеспечивает какой-либо оценки *точности* или *смещения* высказываемых мнений. Этот серьезный недостаток присущ всем известным методам оценки мер предпочтения. Пока что, по существу, отсутствует точное, общепринятое определение смысла понятия «истинного» предпочтения, так что смещение в некотором роде неизмеряемо. В результате практического использования метода в конечном счете, возможно, удастся найти подход к оценке смещения.

Следует добавить, что оценку смещения можно было бы получить, если бы содержание V_j удалось выразить через определенные свойства (вероятности выбора) действительных выборов, осуществляемых в контролируемых условиях. В этом направлении предприняты некоторые попытки [3], но следует еще многое сделать, чтобы разработать процедуры, которые можно было бы использовать в научных исследованиях.

* Отметим, что если существует результат O_j , удовлетворяющий условию А-Зв, то он имеет нулевую оценку при любых методах количественного измерения, рассматриваемых в данной работе, т. е. существует нулевая точка определенной шкалы, инвариантная к любым преобразованиям шкалы оценок V . Это условие не выполняется при оценках (измерениях) так называемой «полезности», рассматриваемых в [11] и [14].

ФОРМАЛЬНОЕ ИЗЛОЖЕНИЕ МЕТОДА

Общее формальное описание метода оценки величин V_j весьма громоздко по внешнему виду (хотя и несложно на практике).

1. Упорядочим результаты в соответствии с их значимостью (ценностью). Пусть O_1 представляет наиболее важный результат, O_2 — следующий по степеням важности и т. д., а O_m — наименее важный.

2. Припишем значение 1,00 результату O_1 (т. е. $v_1 = 1,00$) и другие подходящие значения всем остальным результатам.

3. Сравним O_1 с $O_2 + O_3 + \dots + O_m$ *.

3.1. Если O_1 предпочтительнее $O_2 + O_3 + \dots + O_m$, изменим (в случае необходимости) значение v_1 так, чтобы выполнялось неравенство $v_1 > v_2 + v_3 + \dots + v_m$. При этой корректировке, так же как и при всех остальных, следует стремиться к тому, чтобы относительные значения корректируемого набора (v_2, v_3 и т. д.) оставались без изменений. Далее следует перейти к шагу 4.

3.2. Если O_1 и $O_2 + O_3 + \dots + O_m$ равноценны, то изменим (в случае необходимости) значение v_1 так, чтобы выполнялось равенство $v_1 = v_2 + v_3 + \dots + v_m$. Далее следует перейти к шагу 4.

3.3. Если результат O_1 менее предпочтителен, чем $O_2 + O_3 + \dots + O_m$, то изменим (в случае необходимости) значение v_1 так, чтобы выполнялось неравенство $v_1 < v_2 + v_3 + \dots + v_m$.

3.3.1. Сравним O_1 с $O_2 + O_3 + \dots + O_{m-1}$.

(3.3.1.1.) Если O_1 предпочтительнее, изменим (в случае необходимости) значения оценок так, чтобы выполнялось неравенство $v_1 > v_2 + v_3 + \dots + v_{m-1}$. Далее следует перейти к шагу 4.

(3.3.1.2) Если результат O_1 равноценен всем остальным вместе взятым, изменим (в случае необходимости) значения оценок так, чтобы выполнялось равенство $v_1 = v_2 + v_3 + \dots + v_{m-1}$. Далее следует перейти к шагу 4.

(3.3.1.3) Если O_1 менее предпочтителен, чем все остальные результаты вместе взятые, изменим (в случае необходимости) значения оценок так, чтобы выполнялось неравенство $v_1 < v_2 + v_3 + \dots + v_{m-1}$.

(3.3.1.3.1) Сравниваем O_1 с $O_2 + O_3 + \dots + O_{m-2}$ и т. д. до тех пор, пока или O_1 предпочтительнее, или равноценен всем остальным результатам вместе взятым, а затем переходим к шагу 4, либо до тех пор, пока не закончено сравнение O_1 с $O_2 + O_3$, и затем переходим к шагу 4.

4. Сравним O_2 с $O_3 + O_4 + \dots + O_m$ и выполним далее весь шаг 3.

5. Продолжим шаг 4, пока не выполнено сравнение O_{m-2} с $O_{m-1} + O_m$.

6. Преобразуем каждое полученное значение v_j в нормированное v'_j , разделив соответствующие величины на Σv_j . В итоге сумма $\Sigma v'_j$ должна быть равна 1,00.

Отметим, что окончательные значения оценок являются относительными, т. е. исключение или добавление какого-либо результата может повлиять на полученные значения. Кроме того, полученные оценки для набора результатов могут изменяться во времени, если истинные значения также меняются **.

Описанный метод малоэффективен, если число результатов равно семи или еще больше. В таких случаях метод становится слишком громоздким.

* Знак «+» обозначает здесь логическую конъюнкцию, логическую операцию «и».

** При соответствующем выборе интервала назначения оценок шкалу v_j можно преобразовать с помощью любой линейной функции, проходящей через начало координат (нуль).

Ниже описывается более подходящая процедура. Но, так же как и ранее, предполагается, что оценки потенциально улучшаются. В частности, может измениться исходное упорядочение результатов подобно тому, как эта возможность предусматривалась в изложенном методе.

Процедура 2.

1. Упорядочим все множество результатов в соответствии с предпочтениями, не ставя им в соответствие числовых значений.

2. Выберем случайным образом один результат из множества. Пусть O_s представляет такой (эталонный) результат. Затем также случайным образом разобьем оставшийся набор результатов на группы, содержащие не более пяти результатов в каждой, и желательно (хотя это и не обязательно) на группы примерно одинакового размера. Каждый результат, кроме O_s , должен быть включен только в одну группу. (Возможна альтернатива, когда результату O_s приписывается наивысший ранг.)

3. Добавим O_s в каждую группу и припишем ему оценку 1,00 ($v_s = 1,00$).

4. Используем шаги 1—5 описанного ранее метода, чтобы получить ненормированные значения оценок результатов в группах, образованных на шаге 3 данного метода, но, корректируя значения v_j , оставим v_s без изменений.

5. Сравним оценки, полученные на шагах 2—4 процедуры (метода) 2, с предпочтениями, упорядоченными на шаге 1. Если окажется, что упорядочения не совпадают, то изменим одно из них и в случае необходимости повторим шаги 2—4 данной процедуры.

6. Когда в итоге получим непротиворечивые результаты, пронормируем значения оценок, полученные на шаге 5 данной процедуры, разделив оценку, поставленную в соответствие каждому результату, на сумму оценок всех результатов.

Описанную процедуру можно проиллюстрировать следующим примером. Допустим, что имеется десять целей (результатов).

1. Предположим, что они упорядочены в соответствии со значимостью следующим образом: $O_1, O_2, \dots, O_{10}$.

2. Пусть случайным образом выбрана цель O_7 .

3. Остальные цели можно также случайно разбить на три группы:

(a)	(b)	(c)
O_6	O_5	O_1
O_{10}	O_9	O_3
O_2	O_4	O_8

4. Добавим O_7 к каждой группе, приписав этой цели оценку 1,00.

5. Предположим, что, применяя далее указанные шаги процедуры 1, мы получили следующие ненормированные значения оценок:

(a)	(b)	(c)
$v_6 = 1,35$	$v_5 = 1,50$	$v_1 = 3,60$
$v_{10} = 0,60$	$v_9 = 0,75$	$v_3 = 3,00$
$v_2 = 2,70$	$v_4 = 1,80$	$v_8 = 0,90$
$v_7 = 1,00$	$v_7 = 1,00$	$v_7 = 1,00$

6. Сравнение с шагом 1 показывает, что O_2 и O_3 поменялись местами. Если исходный порядок признается правильным, то значения оценок O_2 и/или O_3 следует в соответствующих группах изменить. Далее все шаги выполняются в прежней последовательности. Однако может оказаться,

что вычисленные значения оценок, в отличие от исходных, признаны более верными. Тогда полученные на шаге 5 значения нормируются (путем деления каждой оценки на число 17,2), и в итоге получаем следующее:

$$\begin{aligned} v'_1 &= 0,21, & v'_2 &= 0,16, & v'_3 &= 0,17, & v'_4 &= 0,10, & v'_5 &= 0,09, \\ v'_6 &= 0,08, & v'_7 &= 0,06, & v'_8 &= 0,06, & v'_9 &= 0,04, & v'_{10} &= 0,03. \end{aligned}$$

В приведенных примерах принято допущение, что оценки определяет одно лицо. Однако довольно часто желательно, чтобы оценки назначались коллективно, особенно тогда, когда принимается коллегиальное решение. В таких случаях каждое сравнение проводится на основе голосования, причем решение принимается простым большинством голосов. Эта процедура на практике вовсе не столь громоздка, как кажется с первого взгляда.

В других случаях может оказаться целесообразным, чтобы каждый член группы экспертов назначал оценки независимо. При таких условиях оценку каждого результата можно брать как среднее от оценок, поставленных данному результату всеми членами группы.

Примеры.

Ниже довольно подробно описываются два реальных примера использования изложенного в данной главе метода. Первый пример не относится к операционным исследованиям, но он содержит ряд интересных подходов к использованию метода определения весовых коэффициентов, не имеющих непосредственного отношения к постановке задачи.

Пример 1*. В самом общем виде дефект можно определить как нежелательное свойство, вызывающее понижение потребительной ценности изделия или продукции. Степень такого понижения, обусловленную каждым из рассматриваемых дефектов, невозможно количественно измерить. Поэтому ее определяют субъективно, исходя из опыта и производственной практики. К сожалению, существует столько же различных мнений относительно значения конкретных дефектов, сколько людей их высказывает. В такой ситуации вполне возможно, что одна и та же партия готовой продукции будет признана либо годной, либо забракована исходя из используемых критериев оценки, а следовательно, субъективных мнений, на которых основаны эти критерии.

Существует насущная необходимость разработки метода, позволяющего определять количественные оценки различной степени «дефектности» продукции. Общепринятые процедуры выборок можно легко применить для контроля качества продукции. На основании этих процедур можно получить обоснованные выводы относительно годности продукции. Ниже излагается одна из таких процедур и пример ее использования на практике.

Метод упорядочения. Рассматриваемый метод представляет собой модификацию количественного метода упорядочения (описанного выше) и служит для построения логической основы принятия стратегических решений. Процедура учитывает субъективные мнения с целью определения количественных соотношений между переменными в случаях, когда эти соотношения нельзя установить из теоретических соображений или на основании накопленных статистических данных. Учитываемые мнения высказываются индивидуально или коллективно и отражают предпочтения, отдаваемые каждому переменному, в сравнении со значимостью различных комбинаций переменных, принадлежащих той же системе. Окончательное решение принимается после ряда переоценок каждого перемен-

* Пример взят из статьи Stills P., A Method for Defect Evaluation, Industrial Quality Control, XI, no. 1, 9—12 (July 1954).

ного и относительного упорядочения их по произвольной шкале, не противоречащей каждому из решений, принятых в процессе оценки.

Первым шагом процедуры является предварительный перечень переменных в порядке их значимости (важности). Переменному, стоящему в начале перечня, приписывается значение 100, а остальным переменным ставятся в соответствие числовые оценки, определяющие их значимость по отношению к самому важному переменному. Эти величины (оценки) качественны по своему содержанию и служат только как промежуточная основа упорядочения, которое в дальнейшем подвергается корректировке.

Эксперту или группе экспертов затем предлагается зафиксировать свое мнение об относительной значимости первого переменного в сравнении с общей значимостью второго и третьего переменных перечня. Три возможных ответа на этот вопрос сводятся к следующему: (1) первое переменное играет более важную роль, чем второе и третье вместе взятые; (2) это переменное менее важное; (3) они равноценны. В случае первого ответа значение оценки 100, приписанной первому переменному, должно превосходить сумму второго и третьего переменных, фигурирующих в данном сравнении. Наоборот, в случае второго ответа эта сумма должна быть больше 100. Наконец, естественно, указанные величины должны быть равны в случае третьего ответа. Эта процедура повторяется, пока первое переменное не сравнивается со всеми комбинациями по два переменных, входящих в перечень. Каждое принятое решение фиксируется наряду с изменениями, вносимыми с целью согласования текущего и предшествующих решений.

Аналогичным образом второе переменное сравнивается с комбинациями переменных, стоящих ниже его в исходном перечне. Процесс сравнений заканчивается тогда, когда дана оценка каждому переменному и значения оценок скорректированы в соответствии с отдельными решениями.

Вследствие неоднократных корректировок исходных числовых оценок вполне возможно, что в итоге первоначальное упорядочение существенно изменится. В таком случае может возникнуть необходимость пересмотра всего перечня переменных и повторения всей процедуры сравнений, хотя эта необходимость становится очевидной уже в самом начале процедуры.

П р и м е н е н и е м е т о д а . Контроль качества готовых упакованных лекарств в фармацевтической промышленности представляет собой типичную операцию, к которой применялся описанный метод. В процессе контроля качества существенную роль играют как дефекты, относящиеся к свойствам продукции, так и дефекты иного рода, причем оба класса дефектов нужно обязательно учитывать в смысле выявления их влияния на качество. Установлено, что все виды дефектов имеют различное значение при принятии решения забраковать партию продукции, и вследствие этого им необходимо приписать соответствующие веса. Отбор учитываемых дефектов производится на основании мнения квалифицированной группы сотрудников отдела контроля качества. В описываемом частном случае каждый дефект рассматривался как свойство продукции и имел место либо отсутствовал в каждом отдельном флаконе. Единственным затруднением при решении вопроса об учете каждого дефекта является необходимость такого его определения, чтобы наличие данного дефекта в большинстве экземпляров партии служило единственным основанием считать всю партию негодной (бракованной).

Применявшиеся ранее нормативы, на основании которых отдельные партии признавались годными или негодными, базировались на классификации дефектов на две группы: основные и незначительные. Была построена произвольная шкала, по которой десять незначительных дефектов незави-

симо от их конкретного характера приравнивались каждому из основных. Такой метод оценки общепринят в производственной практике. Однако наблюдалось резкое расхождение мнений среди специалистов по контролю качества относительно такого метода оценки, а также дифференциации двух видов дефектов. Поэтому было признано целесообразным установить такой порядок определения годности продукции, который был бы приемлем для всех членов экспертной группы и в соответствии с которым можно было бы упорядочить все дефекты по степени их значимости, поставив им в соответствие числовые оценки.

В описываемом примере группа экспертов состояла из девяти человек, мнения которых относительно важности рассматриваемых дефектов существенно расходились. Коллективное мнение группы учитывалось по правилу большинства голосов для каждого сравнения значимости дефектов, причем это коллективное решение принималось за «единую редакцию».

Дефекты, которым нужно было дать оценки, были определены специально таким образом, чтобы все девять членов экспертной группы могли произвести одинаковые сравнения. Решения принимались открытым голосованием, однако обсуждение индивидуальных мнений было запрещено. Коррекция оценок в соответствии с каждым принятым решением осуществлялась с ведома экспертной группы.

Полученные результаты. Все члены экспертной группы должны были произвести исходное упорядочение всех дефектов и поставить им в соответствие числовые оценки. Первому по порядку дефекту было приписано значение 100, а остальные были расположены в порядке убывания их значимости, причем им поставили в соответствие числовые оценки в долях оценки наиболее важного дефекта. Согласно высказанным мнениям были вычислены средние значения оценок каждого дефекта, и эти значения были приняты в качестве исходных данных для последующей коррекции. В общем случае можно обозначить рассматривавшиеся дефекты буквами от *A* до *F*. Исходные оценки имели следующие значения:

Переменное <i>A</i>	100	Переменное <i>D</i>	44
---------------------	-----	---------------------	----

Переменное <i>B</i>	60	Переменное <i>E</i>	34
---------------------	----	---------------------	----

Переменное <i>C</i>	55	Переменное <i>F</i>	27
---------------------	----	---------------------	----

Далее членам группы было предложено в описанном выше порядке сделать ряд сравнений. В соответствии с каждым сравнением (решением) числовые значения оценок в случае необходимости изменялись (корректировались), пока не были исчерпаны все возможные комбинации сравнений. Часть перечня сравнений и принятых решений приведена в таблице 6.1.

Следует отметить, что первое и второе решения группы не требовали внесения каких-либо поправок в исходные числовые оценки, ибо сумма оценок второго и третьего переменных, а также второго и четвертого превышает значение оценки первого (ведущего) переменного. Голосование по третьему сравнению, однако, показало, что общее мнение склоняется в пользу *B* и *E* по отношению к *A*, хотя исходная оценка была 94 к 100 в пользу ведущего переменного. В связи с этим было внесено необходимое изменение, обеспечившее согласование числовых оценок с решением большинства. В данном случае значение оценки переменного *B* было произвольно увеличено с 60 до 70. При этом не следует обращать особого внимания на выбор переменного, для которого корректируется оценка, или на точное значение приращения оценки, так как сама коррекция в свою

Таблица 6.1

Сравнения различных комбинаций дефектов

Номер решения	Сравнение	Решение			Решение по большинству голосов	Номер решения	Сравнение	Решение			Решение по большинству голосов
		да	нет	равнозначно				да	нет	равнозначно	
1	$A > (B + C)$	2	7	0	нет	11	$B > (C + E)$	6	1	2	да
2	$A > (B + D)$	2	6	1	нет	12	$B > (C + F)$	8	0	1	да
3	$A > (B + E)$	3	5	1	нет	13	$B > (D + E)$	9	0	0	да
4	$A > (B + F)$	6	1	2	да	14	$B > (D + F)$	9	0	0	да
5	$A > (C + D)$	7	2	0	да	15	$B > (C + D + E)$	2	5	2	нет
6	$A > (C + E)$	9	0	0	да	16	$C > (D + E)$	2	7	0	нет
7	$A > (C + F)$	9	0	0	да	17	$C > (D + F)$	1	8	0	нет
8	$A > (B + C + D)$	1	8	0	нет	18	$C > (E + F)$	5	2	2	да
9	$A > (B + C + E)$	1	8	0	нет	19	$C > (D + E + F)$	0	7	2	нет
10	$B > (C + D)$	4	2	3	равнозначно	20	$D > (E + F)$	1	4	4	равнозначно

очередь является промежуточной и может подвергнуться изменению в ходе дальнейших сравнений.

Приведенное в таблице четвертое решение было принято исходя из измененной оценки B , равной 70, и не потребовало, как и в двух первых случаях, внесения каких-либо корректив в таблицу оценок дефектов. Следующее, пятое решение не противоречило исходным оценкам, однако оно показало, что существует явно выраженная тенденция пересмотра относительной значимости рассматриваемых переменных. Сумма исходных оценок C и D составляла 99, т. е. всего на единицу меньше оценки ведущего переменного, принятой за 100. Члены экспертной группы, как показали результаты голосования (7 к 2), высказались решительно в пользу A по сравнению с $(C + D)$, т. е. посчитали, что вес A гораздо существеннее. Поэтому в ходе дальнейшей корректировки оценки переменных C и D должны были иметь тенденцию к уменьшению, если бы вообще потребовалось изменение оценок.

После вынесения каждого решения процедура коррекции оценок повторялась. На каждом таком шаге полученная таблица оценок являлась основой для последующей коррекции. Особо следует остановиться на решениях 10 и 20, ибо они относятся к равноценным отношениям между переменными. В обоих случаях внесенные коррективы не противоречили всем предыдущим решениям, а также и данному решению.

Окончательные оценки, полученные на основе описанной процедуры, имели следующие значения:

Переменное A	100	Переменное D	38
Переменное B	82	Переменное E	24
Переменное C	43	Переменное F	13

На этой стадии операционист должен проанализировать все сравнения и проверить решения большинства с окончательными оценками на непротиворечивость. Как видно, решения 10 и 20, т. е. сравнения B с $(C + D)$ и D

с $(E + F)$, весьма близки к эквивалентности и в то же время не противоречат остальным решениям.

Анализ. После получения указанных выше результатов можно применить метод статистических выборок для оценки качества продукции, предположив, что ведущее переменное является наиболее важным дефектом, и приписывая соответствующие оценки всем остальным переменным. Следовательно, переменное B составляет 0,82 основного дефекта, переменное C — 0,43 и т. д. Заметим, что указанное допущение было принято решением большинства еще до того, как оно было введено в процедуру контроля качества.

Если обратиться к единичным процедурам выборочного контроля, описанным в стандартах военного министерства США, то в них указано, что для партии в 40 тыс. изделий при допустимом уровне качества 0,65% в выборке из 450 изделий допускается не более семи дефектов. При выходе за эту норму вся партия признается негодной. Однако максимальное допустимое число дефектов было принято равным 8,000, чтобы исключить область неопределенности между 7,000 и 8,000. Таким образом, 8,000 является максимальным допустимым уровнем для процедуры контроля качества на основании окончательных оценок дефектов. Во время проверки обнаружение в выборке каждого дефекта любого типа фиксируется и общая частота появления дефекта умножается на числовую оценку данного дефекта. Полученные произведения складываются, и общая сумма не должна превышать 8,000, чтобы вся партия была признана годной. Если же эта сумма превысит 8,000, то партия бракуется и возвращается на переработку.

Совершенно очевидно, что анализ карты контроля качества позволяет обнаружить причину брака и установить, какой вид исправлений требуется, чтобы партия была признана годной. Переработке подвергается вся партия, производится 100%-ный контроль качества, а затем берется вторая случайная выборка из 450 изделий для оценки качества.

Примером описанной ситуации может служить недоброкачественная закупорка флаконов при упаковке готовой продукции. Число флаконов в выборке с перекошенными пробками было настолько велико, что этот дефект явился основной причиной брака всей партии. Эта партия была направлена на конвейер, и целиком подвергнута контролю закупорки флаконов. Затем партия была вновь полностью упакована, и была взята вторая случайная выборка для определения качества продукции. Выборка проверялась по всем видам дефектов, был подсчитан общий показатель качества, и партию признали годной.

Пожалуй, наиболее существенным достоинством описанного метода является установление стандартной процедуры контроля качества в том отношении, где прежде существовал значительный произвол. Установление количественных оценок годности и непригодности продукции позволило перевести процедуру контроля качества из сферы субъективных мнений в сферу объективности и обеспечить независимость этой процедуры от индивидуальных мнений по поводу относительной важности каждого из рассматриваемых дефектов. Кроме того, при определении оценок дефектов экспертная группа могла проанализировать каждый дефект в сравнении с различными комбинациями других дефектов и определить более достоверно свои собственные коллективные предпочтения.

Наконец, следует отметить, что каждый член группы был вполне удовлетворен окончательными оценками и считал, что полученные значения оценок совпадали с его собственным априорным мнением относительно важности рассматривавшихся дефектов.

Пример II. Правление фирмы поставило задачу оценить планы различных мероприятий, которые следовало бы провести в деятельности фирмы в течение пятилетнего периода. Исследовательская (операционная) группа на основании предварительного изучения стратегических решений правления в прошлом составила перечень предполагаемых целей руководства фирмы на последующие пять лет. Группа организовала совещание с членами правления, чтобы обсудить содержание этих целей и соответственно изменить их определение в случае неясной формулировки или неполноты. Цели были определены следующим образом:

О₁. Сохранение существующего уровня руководства деятельностью фирмы.

О₂. Обеспечение гарантированной 6%-ной прибыли вкладчикам на их исходные капиталовложения.

О₃. Фирма должна обеспечивать 15% прибыли на капиталовложения, если спрос на ее продукцию сохранится в интервале от 100% до 200% от текущего.

О₄. Исключить случаи увольнения, а также значительного продвижения по службе ведущих сотрудников фирмы.

О₅. Обеспечить устойчивость рабочей силы (например, исключить угрозы забастовок, свести к минимуму наем и увольнение).

О₆. Обеспечить ведущее положение фирмы в смысле технических показателей продукции.

О₇. Предоставление услуг, выходящих за рамки деятельности фирмы, оговоренной в ее правовом статуте (выполнение условий правового статута обязательно для деятельности фирмы и в этом смысле не может рассматриваться как «цель»).

Каждая цель была подвергнута обсуждению, и было достигнуто соглашение относительно содержания, вкладываемого в определение цели. Операционная группа пришла к выводу, что допущение аддитивности вполне оправдано применительно к данному перечню целей.

Далее каждый член правления индивидуально применил описанную выше процедуру 2 оценки целей, и в результате были получены нормированные значения оценок каждой цели каждым членом правления. Затем правление обсудило полученные результаты, и были представлены окончательные оценки (весовые коэффициенты) целей. В описываемом примере голоса всех членов правления были равноправны, а значения v_j по каждой цели усреднялись. (Очевидно, что такая процедура голосования по оценкам целей является лишь одной из большого числа возможных процедур. Ее можно, пожалуй, считать наиболее «демократичной», ибо допускается обсуждение, в результате которого могут изменяться индивидуальные мнения и все участники голосования равноправны.)

Окончательные результаты имели следующий вид:

О₁: стабильность руководства 0,25

О₂: финансовая стабильность 0,30

О₃: финансовые перспективы 0,10

О₄: ведущий персонал 0,15

О₅: стабильность рабочей силы 0,05

О₆: технический прогресс 0,05

О₇: предоставление услуг 0,10

Правление фирмы рассматривало три возможные стратегии:

Стратегия А. Расширение деятельности фирмы на 200% в течение двух лет, включая выпуск новой продукции и завоевание новых рынков сбыта.

Стратегия В. Сохранение прежнего размера фирмы при направлении основных усилий на улучшение выпускаемых моделей продукции.

Стратегия С. Сохранение прежнего размера фирмы при направлении основных усилий на замену недостаточно прибыльной продукции новой продукцией.

Каждая стратегия была описана достаточно подробно. Приведенные формулировки представляют собой лишь сокращенное описание. Затем был сформирован специальный комитет, включавший членов правления, технологов, специалистов по производству и сбыту, а также экономических советников. В задачу этого комитета входила оценка каждой стратегии по отношению к достижению каждой из указанных семи целей. На основании коллективного мнения определялась некоторая оценка в интервале от нуля до единицы, где «0» означал, что данная стратегия серьезно угрожает достижению данной цели (вероятность достижения цели при этой стратегии весьма мала), «1» означала, что данная стратегия, по существу, гарантирует достижение данной цели (вероятность очень велика), а оценка «0,5» свидетельствует о том, что стратегия не влияет на достижение цели. В итоге была построена таблица 6.2.

Таблица 6.2

Эффективность стратегий по отношению к целям

Стратегия	Цель						
	О ₁	О ₂	О ₃	О ₄	О ₅	О ₆	О ₇
А	0,4	0,2	0,8	0,8	0,3	0,6	0,8
В	0,9	0,9	0,2	0,3	0,8	0,4	0,3
С	0,7	0,7	0,4	0,3	0,7	0,8	0,5

Было принято допущение, что стратегия, дающая наибольшее значение «математического ожидания», является оптимальной с точки зрения интересов фирмы. Эта величина для стратегии А была подсчитана по всем целям путем умножения оценки цели на эффективность стратегии и суммирования полученных произведений по всем целям.

$$\text{Полезность стратегии А} = (0,4 \cdot 0,25) + (0,20 \cdot 0,30) + (0,8 \cdot 0,1) + (0,8 \cdot 0,15) + (0,3 \cdot 0,05) + (0,6 \cdot 0,05) + (0,8 \cdot 0,10) = 0,485$$

$$\text{Полезность стратегии В} = 0,650$$

$$\text{Полезность стратегии С} = 0,595$$

Таким образом, было признано, что оптимальной является стратегия В, так как она имеет наибольшее значение «математического ожидания».

Примечание. За последнее десятилетие проделана большая работа в области создания формальной (аксиоматической) теории ценности (полезности). Исходя из идей Парето [11], фон Нейман и Моргенштерн [10] сформулировали набор формальных условий, при выполнении которых обеспечивается введение меры полезности. Затем по их стопам пошли Давидсон, Маккинси и Саппес [5], Давидсон и Саппес [7], Саппес и Вайнет [14] и ряд других ученых, которые также определили различные меры

полезности, предложив для этого различные аксиоматические системы. Большинство из этих работ объединено в сборнике [15]. Этот сборник и работа Саппеса и Вайнета [14] содержат обширную библиографию по проблеме измерения полезности.

Задача построения меры полезности социальной группы по системам ценностей, принятым отдельными личностями, названа задачей объединения. Эту задачу подробно исследовал Эрроу [2]. Кроме того, она рассмотрена в работах Гудмэна, Кумбса, а также Буша, Мостеллера и Томпсона, помещенных в сборнике [15].

Некоторые экспериментальные исследования, выполненные экономистами и психологами, освещены в статье Эдвардса [8]. Сложная задача на данной стадии развития теории связана с трудностью выделения субъективных оценок вероятности и предпочтения. Эксперименты Мостеллера и Ноджи [9] можно рассматривать как попытки измерения полезности в предположении о полном соответствии субъективных и объективных вероятностей. Примерно аналогичные методы изложены в работах Давидсона, Зигеля и Саппеса [6], Зигеля [12] и ряда других авторов сборника [15].

ЛИТЕРАТУРА

1. Ackoff R. L., On a Science of Ethics, *Phil. Phenom. Res.*, IX, no. 4, 663—672 (June 1949).
2. Arrow K. J., *Social Choice and Individual Values*, John Wiley & Sons, New York, 1951.
3. Churchman C. W. and Ackoff R. L., An Experimental Definition of Personality, *Phil. Sci.*, 14, no. 4, 304—332 (Oct. 1947).
4. Churchman C. W. and Ackoff R. L., *Methods of Inquiry*, Educational Publishers, St. Louis, 1950.
5. Davidson D., McKinsey J. C. C. and Suppes P., Outlines of a Formal Theory of Value, I, Report No. 1, Stanford Value Theory Project, Feb. 1954.
6. Davidson D., Siegel S. and Suppes P., Some Experiments and Related Theory on the Measurement of Utility and Subjective Probability, Report No. 4, Stanford Value Theory Project, May 1955.
7. Davidson D. and Suppes P., Finitistic Rational Choice Structures, Report No. 3, Stanford Value Theory Project, Feb. 1955.
8. Edwards W., The Theory of Decision Making, *Psych. Bull.*, 51, no. 4, 380—417 (July 1954).
9. Mosteller F. and Noguee P., An Experimental Measure of Utility, *J. Polit. Econ.*, 59, no. 5, 371—404 (Oct. 1951).
10. Neumann J. von and Morgenstern O., *Theory of Games and Economic Behavior*, 2nd ed., Princeton University Press, Princeton, 1947.
11. Pareto V., *Manuel d'Économie politique*, 2nd ed., M. Giard, Paris, 1927.
12. Siegel S., A Behavioristic Method of Obtaining a Higher Ordered Metric Scale of Utility, Third Annual Meeting, Operations Research Society of America, New York, June 4, 1955.
13. Smith N. M., Jr., Comments, *J. Opns. Res. Soc. Am.*, 2, no. 2, 181—187 (May 1954).
14. Suppes P. and Winet M., An Axiomatization of Utility Based on the Notion of Utility Differences, *Mgmt. Sci.*, I, no. 3—4, 259—270 (Apr.—July 1955).
15. Thrall R. M., Coombs C. H. and Davis R. L. (eds.), *Decision Processes*, John Wiley & Sons, New York, 1954.

В этом разделе дается понятие модели: что такое модель, какие типы моделей существуют, как строить модель и как использовать ее для решения задачи.

Показывается, что модель является отображением исследуемой системы, с помощью которого можно предсказывать, как повлияют на эффективность системы возможные изменения ее характеристик. Рассматриваются три вида моделей: физические модели (модели геометрического подобия), модели прямой аналогии и символические (математические модели), причем наиболее важной является символическая модель. Выполнив определенные математические или логические операции, можно приспособить символическую модель для решения изучаемой задачи. Различаются два вида математических методов решения задачи: аналитический и численный. В некоторых символических моделях есть параметры или неизвестные, которые нельзя точно оценить. В этих случаях применяют метод Монте-Карло. Все три метода (аналитический, численный и метод Монте-Карло) рассмотрены на соответствующих примерах. В конце главы кратко описываются модели некоторых наиболее распространенных систем. В других частях книги эти модели рассматриваются более подробно.

Глава 7

ПОСТРОЕНИЕ И РЕШЕНИЕ МОДЕЛИ

ВВЕДЕНИЕ

Научная модель является *отображением* некоторых интересующих нас явлений (например, определенных объектов, событий, процессов, систем) и используется в целях управления и предсказания. Основная функция научной модели заключается не в том, чтобы описывать явления, а в том, чтобы объяснить их. Модель должна помочь выяснить, каким образом некоторые стороны явления влияют на другие стороны или же на явление в целом. Если построена достаточно верная модель, то эти вопросы можно выяснить, производя соответствующие опыты на модели, не меняя характеристик изучаемого объекта.

Преимущества использования модели для этих целей особенно очевидны, когда опыты на самом объекте или невозможны, как, например, в астрономии, или очень дороги, как в сложных промышленных организациях *. Но значение моделей этим далеко не исчерпывается. В самом деле, в некотором смысле научные теории, объясняющие определенные явления, аналогичны моделям этого явления, поэтому наука не могла бы существовать без моделей, как она не могла бы существовать без теории.

Так как модели играют такую важную роль в исследовательском процессе**, следует подробно рассмотреть различные типы моделей, их важные логические свойства и некоторые наиболее существенные зависимости между этими типами моделей, а также их связь с отображаемыми явлениями***.

Существующие модели можно разделить на три типа: изобразительные (модели геометрического подобия), модели-аналоги и символические (математические). Дадим их краткую характеристику. 1. Изобразительная модель отражает внешние характеристики системы (как фотография или модель самолета). 2. Модель-аналог использует ряд свойств одного явления для отображения свойств другого явления (например, в некоторых случаях поток воды через трубы можно принять за аналог «потока» электричества по проводам). 3. Символическая модель использует символы для отображения свойств изучаемой системы (с помощью математического уравнения или системы уравнений).

Это, конечно, очень краткая характеристика типов моделей. Чтобы дать полное и точное описание каждого типа и особенно их отличительных характеристик, потребовалось бы сложное исследование, включающее символическую логику и формальную семантику. Однако здесь мы обойдемся

* Это, конечно, не означает, что следует отказаться от всяких экспериментов с реальной системой. При длительном использовании модели всегда необходимо испытание модели, для чего обычно требуется проводить определенные опыты с реальной системой. Кроме того, часто бывает необходимо провести эксперименты с системой ввиду отсутствия необходимых данных для построения и оценки модели, а также для проверки модели путем изучения работы реальной системы в прошлом.

** В некоторых областях науки, например в физике и химии, используют ряд общепринятых моделей (например, модели атома). Существует также ряд уже разработанных моделей, которые можно применить для решения задач исследования операций (ниже дается обзор этих моделей).

*** Следует подчеркнуть, что сходство модели с оригиналом всегда является неполным. Модель лишь приблизительно отражает некоторые свойства оригинала. Такие модели называются *гомоморфными* в отличие от *изоморфных*, не отличимых от оригинала в пределах допустимой системы экспериментов. (Прим. ред.)

некоторые сложные задачи, связанные с таким описанием, и ограничимся приведенной характеристикой. Тем не менее нужно всегда помнить, что о моделях можно сказать гораздо больше, чем сказано в этой главе.

Следует отметить, что три упомянутых типа моделей представляют своего рода определенную последовательность. Изобразительная модель обычно является простейшей и наиболее конкретной моделью. Обычно она описывает явления, а не объясняет их, а потому с ее помощью редко можно установить причинные связи. Точно так же трудно с ее помощью определить или предсказать последствия изменения различных параметров реальной системы. Символическая модель обычно является самой сложной, а также наиболее общей и абстрактной. Она чаще служит для объяснения явлений, а не описания их. Кроме того, с помощью этой модели можно определять и предсказывать последствия изменения параметров реальной системы. Модели-аналоги по двум этим характеристикам можно поместить между изобразительными и символическими моделями.

Изобразительные модели (модели геометрического подобия). Изобразительная модель подобна оригиналу. Многие фотографии, картины и скульптуры являются изобразительными моделями людей, различных предметов или сцен. Игрушечный автомобиль является изобразительной моделью «настоящего» автомобиля. Глобус является моделью земного шара. В астрономии пользуются моделями солнечной системы, а в физике — моделями атомов и молекул. В общем случае всякое отображение представляет собой изобразительную модель в той мере, в какой его свойства совпадают со свойствами оригинала. Правда, эти свойства обычно подвергаются метрическому преобразованию, т. е. берется определенный масштаб. Например, глобус имеет уменьшенный диаметр по сравнению с земным шаром, хотя форма и относительные размеры континентов, морей и т. д. приблизительно правильны. Модель атома, наоборот, имеет увеличенные размеры, чтобы его можно было разглядеть невооруженным глазом. Масштаб в модели вводится для экономии и удобства пользования. В обычных условиях гораздо легче работать с моделью здания, атома или производственной системы, чем с самим объектом. Так, с опытным заводом, который является уменьшенной моделью полного завода, работать гораздо легче, чем с настоящим заводом.

Изобразительные модели хорошо приспособлены для отображения статического или динамического явления в определенный момент времени. Например, фотография или схема производственных потоков может дать хорошую «картину» работы завода. Но такие модели не подходят для отображения динамики явлений, например для отображения рабочих операций на заводе. Поэтому они не годятся для изучения изменяющегося процесса, или динамики системы. Можно, конечно, построить маленькую действующую модель завода. Однако было бы слишком дорого строить и изменять модель для изучения возможных усовершенствований в работе завода.

Однако, хотя изобразительная модель и подобна оригиналу, она, как и другие типы моделей, отличается от оригинала и не может отразить всех его свойств. В ней отображаются только свойства оригинала, существенные для задач, решаемых с помощью данной модели. Этой избирательностью во многом определяется экономичность использования любой научной модели. Нет надобности отображать в модели свойства, несущественные для определенного исследования. Это справедливо не только для научных исследований. Так, для изучения задачи стоянки автомобилей вполне подходят модели автомобилей без обивки и моторов. Точно так же для некоторых

целей на глобусе или карте не нужно отображать рельеф (т. е. возвышенности и низменности).

Модели-аналоги. При построении модели различных объектов, событий, процессов или систем не всегда можно простым изменением масштаба изобразить все интересующие нас свойства. Например, мы не можем наглядно представить на глобусе геологическую структуру Земли. Но мы легко можем представить различные геологические формации с помощью разноцветной окраски. При этом мы производим подмену одного свойства (цвет) другим (геологическая структура) в соответствии с некоторыми правилами преобразования. В картографии, например, такое преобразование является узаконенным, причем правила для преобразования приводятся в легенде. В легенде на карте приводится также перечень обозначений: например, сплошной линией обозначается грунтовая дорога, а пунктирной — шоссе. Такая модель называется моделью-аналогом, поскольку в ней совокупность одних свойств представляется с помощью совокупности других свойств.

Примером простой аналогии являются графика. На графиках пользуются расстоянием для отображения таких свойств, как время, число, проценты, возраст, вес, и многих других. Графики часто удобны для представления количественных соотношений и дают возможность предсказывать, как изменения одного свойства сказываются на другом свойстве.

Используя модели-аналоги, мы увеличиваем наши возможности проверять на модели изменения различных параметров. Обычно проще изменить модель-аналог, чем изобразительную модель. Например, горизонталь на карте отображают рельеф местности. Проще изменить такие горизонталь в двумерной карте, чем переделать рельеф трехмерной изобразительной модели.

Модели-аналоги удобны для отображения динамических процессов или систем. Можно построить модель, работа которой будет аналогична работе конвейера на заводе. Или можно отобразить колебания спроса путем соответствующего изменения некоторой входной величины модели. Однако на изобразительной модели, например уменьшенной действующей модели цеха, такое изменение провести трудно.

Другим преимуществом модели-аналога по сравнению с изобразительной моделью является большая универсальность этой модели. Так, незначительно изменяя модель, можно отобразить различные процессы одного класса.

Ярким примером модели-аналога является схема информационных и материальных потоков, изображенная на рис. 2.1. Уменьшенную модель предприятия нельзя эффективно использовать для изучения изменений информационной системы. А модель-аналог (т. е. схема) является очень простой и эффективной моделью, пригодной для этих целей *.

Символические модели. В символической модели элементы и их взаимосвязь задаются с помощью символов (обычно математического или логического характера). Чтобы показать построение символической модели про-

* Приведенные здесь примеры моделей-аналогов следует дополнить наиболее гибким и разнообразным видом таких моделей — универсальными аналоговыми устройствами. Такие устройства позволяют строить для изучения процессов в широком классе динамических систем модели-аналоги, основывающиеся на аналогии между протеканием переходных процессов в механических, тепловых, электрических и других динамических системах. (Прим. ред.)

песса, рассмотрим задачу управления запасами, описанную в главе 2. Помесячное удовлетворение производственных требований можно представить с помощью следующей графической аналогии:

Вход *	1	1	1	1	1	1	1	1	1	1	1	1
Уровень запаса	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0	↓ 0
Выход	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1

Отсюда можно определить, что средние запасы равны нулю. Если же через месяц поступают двухмесячные запасы, то получим другую модель:

2	2	2	2	2	2
↓ 1 0	↓ 1 0	↓ 1 0	↓ 1 0	↓ 1 0	↓ 1 0
↓ 1	↓ 1	↓ 1	↓ 1	↓ 1	↓ 1

В этом примере средний уровень запасов равен $1/2$ от месячной потребности. Можно продолжать и дальше такие построения и давать для каждого процесса новую схему. Но гораздо проще представить этот процесс символически.

Обозначим через x число месячных потребностей производства, удовлетворяемых одним пополнением запасов. Тогда можно представить средний уровень месячных запасов с помощью следующего символического выражения:

$$\frac{x-1}{2}.$$

Во многих случаях построение моделей-аналогов затруднительно, поскольку изучение динамики явления отнимает много времени. Например, чтобы изучить с помощью аналоговой модели влияние колебаний спроса на производственный процесс, нужно проделать на модели много опытов. Если же систему можно представить с помощью математического выражения (как в главе 2), то влияние изменений какого-нибудь параметра можно установить с помощью математической дедукции за несколько шагов. Поэтому в этой главе и в остальных главах книги говорится в основном о символических моделях. Однако следует помнить, что могут возникнуть задачи, когда аналоговые модели оказываются предпочтительнее символических, например, когда изучается настолько сложная система, что для построения символической модели потребовалось бы провести очень большую работу.

ПОСТРОЕНИЕ МОДЕЛИ

Следует напомнить, что для постановки задачи необходим анализ системы, исследование ее особенностей и возможных методов управления системой. Схема, построенная в результате такого анализа (см., например, схему рис. 2.1), является либо изобразительной, либо аналоговой моделью. Таким образом, первый этап построения модели выполняется в процессе постановки задачи. После такого анализа системы уточняется перечень различных вариантов решения, которые надо оценить. Затем определяются меры общей эффективности этих вариантов. Следовательно, следующий этап заключается в построении такой модели, в которой эффективность системы можно выразить в функции переменных, определяющих систему. Некоторые из этих переменных в реальной системе можно менять (например, размер партии деталей в задаче главы 2), другие переменные менять нельзя

* Все величины даются в пересчете на один месяц.

(например, требования потребителя). Те переменные, которые можно изменять, назовем «управляемыми». Различные варианты решения задачи необходимо выразить с помощью управляемых переменных.

Можно кратко охарактеризовать роль символической модели в исследовании операций, также прибегая к помощи символов.

Допустим, что E означает меру общей эффективности, X_i представляют собой управляемые переменные, а Y_j — неуправляемые переменные системы. При построении модели мы пытаемся вывести одно или несколько уравнений вида

$$E = f(X_i Y_j).$$

Чтобы получить решение из такой модели, надо определить значения управляемых переменных X_i , при которых мера общей эффективности максимальна.

В задаче может быть даже одна управляемая переменная (например, число партий за год в уравнении, определяющем оптимальный размер партии). Кроме того, в некоторых случаях удобно использовать меру «неэффективности» (например, ожидаемые издержки, а не ожидаемую прибыль). В таких случаях следует минимизировать эту меру.

Элементы системы. Построение символической модели системы можно начать с перечисления всех элементов системы, которые влияют на эффективность работы системы. Если в качестве меры общей эффективности используются «общие ожидаемые издержки» (как в задаче главы 2), то можно начать с исследования изобразительной или аналоговой модели системы, полученной на стадии постановки задачи. Можно выделить операции и материалы, которым сопоставляются некоторые затраты. При этом получим, например, следующий исходный список:

1. **Производственные затраты:** (а) закупочная цена сырья, (б) издержки перевозки сырья, (в) стоимость приемки сырья, (г) стоимость хранения сырья, (д) стоимость планирования производства (организация группы планирования), (е) стоимость наладочных работ в цехе, (ж) стоимость процесса обработки, (з) стоимость хранения запасов в процессе производства, (и) стоимость завершения производства и передачи готовых изделий на склад, (к) стоимость анализа результатов работы группой планирования, (л) стоимость хранения готовых изделий.

2. **Затраты на сбыт.**

3. **Накладные расходы.**

Соотношение элементов. После того как получен полный список элементов системы, следует определить, нужно ли учитывать каждый из этих элементов. Для этого надо определить, на каждый ли из этих элементов влияет выбор одной из возможных стратегий. Часто один или несколько элементов (например, постоянные затраты) не зависят от выбора стратегии. Если задача заключается в определении наиболее экономичного размера партии, то можно не учитывать затраты на сбыт, если на них не влияет размер партии. Если затраты на сбыт не учитываются, то «общие издержки» нужно заменить новой функцией стоимости «общие производственные издержки». Во многих случаях, хотя на отдельные элементы влияет выбор стратегии, это влияние может быть очень мало по сравнению с влиянием на другие элементы. В примере, приведенном в главе 2, была опущена стоимость хранения запасов в процессе производства, потому что она составляет незначительную часть от общих производственных издержек.

На этом этапе построения модели может быть неизвестно, насколько сильно влияет каждый элемент на целевую функцию. Можно предположить, что влияние определенного элемента незначительно, и временно исключить его из рассмотрения. Но это предположение следует проверить, когда появится соответствующая информация или средства исследования.

В некоторых случаях наши знания системы недостаточны, чтобы сказать с уверенностью, что составлен верный список элементов. В таких случаях желательно проверить экспериментально или статистически, являются ли учтенные элементы существенными. Другими словами, мы хотим экспериментально или с помощью анализа определить, влияют ли перечисленные переменные на общую эффективность системы. Нужно провести определенные исследования, выдвигать различные предположения, проверять их и объяснять, *почему* система действует именно так, а не иначе. Чем объясняются наблюдаемые явления? Какими параметрами можно управлять, чтобы получить желаемый результат?

При таком анализе часто полезны методы планирования эксперимента (описанные в IX части). Например, когда исследовались причины «ложных вызовов» в системе защиты от грабителей (см. главу 5), то сначала было мало что известно о причинах неисправностей в этих системах. Предполагалось, что в случае появления грабителей в охраняемых помещениях с помощью электрических устройств, установленных в магазинах, на складах и предприятиях, в центральную контору будут передаваться сигналы по телефонному кабелю. В качестве детекторов использовались фотоэлементы, реле и другие устройства. Однако оказалось, что в большинстве случаев поступают ложные сигналы. Выдвигались различные объяснения возможных причин таких ложных сигналов, например: птицы, кошки, незапертые окна, природные условия. Был составлен перечень приблизительно 100 причин. С помощью методов наблюдения и планирования эксперимента были устранены несущественные причины и составлен окончательный небольшой список действительно важных причин.

Объединение и расчленение элементов. Иногда может оказаться удобным группировать определенные элементы системы. Например, можно объединить такие элементы, как закупочная цена сырья, издержки перевозки и стоимость приемки сырья, в одну категорию: «стоимость приобретения сырья». Возможно, что на эту стоимость приобретения не влияет размер партии. Но даже в этом случае нельзя опускать эту стоимость, так как она входит в стоимость производства, а стоимость запасов готовой продукции в свою очередь зависит от вложенного в продукцию капитала. Следовательно, чтобы рассчитать, какую часть составляет вложенный капитал от стоимости запасов, нужно учесть стоимость сырья. Нельзя пропускать такие факторы, которые влияют хотя бы косвенно на меру общей эффективности.

Введение символов. Для каждого элемента, оставшегося в окончательном списке, необходимо определить, является ли он постоянным или переменным. Если некоторый элемент является переменным, то необходимо определить те характеристики системы, которые влияют на его величину. Например, стоимость обработки обычно складывается из (1) числа обрабатываемых деталей и (2) стоимости обработки одной детали.

Аналогично стоимость хранения готовых деталей зависит от (1) числа деталей на складе, (2) времени хранения, (3) стоимости хранения одной детали.

После того как проведено такое разбиение каждого переменного элемента в окончательном списке, следует приписать соответствующий символ каждому подэлементу. В примере, приведенном в этой главе, получен следующий список символов:

c_1 — средняя стоимость подготовительно-заключительных работ на партию, c_2 — средняя стоимость сырья и стоимость обработки на деталь, P — средняя стоимость хранения готовой детали в месяц, выраженная в процентах от вложенного капитала, L — средняя месячная потребность в деталях, R — число одинаковых партий за год, K — общие ожидаемые затраты на производство деталей, требующихся на 1 год (без учета стоимости хранения запасов в процессе производства). Последний символ используется для обозначения меры общей эффективности.

В некоторых случаях мы можем вывести одно выражение, показывающее общую эффективность процесса или системы в функции от различных элементов, для которых были введены символы. Иногда же требуется система уравнений. Для иллюстрации рассмотрим простой случай с одним уравнением. Для этого вернемся к примеру главы 2 и покажем, как получается первая (пробная) символическая модель системы производства.

Пример построения символической модели *. В построенной модели используются вышеперечисленные символы, а также дополнительные, введенные для удобства: $N = 12L/R$ — число деталей в партии (12 означает число месяцев в году), $K_R = K/R$ — общие ожидаемые затраты на партию.

Общие затраты на партию складываются из трех составных частей: стоимость подготовительно-заключительных работ, стоимость производства и материалов и стоимость капитала, вложенного в запасы. (Считается, что колебания стоимости запасов в процессе производства незначительны и не входят в общие затраты.) Уравнение модели выводится так:

$$\text{Средняя стоимость наладки на партию} = c_1, \quad (1)$$

$$\text{Средняя стоимость производства и материалов на партию} = Nc_2, \quad (2)$$

Средняя стоимость капитала в запасах:

$$\left(\frac{c_1}{N} + c_2 \right) \text{ — средняя стоимость капитала на деталь,} \quad (3a)$$

поступающую на склад,

$$P \left(\frac{c_1}{N} + c_2 \right) \text{ — средняя стоимость хранения детали в месяц,} \quad (3б)$$

$$N/L \text{ — число месячных потребностей, удовлетворяемых} \quad (3в)$$

партийей.

Затем необходимо определить величину запасов после выпуска партии. Для этого воспользуемся графической аналогией, приведенной выше. Здесь следует напомнить предположение, что детали поступают со склада на сборку помесечно и, кроме того, что обработка деталей заканчивается точно в момент выдачи деталей со склада на сборку **.

* Вывод уравнений для определения оптимального размера партии при условии известного спроса производился еще Ф. В. Харрисом в 1915 году. Только недавно экономисты, математики и операционисты обобщили эти результаты на случай переменных сроков поставки (см. часть IV). В литературе можно найти много моделей, аналогичных приведенной модели (см., например, [2, 4, 9, 12, 14—17, 19 и 25]).

** Таким образом, в нашей модели предполагается дискретное поступление на склад и дискретная выдача со склада, причем выдача происходит непосредственно за поступлением. Интересно отметить, что полученная символическая модель справедлива также для непрерывных поступлений и выдачи. Для третьей модели (дискретные поступления и непрерывная выдача) получается аналогичное в численном отношении выражение, приведенное во многих работах, ссылки на которые даются в данной главе.

Таким образом, если поставки выполняются ежемесячно, то получим следующую схему:

Вход	1	1	1	
Запас	↓ 0	↓ 0	↓ 0	
Выход	1	1	1	

Если поставки производятся через месяц, то получим другую картину:

Вход	2		2		2
Запас	↓ 1	0	↓ 1	0	↓
Выход	1	1	1	1	1

Если поставки выполняются раз в четыре месяца, то получим:

Вход	4				4
Запас	↓ 3	2	1	0	↓
Выход	1	1	1	1	1

Из этой схемы видно, что если поставка удовлетворяет потребность на x месяцев, то общее число месячных запасов в партии составляет

$$(x-1) + (x-2) + \dots + 1,$$

что эквивалентно выражению

$$\frac{x(x-1)}{2}.$$

Но одна партия потребляется в течение N/L месяцев, т. е. $x = N/L$. Тогда

$$\left(\frac{N}{2L}\right) \left(\frac{N}{L} - 1\right) - \text{число месячных запасов в партии}^*, \quad (3г)$$

$$L \frac{N}{2L} \left(\frac{N}{L} - 1\right) = \frac{N}{2} \left(\frac{N}{L} - 1\right) - \text{величина запасов на партию,} \quad (3д)$$

выраженная в деталях,

$$\frac{N}{2} \left(\frac{N}{L} - 1\right) P \left(\frac{c_1}{N} + c_2\right) - \text{средняя общая стоимость хранения на партию.} \quad (3е)$$

Следовательно, ожидаемые общие затраты на партию равны

$$\begin{aligned} K_R &= c_1 + Nc_2 + \frac{N}{2} \left(\frac{N}{L} - 1\right) P \left(\frac{c_1}{N} + c_2\right) = \\ &= c_1 + Nc_2 + \frac{Pc_1}{2} \left(\frac{N}{L} - 1\right) + \frac{NPc_2}{2} \left(\frac{N}{L} - 1\right), \end{aligned} \quad (4а)$$

или так как $N = 12L/R$, то

$$K_R = c_1 + \frac{12L}{R} c_2 + \frac{Pc_1}{2} \left(\frac{12}{R} - 1\right) + \frac{12LPc_2}{2R} \left(\frac{12}{R} - 1\right). \quad (4б)$$

Но

$$K = RK_R - \text{ожидаемые общие затраты за год.} \quad (5а)$$

* Выражение (3г) является точным, только если N/L — целое число; в противном случае получаем приближенное выражение, но приближение достаточно хорошее.

Следовательно,

$$K = Rc_1 + 12Lc_2 + \frac{Pc_1R}{2} \left(\frac{12}{R} - 1 \right) + \frac{Pc_212L}{2} \left(\frac{12}{R} - 1 \right), \quad (5б)$$

и окончательно

$$K = Rc_1 + 12Lc_2 + \frac{Pc_1R}{2} \left(\frac{12}{R} - 1 \right) + 6Pc_2L \left(\frac{12}{R} - 1 \right). \quad (5в)$$

Выражение (5в) представляет собой математическую модель простой и очень частной задачи управления системой производства и хранения, выраженной через ожидаемые общие ежегодные затраты производства. Эта модель получается значительно сложнее, если учесть переменный спрос и переменное «запаздывание» (лаг) производства. Более реальные модели запасов будут даны в IV части. Но здесь не ставится задача разработать общую модель процессов производства и хранения; мы хотели лишь показать метод построения символической модели.

ОТ МОДЕЛИ К РЕШЕНИЮ

Модель представляет собой инструмент, с помощью которого можно эффективно оценивать различные стратегии управления. Выбор метода решения задачи зависит от характеристик модели. Существующие методы решения можно разбить на два типа: аналитический и численный. Аналитический метод по своему характеру является в основном дедуктивным, а численный метод (метод проб и ошибок) в основном является индуктивным. В некоторых случаях нельзя применить ни один из методов, так как некоторые параметры необходимо оценить при помощи так называемого метода Монте-Карло. Ниже рассматриваются последовательно аналитический и численный методы, а затем метод Монте-Карло.

Переходя от модели к решению, следует помнить, что стратегия, которая является оптимальной для модели, может оказаться неприемлемой для реальной системы, поскольку модель не может быть точным отображением реального явления. Например, могут быть пропущены некоторые важные переменные или функция цели (такая, как K) может быть выбрана неправильно (например, минимизация времени может быть гораздо важнее минимизации затрат). Поэтому, когда мы говорим о «решении», то имеется в виду решение для модели, а не обязательно для реальной системы, отображением которой является данная модель.

Аналитическое решение. Рассмотрим полученную модель управления запасами (выражение (5в)). Необходимо установить, сколько деталей следует производить за один производственный цикл, т. е. определить оптимальный размер партии.

Из выражения (5в) легко определить, что эта задача сводится к другой: какова должна быть величина R (число равных партий в году), чтобы K было минимальным (общие ожидаемые затраты за год)?

Эту задачу можно решать двумя способами. Можно изобразить графически зависимость K от различных значений R и выбрать ту величину R , для которой K минимально. Или можно получить общее решение, используя методы дифференциального исчисления. При этом для минимального K получим значение R в функции от других переменных. Это значение R можно определить, найдя оптимальный размер партии.

Сначала берется производная от K (см. (5в)) по R . Имеем

$$\frac{dK}{dR} = c_1 - \frac{Pc_1}{2} - \frac{72LPc_2}{R^2}. \quad (6a)$$

Эта производная приравняется нулю:

$$c_1 - \frac{Pc_1}{2} - \frac{72LPc_2}{R^2} = 0, \quad (6б)$$

откуда определяется

$$R^2 = \frac{72LPc_2}{(c_1 - Pc_1/2)} = \frac{144LPc_2}{c_1(2-P)}, \quad (6в)$$

или

$$R = \sqrt{\frac{144LPc_2}{c_1(2-P)}}. \quad (6г)$$

— оптимальное число партий за год, т. е. величина R , которая минимизирует K при условии

$$\frac{d^2K}{dR^2} > 0, \quad (6д)$$

но

$$\frac{d^2K}{dR^2} = 144LPc_2R^{-3} = \frac{144LPc_2}{R^3} > 0. \quad (6е)$$

Поскольку $N = 12L/R$, то оптимальный размер партии N_0 задается в виде

$$N_0 = 12L : \sqrt{\frac{12^2LPc_{12}}{c_1(2-P)}} = 12L : 12 \sqrt{\frac{LPc_2}{c_1(2-P)}} = L : \sqrt{\frac{LPc_2}{c_1(2-P)}},$$

т. е.

$$N_0 = \sqrt{\frac{Lc_1(2-P)}{c_2P}}. \quad (6ж)$$

Таким образом, выражение (6ж) определяет оптимальный размер партии в соответствии с принятой моделью.

При изменении модели меняются и методы ее решения. В только что приведенном решении используются элементарные методы дифференциального исчисления и алгебры. Предположим, что в полученной модели есть два параметра, которые можно менять. Например, в модели химического процесса планируются два продукта, которые нераздельно проходят несколько стадий производственного процесса. В этом случае получается результирующее выражение для стоимости в виде

$$K = f(X_1, X_2, Y_1, Y_2, \dots, Y_m),$$

где X_1 и X_2 — управляемые переменные, а Y_j ($j = 1, 2, \dots, n$) — неуправляемые переменные. Чтобы получить решение такой модели, можно взять частные производные от K по X_1 ($\partial K / \partial X_1$) и по X_2 ($\partial K / \partial X_2$). Затем надо приравнять нулю оба полученных выражения и решить эти уравнения относительно X_1 и X_2 *.

Этот метод (дифференцирование и решение системы однородных уравнений) применим к моделям, содержащим любое число управляемых переменных. Но с увеличением числа переменных вычисления становятся очень

* Приравнивание нулю частных производных является необходимым условием. Необходимые и достаточные условия минимума или максимума функции двух переменных можно найти в специальных учебниках по дифференциальному исчислению.

сложными. Правда, во многих случаях можно использовать для расчетов вычислительные машины.

Необходимо также учесть, что при наличии различных ограничений в системе модель получается значительно сложнее. Например, предположим, что в представленной выше модели необходимо учесть ограничения по мощности машин. Для этого необходимо добавить к модели неравенство *, отображающее ограничение общего времени на определенной группе машин. Такое неравенство имеет, например, следующий вид:

$$n_A t_A + n_B t_B + \dots + n_K t_K \leq T, \quad (7)$$

где T — резерв времени определенной группы машин, n_A — число планируемых деталей A , t_A — время обработки одной детали A на данной машине. Чтобы учесть такое ограничение при решении модели, необходимо использовать метод множителей Лагранжа или какой-либо вариант этого метода. В главе 10 представлены модели этого типа и их решение.

Во многих случаях удобнее описывать изучаемую систему не одним уравнением, а системой уравнений. Например, предположим, что у нас есть два источника снабжения A и B с запасом материалов в количествах Q_A и Q_B . Кроме того, допустим, что есть два пункта назначения 1 и 2, потребности которых выражаются в количествах R_1 и R_2 , причем $R_1 + R_2 \leq Q_A + Q_B$. И наконец, заданы различные стоимости перевозки грузов из A в 1 (C_{A1}), из A в 2 (C_{A2}), из B в 1 (C_{B1}) и из B в 2 (C_{B2}). Задача заключается в том, чтобы осуществить перевозку грузов из пунктов отправления A и B в пункты назначения 1 и 2 с минимальными транспортными издержками.

Эту систему нетрудно описать с помощью системы уравнений и неравенств. Предположим, N_{A1} представляет собой количество перевезенного груза из A в 1 и т. д. Тогда общая стоимость перевозки K выражается в виде

$$K = N_{A1}C_{A1} + N_{A2}C_{A2} + N_{B1}C_{B1} + N_{B2}C_{B2}, \quad (8)$$

но необходимо учесть ограничения

$$\left. \begin{aligned} N_{A1} + N_{A2} &\leq Q_A, & N_{B1} + N_{B2} &\leq Q_B, \\ N_{A1} + N_{B1} &= R_1, & N_{A2} + N_{B2} &= R_2. \end{aligned} \right\} \quad (9)$$

Эту систему можно представить с помощью матрицы, показанной в таблице 7.1.

Таблица 7.1

Пункты отправления	Количество груза в пунктах отправления	Пункты назначения		Общее количество перевезенного груза
		1	2	
A	Q_A	N_{A1}	N_{A2}	$\leq Q_A$
B	Q_B	N_{B1}	N_{B2}	$\leq Q_B$
Общее количество потребного груза		R_1	R_2	

* Неравенство вытекает из ограничений, выраженных в форме «меньше чем», «больше чем», «самое большое», «по крайней мере» и т. д. Термин «меньше чем» обозначается символом $<$, «меньше или равно» обозначается $\leq$, «больше чем» — символом $>$ и «больше или равно» обозначается символом $\geq$.

Используя приведенную матрицу, данную задачу можно выразить следующим образом: определить значения N_{A1} , N_{A2} , N_{B1} и N_{B2} , при которых общие транспортные издержки K будут минимальны при условии, что выполняются поставленные ограничения. Такую задачу можно решить с использованием матричной алгебры. Модели такого типа и их решение приводятся в части V. Хотя эти задачи можно решать аналитически путем математической дедукции, однако их удобнее решать с помощью метода проб и ошибок или каким-либо иным итеративным методом. Эти вопросы освещены в следующем разделе.

Модель можно представить в различных математических формах, а следовательно, для получения решения применяется различный математический аппарат, что требует высокой математической подготовки. Но вывести математическое выражение модели гораздо проще, чем решить ее. Поэтому группа исследования операций не должна состоять только из математиков, но математики обязательно должны входить в эту группу.

Численные решения. Численные методы получения решения модели предполагают подстановку чисел вместо символов и определение таких чисел, которые дают оптимальное решение. Например, можно определить оптимальную величину управляемой переменной (или нескольких переменных) в символической (математической) модели, если подставлять различные значения переменной и вычислять при каждой подстановке общую эффективность. Затем надо выбрать такую совокупность переменных, которой соответствует максимальное значение общей эффективности системы. Однако эта процедура очень длительна и дорога даже при использовании электронных вычислительных машин, хотя обычно и нет необходимости пробовать все возможные значения, так как можно разработать методы, в которых последующая подстановка улучшает результаты, полученные с помощью предыдущей подстановки. Решение проводится до тех пор, пока дальнейшая подстановка уже не дает значительных улучшений. С помощью этого метода можно получить решение за меньшее число шагов по сравнению с длительным процессом решения по методу проб и ошибок. Такой метод, который является, по существу, улучшенным методом проб и ошибок, называется *итеративным*.

В качестве примера использования итеративного процесса рассмотрим простую модель, которую можно представить в виде матрицы и одного неравенства. Предположим, что коммивояжер имеет дело с двумя клиентами A и B , от каждого из которых он получает определенный доход в зависимости от потраченного на него времени. Но коммивояжер может истратить на обоих клиентов только 6 часов. Каким образом он должен распределить свое время, чтобы получить максимальный доход?

Все данные приведены в таблице 7.2. Для заданной функции дохода (параболической или квадратической) можно воспользоваться следующим итеративным процессом.

1. Можно начать с распределения времени поровну между двумя клиентами, т. е. по 3 часа на каждого. Тогда общий доход будет $\$18 + \$28 = \$46$.

2. Определим, для какого клиента дополнительный час времени даст большее увеличение дохода. Получаем, что для B увеличение дохода равно $\$34 - \$28 = \$6$.

3. Сравним возможное увеличение дохода, полученное на втором шаге, с уменьшением дохода за счет уменьшения времени на 1 час на клиента A . Получим $\$18 - \$14 = \$4$.

Таблица 7.2

Время, потрачен- ное на клиента	Клиент А		Клиент В		Время, потрачен- ное на клиента	Клиент А		Клиент В	
	ожидае- мый до- ход	при- рост дохода	ожидае- мый до- ход	при- рост дохода		ожидае- мый до- ход	при- рост дохода	ожидае- мый до- ход	при- рост дохода
1	\$ 8	\$ 8	\$ 12	\$12	5	\$ 21	\$ 1	\$ 38	\$ 4
2	14	6	21	9	6	21	0	41	3
3	18	4	28	7	7	21	0	42	1
4	20	2	34	6	8	21	0	42	0

4. Поскольку уменьшение меньше, чем увеличение дохода ($\$4 < \6), перераспределим время следующим образом: 2 часа на клиента А и 4 часа на клиента В, общий доход равен \$48.

5. Продолжаем до тех пор, пока получение чистого прироста дохода становится невозможным. В нашем примере больший доход получить нельзя. Увеличение времени по одному часу на каждого клиента дает прирост дохода в \$4, в то время как сокращение времени на один час у каждого клиента дает уменьшение дохода на \$6. Следовательно, при дальнейшем перераспределении времени в любом случае получим сокращение дохода на \$2.

Таким образом, оптимальным распределением времени в данном примере * будет 2 часа на клиента А и 4 часа на клиента В.

Более точное описание итеративного процесса приведено в приложении в конце этой главы. В части V обсуждаются различные более сложные итеративные методы для решения более реальных задач.

Метод Монте-Карло. При построении модели системы желательно использовать переменные, значения которых достаточно легко получить. Однако некоторые выражения в модели, включающие очень простые переменные, могут быть сами по себе очень сложными. Особенно часто это имеет место, когда переменные являются случайными величинами или функциями.

Рассмотрим, например, новое изделие, состоящее из двух деталей, которые с течением времени могут выходить из строя. Допустим, что детали представляют собой лампу и конденсатор. Из предыдущих опытов мы знаем вероятность выхода из строя каждого элемента в функции времени, т. е. мы знаем так называемую «кривую срока службы» каждого элемента. Но мы хотим знать «кривую срока службы» всего изделия, в которое входят оба элемента. Другими словами, если $f(t)$ является «кривой срока службы» одного из элементов, а $g(t)$ — другого элемента, то результирующая «кривая срока службы» всего изделия является функцией этих двух кривых, т. е. $h[f(t), g(t)]$ или просто $h(t)$. Если в модели системы появляется такое выражение, то до тех пор, пока не произведена его оценка, нельзя получить решение модели ни аналитическим путем, ни с помощью итеративного процесса. Таким образом, если известны функции $f(t)$ и $g(t)$, то, прежде чем получить решение, надо получить функцию $h(t)$ в явном виде.

В некоторых случаях $h(t)$ можно получить математическим путем, когда $f(t)$ и $g(t)$ представляют собой нормальные плотности вероятности.

* Решение можно было бы получить непосредственно из таблицы 7.2. Однако нашей целью является не решение этой или любой другой задачи. Мы стремимся показать процедуру решения с помощью итеративного процесса.

Однако это не всегда возможно, а иногда практически неосуществимо. Для примера можно привести вычисление некоторых важных выражений в моделях игр и замены (см. части VI и VII). Выход из этого затруднительного положения заключается в использовании метода Монте-Карло (статистических испытаний), который дает возможность приближенно оценить искомые выражения, состоящие из одной или нескольких функций распределения вероятностей. Такие выражения очень часто встречаются в моделях исследования операций. Используя метод Монте-Карло совместно с аналитическим и итеративным методами, можно получать решения для моделей, содержащих выражения такого типа.

Метод Монте-Карло представляет собой новое применение старой процедуры «неограниченной случайной выборки», т. е. выборки отдельных элементов из множества таким образом, чтобы вероятность выбора каждого элемента была одинакова. «Новый» вариант этой процедуры заключается в использовании случайной выборки для имитации игры с природой или искусственной системой. Эта игра моделирует эксперимент. По существу метод Монте-Карло представляет собой моделирование эксперимента для определения некоторых вероятностных свойств множества объектов или событий. Это абстрактное положение можно пояснить на примере, а также показать, каким образом был разработан указанный метод.

Если существует легенда об открытии Ньютоном законов тяготения при наблюдении за падающим яблоком, то и открытие метода Монте-Карло приписывают легендарному наблюдению математика за блужданием пьяного. Предполагается, что каждый шаг пьяного может быть с равной вероятностью совершен в любом направлении. Математика интересует вопрос, сколько шагов в среднем должен сделать пьяный, чтобы пройти определенное расстояние от исходной точки. Эта задача была названа задачей о «случайном блуждании». Для решения этой задачи был применен метод случайной выборки, названный «стохастической выборкой». Со временем было обнаружено, что этот метод имеет широкое практическое применение, и постепенно ему было дано другое название — *метод Монте-Карло*.

Вернемся к задаче о «случайном блуждании». Предположим, что пьяный начинает свои блуждания от фонаря, расположенного в центре большой городской площади. Он решает идти, но без определенного направления. Если наблюдать за ним, то можно заметить, что он может сделать несколько шагов в одном направлении, затем несколько шагов в другом направлении и т. д. непредвиденным, или случайным, образом. Задача заключается в том, чтобы определить, как далеко он будет от фонаря после n беспорядочных зигзагов. То есть каково наиболее вероятное расстояние от фонаря после n шагов?

Но каким образом можно оценить это наиболее вероятное расстояние без наблюдения за многими пьяными в аналогичных обстоятельствах? А такие широкие опыты провести практически невозможно. Но можно воспользоваться тем обстоятельством, что пьяный движется случайным образом, и промоделировать кривую его движения с помощью таблицы случайных чисел (см. таблицу 7.3) *. После моделирования большого числа таких опытов мы можем оценить вероятное расстояние от исходной точки после n беспорядочных шагов.

Для примера покажем, как можно применить метод Монте-Карло к задаче о «случайном блуждании». Для этого оценим наиболее вероятное

* Об этих таблицах см. [1, гл. V]. Относительно полной таблицы см. [21], где содержится также библиография по таблицам и работам по этому вопросу.

Таблица 7.3

Случайные числа

09 73 25 33	76 53 01 35 86	34 67 35 48 76	80 95 90 90 17	39 29 27 49
54 20 48 05	64 89 47 42 96	24 80 52 40 37	20 63 61 04 02	00 82 29 16
42 26 89 53	19 64 50 93 03	23 20 90 25 60	15 95 33 47 64	35 08 03 36
01 90 25 29	09 37 67 07 15	38 31 13 11 65	88 67 67 43 97	04 43 62 76
80 79 99 70	80 15 73 61 47	64 03 23 66 53	98 95 11 68 77	12 17 17 68
06 57 47 17	34 07 27 68 50	36 69 73 61 70	65 81 33 98 85	11 19 92 91
06 01 08 05	45 57 18 24 06	35 30 34 26 14	86 79 90 74 39	23 40 30 97
26 97 76 02	02 05 16 56 92	68 66 57 48 18	73 05 38 52 47	18 62 38 85
57 33 21 35	05 32 54 70 48	90 55 35 75 48	28 46 82 87 09	82 49 12 56
79 64 57 53	03 52 96 47 78	35 80 83 42 82	60 93 52 03 44	35 27 38 84
52 01 77 67	14 90 56 86 07	22 10 94 05 58	60 97 09 34 33	50 50 07 39
80 50 54 31	39 80 82 77 32	50 72 56 82 48	29 40 52 42 01	52 77 56 78
45 29 96 34	06 28 89 80 83	13 74 67 00 78	18 47 54 06 10	68 71 17 78
68 34 02 00	86 50 75 84 01	36 76 66 79 51	90 36 47 64 93	29 60 91 01
59 46 73 48	87 51 76 49 69	91 82 60 89 28	93 78 56 13 68	23 47 83 41
48 11 76 74	17 46 85 09 50	58 04 77 69 74	73 03 95 71 86	40 21 81 65
12 43 56 35	17 72 70 80 15	45 31 82 23 74	21 11 57 82 53	14 38 55 37
35 09 98 17	77 40 27 72 14	43 23 60 02 10	45 52 16 42 37	96 28 60 26
91 62 68 03	66 25 22 91 48	36 93 68 72 03	76 62 11 39 90	94 40 05 64
89 32 05 05	14 22 56 85 14	46 42 75 67 88	96 29 77 88 22	54 38 21 45
49 91 45 23	68 47 92 76 86	46 16 28 35 54	94 75 08 99 23	37 08 92 00
33 69 45 98	26 94 03 68 58	70 29 73 41 35	53 14 03 33 40	42 05 08 23
10 48 19 49	85 15 74 79 54	32 97 92 65 75	57 60 04 08 81	22 22 20 64
55 07 37 42	11 10 00 20 40	12 86 07 46 97	96 64 48 94 39	28 70 72 58
60 64 93 29	16 50 53 44 84	40 21 95 25 63	43 65 17 70 82	07 20 73 17
19 69 04 46	26 45 74 77 74	51 92 43 37 29	65 39 45 95 93	42 58 26 05
47 44 52 66	95 27 07 99 53	59 36 78 38 48	82 39 61 01 18	33 21 15 94
55 72 85 73	67 89 75 43 87	54 62 24 44 31	91 19 04 25 92	92 92 74 59
48 11 62 13	97 34 40 87 21	16 86 84 87 67	02 07 11 20 59	25 70 14 66
52 37 83 17	73 20 88 98 37	68 93 59 14 16	26 25 22 96 63	05 52 28 25
49 35 24 94	75 24 63 38 24	45 86 25 10 25	61 96 27 93 35	65 33 71 24
54 99 76 54	64 05 18 81 59	96 11 96 38 96	54 69 28 23 91	23 28 72 95
96 31 53 07	26 89 80 93 54	33 35 13 54 62	77 97 45 00 24	90 10 33 93
80 80 83 91	45 42 72 68 42	83 60 94 97 00	13 02 12 48 92	78 56 52 01
05 88 52 36	01 39 09 22 86	77 28 14 40 77	93 91 08 36 47	70 61 74 29
17 90 02 97	87 37 92 52 41	05 56 70 70 07	86 74 31 71 57	85 39 41 18
23 46 14 06	20 11 74 52 04	15 95 66 00 00	18 74 39 24 23	97 11 89 63
56 54 14 30	01 75 87 53 79	40 41 92 15 85	66 67 43 68 06	84 96 28 52
15 51 49 38	19 47 60 72 46	43 66 79 45 43	59 04 79 00 33	20 82 66 85
86 43 19 94	36 16 81 08 51	34 88 88 15 53	01 54 03 54 56	05 01 45 11
08 62 48 26	45 24 02 84 04	44 99 90 88 96	39 09 47 34 07	35 44 13 18
18 51 62 32	41 94 15 09 49	89 43 54 85 81	88 69 54 19 94	37 54 87 30
95 10 04 06	96 38 27 07 74	20 15 12 33 87	25 01 62 52 98	94 62 46 11

расстояние, пройденное после пяти одинаковых шагов (т. е. $n=5$). Обратимся к таблице 7.3, в которой приведены двузначные случайные числа. Кроме того, используем следующую символику:

1. Будем считать, что фонарь расположен в начале координат X , Y (см. рис. 7.1).

2. Примем, что первая цифра двузначного числа, выбранного из таблицы, означает положительную единицу по оси X , если получаем нуль или четное число, и отрицательную, если число нечетное.

3. Будем считать, что вторая цифра этого числа соответствует единице по оси Y , знак которой определяется так же, как в п. 2.

4. Точка (x_n, y_n) определяет положение пьяного после n шагов.

5. Расстояние пьяного от фонаря после n шагов определяется выражением

$$d_n^2 = x_n^2 + y_n^2.$$

Если мы начнем со случайной точки и выберем двузначное число, например, в 10-м столбце и в 6-й строке таблицы 7.3, а затем начнем выбирать числа сверху вниз, то мы получим следующие 5 чисел: 36, 35, 68, 90 и 35. С помощью этих чисел получим координаты движения пьяного, показанные в таблице 7.4. Кривая движения показана на рис. 7.1.

Таким образом, в этом примере получена одна оценка расстояния от фонаря после 5 шагов. Расстояние равно 3,16 единиц и получено следующим образом:

$$d_5^2 = x_5^2 + y_5^2 = 9 + 1 = 10,$$

$$d_5 = \sqrt{10} = 3,16.$$

Затем повторим указанный процесс для различных случайных чисел в таблице, чтобы получить несколько оценок искомого расстояния. Затем эти оценки следует усреднить для получения среднего расстояния. В общем случае наши оценки будут улучшаться по мере увеличения числа опытов. После проведения серии опытов мы можем оценить вероятность того, что пьяный будет находиться на определенном расстоянии от фонаря после n беспорядочных шагов.

Чтобы читатель мог сравнить свое собственное решение этой задачи методом Монте-Карло с правильным решением, отметим, что здесь можно получить аналитическое решение в виде

$$d_n = a \sqrt{n},$$

т. е. наиболее вероятное расстояние пьяного от фонаря после большого числа беспорядочных шагов равно средней длине одного шага a , умноженной на корень квадратный из числа n его шагов.

Таблица 7.4

Шаг	Первая цифра	Вторая цифра	Расположение точки (x_n, y_n)
1	3	6	$(-1, 1)$
2	3	5	$(-2, 0)$
3	6	8	$(-1, 1)$
4	9	0	$(-2, 2)$
5	3	5	$(-3, 1)$

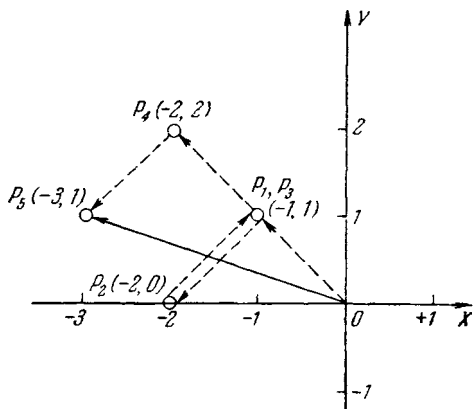


Рис. 7.1. Траектория «случайного блуждания».

Рассмотрев пример использования метода Монте-Карло, вернемся к задаче, с которой мы начали обсуждение этого метода, т. е. к задаче определения кривой срока службы изделия, содержащего две детали, кривые сроков служб которых известны. Предположим, что кривые сроков служб обеих деталей представляют собой нормальные кривые, первая из которых $f(t)$ имеет математическое ожидание 100 часов, среднее квадратическое отклонение 20 часов, а вторая $g(t)$ имеет математическое ожидание 90 часов и среднее квадратическое отклонение 10 часов. Эти две кривые показаны на рис. 7.2. Вероятностные характеристики сроков службы деталей можно также отразить в виде интегральных функций распределения (см. рис. 7.3).

Когда рассматривается множество изделий, то из каждого класса деталей выбирается случайным образом одна из деталей. Используя метод Монте-Карло, можно выбирать эти детали случайным образом и рассматривать

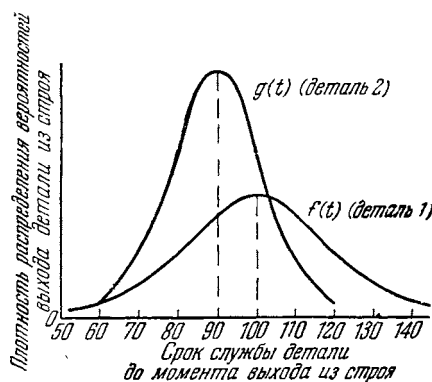


Рис. 7.2. Кривые срока службы деталей.

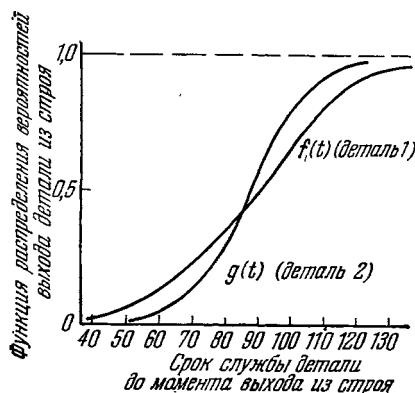


Рис. 7.3. Интегральные функции распределения срока службы деталей.

соответствующие сроки службы изделия. Однако предварительно рассмотрим, как сделать случайную выборку при нормальном распределении *. Поскольку (а) детали нужно выбирать таким образом, чтобы вероятность выбора каждой была одинаковой, и (б) гораздо больше деталей имеют срок службы в одном из временных интервалов (например, от 95 до 105), чем в других интервалах (например, от 75 до 85), то процесс выбора должен быть таким, чтобы вероятность выбора детали в любом временном интервале была пропорциональна количеству деталей, попадающих в этот интервал. Это означает, что мы не можем производить случайные выборки по оси абсцисс (горизонтальной оси) рис. 7.2 или 7.3, поскольку в этом случае мы имели бы равную вероятность выбора детали в любом временном интервале. Следовательно, мы должны производить случайные выборки по оси ординат (вертикальной оси). Для кривых распределения (рис. 7.3) выборка производится следующим образом.

Вертикальная ось от начала координат до верхней точки кривой делится, например, на 100 равных частей. Затем, используя таблицу случайных чисел (например, таблицу 7.3), выбираем числа по вертикальной оси. Для каж-

* Поскольку в нашем примере срок службы каждой из деталей подчиняется нормальному закону распределения, все объяснения относятся к нормальному закону распределения. Однако метод получения случайной выборки является общим и применим к любому закону распределения.

Таблица 7.5

Случайные нормальные числа *

$$\mu = 0, \sigma = 1$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	0,464	0,137	2,455	-0,323	-0,068	0,296	-0,288
2	0,060	-2,526	-0,531	-1,940	0,543	-1,558	0,187
3	1,486	-0,354	-0,634	0,697	0,926	1,375	0,785
4	1,022	-0,472	1,279	3,521	0,571	-1,851	0,194
5	1,394	-0,555	0,046	0,321	2,945	1,974	-0,258
6	0,906	-0,513	-0,525	0,595	0,881	-0,934	1,579
7	1,179	-1,055	0,007	0,769	0,971	0,712	1,090
8	-1,501	-0,488	-0,162	-0,136	1,033	0,203	0,448
9	-0,690	0,756	-1,618	-0,445	-0,511	-2,051	-0,457
10	1,372	0,225	0,378	0,761	0,181	-0,736	0,960
11	-0,482	1,677	-0,057	-1,229	-0,486	0,856	-0,491
12	-1,376	-0,150	1,356	-0,561	-0,256	0,212	0,219
13	-1,010	0,598	-0,918	1,598	0,065	0,415	-0,169
14	-0,005	-0,899	0,012	-0,725	1,147	-0,121	-0,096
15	1,393	-1,163	-0,911	1,231	-0,199	-0,246	1,239
16	-1,787	-0,261	1,237	1,046	-0,508	-1,630	-0,146
17	-0,105	-0,357	-1,384	0,360	-0,992	-0,116	-1,698
18	-1,339	1,827	-0,959	0,424	0,969	-1,141	-1,041
19	1,041	0,535	0,731	1,377	0,983	-1,330	1,620
20	0,279	-2,056	0,717	-0,873	-1,096	-1,396	1,047
21	-1,805	-2,008	-1,633	0,542	0,250	0,166	0,032
22	-1,186	1,180	1,114	0,882	1,265	-0,202	0,151
23	0,658	-1,141	1,151	-1,210	-0,927	0,425	0,290
24	-0,439	0,358	-1,939	0,891	-0,227	0,602	0,973
25	1,398	-0,230	0,385	-0,649	-0,577	0,237	-0,289
26	0,199	0,208	-1,083	-0,219	-0,291	1,221	1,119
27	0,159	0,272	-0,313	0,084	-2,828	-0,439	-0,792
28	2,273	0,606	0,606	-0,747	0,247	1,291	0,063
29	0,041	-0,307	0,121	0,790	-0,584	0,541	0,484
30	-1,132	-2,098	0,921	0,145	0,446	-2,661	1,045
31	0,768	0,079	-1,473	0,034	-2,127	0,665	0,084
32	0,375	-1,658	-0,851	0,234	-0,656	0,340	-0,086
33	-0,513	-0,344	0,210	-0,736	1,041	0,008	0,427
34	0,292	-0,521	1,266	-1,206	-0,899	0,110	-0,528
35	1,026	2,990	-0,574	-0,491	-1,114	1,297	-1,433
36	-1,334	1,278	-0,568	-0,109	-0,515	-0,566	2,923
37	-0,287	-0,144	-0,254	0,574	-0,451	-1,181	-1,190
38	0,161	-0,886	-0,921	-0,509	1,410	-0,518	0,192
39	-1,346	0,193	-1,202	0,394	-1,045	0,843	0,942
40	1,250	-0,199	-0,288	1,810	1,378	0,584	1,216

* Таблица частично воспроизведена из таблиц Рэнд Корпорейшн.

дого числа определяем по кривой соответствующую абсциссу. Например, на рис. 7.4 показано, как для выбранного из таблицы числа 0,55 получено число 88 часов для данной кривой распределения.

Повторяя этот процесс, мы можем получить случайную выборку из заданного нормального распределения. Однако нам не надо проделывать эту утомительную процедуру, поскольку существуют таблицы случайных нормальных чисел (см., например, таблицу 7.5). Таблицей 7.5 удобно пользоваться для определения интересующей нас кривой срока службы изделия. Для этого построим таблицу, аналогичную таблице 7.6.

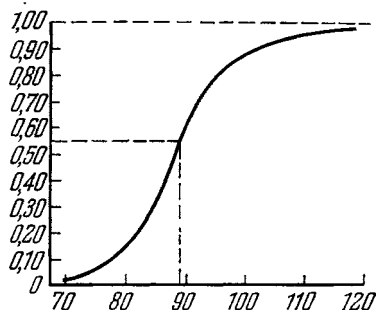


Рис. 7.4. Случайная выборка при нормальном распределении.

В столбцах (1) и (3) мы просто перечисляем числа, взятые из таблицы случайных нормальных чисел. Можно начать с любого столбца. В нашем примере выбраны первые два столбца таблицы 7.5, начиная сверху. В таблице приведены числа в долях от среднего квадратического отклонения. Следовательно, их можно превратить в часы. Для этого надо умножить величины, взятые из таблицы, на соответствующее среднее квадратическое отклонение (20 для детали 1 и 10 для детали 2) и прибавить полученное число к соответствующему среднему значению (100 для

детали 1 и 90 для детали 2). В столбце (5) проставляются величины первого выхода из строя, соответствующие сроку службы всего изделия. По этим величинам можно построить кривую срока службы изделия.

Таблица 7.6

Данные для построения кривой срока службы изделия

Деталь 1		Деталь 2		Срок службы изделия — минимальная величина из (2) и (4)	Деталь 1		Деталь 2		Срок службы изделия — минимальная величина из (2) и (4)
Случайные нормальные числа	Срок службы детали 100 ± 20 (1)	Случайные нормальные числа	Срок службы детали 90 ± 10 (3)		Случайные нормальные числа	Срок службы детали 100 ± 20 (1)	Случайные нормальные числа	Срок службы детали 90 ± 10 (3)	
(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
0,464	109,28	0,137	91,37	91,37	0,906	118,12	-0,513	84,87	84,87
0,060	101,20	-2,526	64,74	64,74	1,179	123,58	-1,055	79,45	79,45
1,486	129,72	-0,354	86,46	86,46	-1,501	69,98	-0,488	85,12	69,98
1,022	120,44	-0,472	85,28	85,28	-0,690	86,20	0,756	97,56	86,20
1,394	127,88	-0,555	84,45	84,45	1,372	127,44	0,225	92,25	92,25

На практике для построения кривой нужно иметь больше 10 значений. Чем больше требуемая точность построения кривой, тем больше объем выборки. Подробнее связь точности с объемом выборки освещена в части IX. Здесь мы только приводим пример. Кроме того, следует отметить, что в большинстве случаев использование метода Монте-Карло невозможно без взаимодействующих электронных машин.

Предположим, что после получения данных о сроке службы двух деталей оказалось, что невозможно получить хорошую кривую для всего изделия. Метод Монте-Карло можно применить непосредственно к опытным

данным, не прибегая к кривой срока службы. Каждый опыт для каждой детали можно последовательно пронумеровать. Затем, используя таблицу случайных чисел (например, таблицу 7.3), можно выбрать срок службы для каждой детали, соответствующий опытному данным. Затем эти числа можно использовать в столбцах (2) и (4) таблицы 7.6 вместо чисел, полученных из таблицы случайных нормальных чисел.

После того как математические выражения, входящие в модель, оценены с помощью метода Монте-Карло, можно найти решение модели или путем итераций, или с помощью анализа. В моделях очередей и замены (части VI и VII) приводятся примеры совместного использования метода Монте-Карло с аналитическим и численным решением моделей.

И наконец, необходимо отметить, что в некоторых выражениях все величины таковы, что можно (или необходимо) использовать метод Монте-Карло для их оценки. В таких случаях и само выражение и его составляющие оцениваются с помощью метода Монте-Карло. Более подробно об этом методе см. [7, 10, 11, 13, 20 и 23].

МОДЕЛИ ДЛЯ РЕШЕНИЯ НАИБОЛЕЕ РАСПРОСТРАНЕННЫХ ЗАДАЧ

В промышленности и государственном управлении можно выделить некоторые наиболее распространенные задачи. Для создания моделей этих задач была проделана значительная работа. И хотя реальные задачи обычно отличаются от идеализированных моделей, однако это отличие не слишком велико, и гораздо проще решать новые задачи, используя существующие модели, чем создавать новые модели. Несколько глав книги (от части IV до части VIII) посвящено описанию таких моделей и их анализу. Знакомство с такими моделями и умение пользоваться ими и видоизменять их в определенных ситуациях является необходимым для всякого операциониста. В этой главе мы кратко перечислим классы задач, для которых были разработаны соответствующие модели. В каждой задаче можно отметить «уравновешивание» по крайней мере двух конфликтных ситуаций. Такое уравновешивание или балансировка характерны для всех моделей исследования операций. Рассматриваются пять типов задач: управления запасами, распределения, массового обслуживания, замены и задачи состязания.

Задачи управления запасами. В задачах управления запасами можно выделить некоторые общие особенности. Прежде всего, очевидно, что с увеличением запасов увеличивается стоимость хранения, но потери из-за нехватки товаров уменьшаются. Следовательно, одна из задач управления запасами заключается в определении такого уровня запасов, чтобы минимизировать сумму ожидаемых затрат по хранению товаров и потерь из-за нехватки товаров. Однако во многих случаях, например, как в описанном выше в этой главе, стоимость хранения и затраты производства не являются независимыми величинами, и их следует рассматривать совместно. Чем больше партия, тем меньше производственные затраты, поскольку расходы на подготовительно-заключительные работы в расчете на единицу продукции уменьшаются. Но с увеличением размера партии увеличивается стоимость хранения запасов. С другой стороны, чем меньше размер партии, тем меньше стоимость хранения, но стоимость производства единицы продукции увеличивается.

В модели производства и хранения общие производственные затраты выражаются через затраты на подготовку производства, обработку материалов, расходы на хранение и потери от нехватки товаров. Кроме того,

часто следует учитывать потери, связанные с изменением уровня производства (т. е. затраты в связи с наймом и увольнением). Модели хранения покупных товаров отличаются в основном тем, что стоимость материалов и стоимость обработки заменяются покупной ценой, на которую влияют различные скидки, т. е. эта цена изменяется дискретно.

Задачи управления запасами делятся на три класса.

1. Интервалы времени производства или закупки являются фиксированными, следует определить количество производимой или закупаемой продукции.

2. Количество произведенной или закупленной продукции является фиксированным, следует определить время производственного или закупочного циклов.

3. Требуется определить и время производственного или закупочного циклов и количество продукции. В тех случаях, когда спрос постоянен (см. пример главы 2), эти задачи становятся идентичными. Они различны только при изменяющемся спросе.

Символические модели построены для различных систем производства-хранения и закупки-хранения, в том числе для систем с известным спросом или спросом, изменяющимся дискретным или непрерывным образом, а также для систем, в которых время, потребное для получения товаров (после размещения заказов), или приравнивается нулю, или настолько значительно, что его необходимо учесть. Эти модели рассматриваются в части IV.

Задачи распределения. Задачи распределения возникают в тех случаях, когда существует определенный набор операций (работ или задач), которые надо выполнить, а наличных ресурсов (средств или возможностей) для выполнения каждой работы наилучшим образом не хватает *. Эти задачи можно разделить на три типа.

1. И работы и ресурсы заданы. Требуется так распределить ресурсы по работам, чтобы максимизировать некоторую меру эффективности (т. е. ожидаемую прибыль) или минимизировать некоторую меру отрицательной эффективности (т. е. ожидаемые затраты или время).

2. Заданы только наличные ресурсы. Требуется определить состав работ, которые можно выполнить с учетом наличных ресурсов при условии, что будет максимизирована некоторая мера эффективности или минимизирована некоторая мера отрицательной эффективности.

3. Заданы только потребные работы. Требуется определить, какие ресурсы следует использовать для проведения данных работ, чтобы максимизировать некоторую меру эффективности или минимизировать некоторую меру отрицательной эффективности.

Можно привести следующие примеры этих трех типов задач.

1. Известно производственное задание. Известны производственные мощности предприятия. Существуют различные способы получения каждого из потребных продуктов. Ограничения по мощности не позволяют производить каждый из продуктов наилучшим способом. Какой способ производства выбрать для различных продуктов, чтобы выполнить задание с минимальными затратами?

2. Рассматривается предприятие с определенными производственными мощностями. Какую продукцию следует производить, чтобы получить максимальный доход?

* Задачи распределения невозможно четко сформулировать, не вводя определения для понятия «ресурсы» и не давая их классификации, что, к сожалению, авторами не сделано. (*Прим. ред.*)

3. Известно месячное расписание полетов на одной из воздушных линий. Какое количество экипажей необходимо, чтобы минимизировать расходы по эксплуатации?

Для решения задач распределения используются методы линейного программирования и другие сходные методы. Эти методы и модели рассматриваются в главах 11, 12 и 13.

Задачи массового обслуживания (задачи ожидания в очередях). Можно привести много примеров явлений с ожиданием в очередях, например: вкладчики в банке, посетители кафе, самолеты, прибывающие на аэродром, и изделия в процессе их производства. В этих явлениях можно отметить некоторые общие для них особенности. Элементы, ожидающие обработки или обслуживания, поступают в определенном порядке. Возможен вариант, когда происходит скопление элементов, так что возникает очередь, неизбежно ожидание. А такое ожидание влечет за собой затраты за счет пролеживания деталей в процессе производства, задержки в погрузке, раздражение клиентов и т. д. В то же время элементы, поступающие на обработку или обслуживание, могут прибывать с большими интервалами, так что неизбежен простой оборудования или обслуживающего персонала, что влечет за собой дополнительные расходы.

Можно сформулировать два различных класса задач: (1) поступление является случайным и неуправляемым, требуется определить оптимальный объем оборудования (возможностей обслуживания); (2) известен объем оборудования (возможностей обслуживания), требуется определить оптимальный график работы наличного оборудования. Для решения первого класса задач используется теория массового обслуживания (см. главы 14 и 15). Задачи второго типа относятся к задачам «балансировки линий сборки» или определения «загрузки линий». Для решения этих задач используются комбинаторные методы. В главе 16 приводятся соответствующие модели, для которых даются точные и приближенные решения.

Сложные задачи, включающие управление запасами, распределение и очереди. Ко многим практическим задачам можно применить несколько моделей. Например, полный цикл производства и управления включает все три типа рассмотренных выше задач.

В некоторых случаях для управления производством требуется решать следующие вопросы.

1. Сколько изделий каждого типа необходимо выпустить? Другими словами, требуется запланировать партию оптимального размера. Эту задачу можно рассматривать как задачу производственного планирования, для которой применимы различные модели управления запасами и определения оптимального размера партии.

2. Как распределить выполнение производственного заказа по видам оборудования? Или из каких материалов изготовить данные изделия? Это распределительная задача, которая решается с помощью моделей распределения или программирования.

3. В какой последовательности и когда следует начать изготовление производственной партии? Это задача календарного планирования, для которой применяются модели очередей или модели массового обслуживания.

Ответы на эти вопросы нельзя дать независимо один от другого. Например, мы можем сначала получить оптимальный ответ на задачу производственного планирования и затем в зависимости от этого оптимума получить

наилучшее распределение; затем с учетом этого распределения мы можем получить оптимальный календарный план. Однако окончательный результат может отличаться от оптимального решения задачи в целом, т. е. последовательное решение этих трех задач не всегда приводит к оптимальному решению всей задачи. Возможно, что полученный в результате план нереализуем. Например, нельзя произвести все продукты в оптимальных количествах, потому что потребные для этого мощности превосходят наличный резерв. Или полученная программа не может быть реализована в заданное время. В настоящее время не существует метода для одновременного получения трех оптимальных решений. Поэтому нам приходится или воспользоваться приемлемым приближением к общему оптимальному решению, или проводить последовательную оптимизацию. При последовательной оптимизации мы пересматриваем каждое полученное оптимальное решение после получения нового очередного решения и производим необходимые корректировки. В результате можно получить решение, достаточно близкое к оптимальному.

Модели замены и ремонта. В сущности, процессы замены и ремонта не отличаются друг от друга. Разница заключается в самом объекте. Например, мы можем считать, что замена шин на грузовике необходима для правильной эксплуатации грузовика, а замена старого грузовика новым является составным элементом ремонта машинного парка. Таким образом, ремонт является процессом замены элементов. В дальнейшем используется только термин «замена», поэтому его следует пояснить более подробно, чтобы он отражал также понятие ремонта.

Задачи замены можно разделить на два класса: задачи с заменой элементов, эффективность которых со временем уменьшается (т. е. элементы со старением), и задачи, в которых элементы имеют приблизительно постоянную эффективность, а затем сразу выходят из строя. Револьверные станки можно считать примером элементов первого типа, а электрические лампы являются примером элементов второго типа.

Решая задачи первого типа, необходимо сравнивать дополнительные затраты на установку нового оборудования и прибыль за счет увеличения эффективности этого оборудования по сравнению со старым. Такое сравнение дает различные результаты в зависимости от эффективности старого оборудования (и, следовательно, обычно от его возраста) и от качеств нового оборудования.

Задачи второго типа обычно отличаются следующими особенностями. Оборудование или элементы выходят из строя в разные моменты времени. Каждый элемент можно заменить или отремонтировать после его выхода из строя, что связано с определенными расходами. С другой стороны, можно, не дожидаясь выхода элементов из строя, производить их замену или профилактический ремонт. Если же учесть не один элемент, а группу элементов, то получим, что стоимость единичного ремонта обычно падает с увеличением числа элементов, но общее число неисправностей возрастает. Поэтому, если учитывать группу элементов, то оказывается, что существует много возможных вариантов ремонта (единичный ремонт после каждой неисправности сравнивается с групповым профилактическим ремонтом) по времени проведения группового ремонта. Каждому времени можно поставить в соответствие ожидаемое число неисправностей. Требуется определить время для проведения группового ремонта, чтобы сумма ожидаемых расходов по всем видам ремонта была минимальной.

Модели замены и их применение рассматриваются в главе 17.

Задачи состязания (конкуренции). Под состязанием в общем случае понимаем ситуацию, в которой (1) две партии или группы конфликтуют между собой относительно определенной цели или нескольких целей и (2) эти партии или группы сотрудничают относительно определенной общей цели (или нескольких целей) или в борьбе с общим конкурентом. Примером такой ситуации является игра двух лиц в шахматы. Противники конфликтуют между собой в смысле выигрыша. Так, увеличение вероятности выигрыша игрока *A* неизбежно вызывает уменьшение вероятности выигрыша игрока *B*. Но в то же время можно считать, что *A* и *B* сотрудничают друг с другом, если их целью считать развлечение. В промышленности конкуренты могут захватывать друг у друга рынки сбыта, но они могут совместно выступать в интересах потребителей, снижая цену и улучшая качество товаров. Если же потребителя не обслуживают должным образом, то это будет уже конфликт, но не конкуренция. Конкуренты могут объединиться и образовать кооперацию, отношения которой складываются на основе сотрудничества, а не конкуренции.

Математические модели таких идеализированных ситуаций разрабатываются в теории игр, так же как в теории информации строятся свои идеальные модели (см. главу 4), но эти модели не находят широкого практического применения в промышленности. И хотя в некоторых случаях эти модели используются, но при этом соблюдается строгая коммерческая тайна и подробности применения (а иногда даже и общие положения) не публикуются. И все же идеи теории игр можно уже в настоящее время использовать для решения практических задач гораздо шире, чем математический аппарат этой теории.

Правда, большинство игр, рассматриваемых в теории игр, намного проще в сравнении с реальными проблемами промышленной конкуренции. Во многих играх можно точно определить правила игры, перечислить возможные ходы или последовательность ходов, а также результаты каждого хода. Используя эту информацию, а также сведения о возможных ходах противника, можно разработать стратегию поведения игрока в игре. Фон Нейман и Моргенштерн [18] показали, как выбрать оптимальную стратегию для некоторых игр.

Однако многие промышленные и военные задачи очень трудно формализовать из-за их сложности и отсутствия соответствующей информации о последствиях различных действий и о поведении противника. Но даже в этих случаях можно с успехом использовать многие положения теории игр. Например, можно определить оптимальную возможную стратегию противника. Затем выбираем такую стратегию, которая обеспечивает выигрыш даже при оптимальной стратегии противника. Тем самым мы гарантируем по крайней мере этот минимальный выигрыш. Очевидно, что этот подход будет тем удачнее, чем точнее мы сумеем определить оптимальную вероятную стратегию противника. Во многих военных задачах этот метод был использован с большим успехом.

Примером состязания являются аукцион и различные ситуации предложения цены при заключении контрактов, получении концессий, лицензий и др. В таких ситуациях основной конфликт заключается в том, что при увеличении цены шансы на выигрыш увеличиваются, но ожидаемые доходы (от выигрыша) уменьшаются. С другой стороны, при уменьшении цены шансы на выигрыш уменьшаются, но доходы увеличиваются. Задача заключается в том, чтобы наилучшим образом разрешить этот конфликт. Для этой цели разработаны модели торгов.

В главе 18 описывается теория игр, а в главе 19 — модели торгов.

ЗАКЛЮЧЕНИЕ

В этой главе были описаны различные способы использования модели для оценки возможных изменений реального процесса. Однако неправильно думать, что это отношение (модель как средство оценки процесса) является односторонним. На практике всякая информация об изучаемом процессе используется для оценки модели. Например, для прогнозирования затрат используется определенная модель. Однако на практике оказывается, что фактические расходы отличаются от прогнозируемых. Это указывает на неполноту модели и дает материал для ее переоценки. Следовательно, модель и процесс можно использовать для взаимной оценки. Естественно, что чем больше мы узнаем о процессе, тем больше изменений вводится в модель. Это справедливо для моделей, используемых в различных областях науки. Ни одна модель не является совершенной, поскольку наши знания о различных явлениях всегда остаются неполными. Вопросы проверки и испытания моделей рассмотрены в главах 20 и 21.

Вечным вопросом всех исследователей остается вопрос о том, как начинать исследования. Некоторые полагают, что необходимо сначала собрать факты, прежде чем строить какую-либо модель, чтобы исследователь не рассматривал реальный процесс с точки зрения модели и чтобы не было предвзятого объяснения фактов. Другие же считают, что мы вообще до построения модели не знаем, какие данные следует собирать. Существуют и более сложные представления о процессе построения модели, предполагающие непрерывную связь процесса сбора и обновления данных с процессом построения и корректировки модели. Более подробно эти вопросы освещены в других источниках [5, 6].

Следует также еще раз подчеркнуть, что решение, полученное с помощью модели, не обязательно является решением задачи. Например, оптимальный размер партии можно определить из соответствующего выражения только в том случае, если точно известна, например, стоимость подготовки производства и если она остается постоянной. Но очевидно, что эта стоимость не остается неизменной в течение длительного времени. Поэтому необходимо предусмотреть регулярную проверку различных переменных модели, чтобы можно было получить новое решение по мере изменения данных. Вопросы проверки более подробно рассмотрены в главе 21.

Важность такого контроля трудно переоценить. Общеизвестно, что многие характеристики промышленных процессов и процессов управления сильно меняются за сравнительно небольшие периоды. Поэтому необходимо предусмотреть методы проверки правильности решения и его корректировки.

П р и м е ч а н и е

ПРИМЕР ИТЕРАТИВНОГО ПРОЦЕССА

В качестве примера итеративного процесса опишем метод Ньютона решения уравнений. Этот метод дает возможность определить с любой степенью точности корни алгебраического уравнения. Для небольшого участка гладкой кривой касательная мало отличается от самой кривой. На этом положении основан описываемый метод решения уравнений.

Метод Ньютона можно сформулировать следующим образом. Требуется найти решение уравнения $f(X) = 0$. Корнем этого уравнения является абсцисса точки пересечения оси X кривой $Y = f(X)$.

Выберем начальное решение, например X_0 (см. рис. 7.5). Эта величина X_0 определяет точку P на кривой с координатами (X_0, Y_0) . Проведем касательную к кривой в точке P , которая пересечет ось X в точке $(X_1, 0)$. Если мы считаем, что на участке (X_0, X_1) кривая и касательная почти совпадают, то можно принять X_1 за приближенный корень уравнения. Наклон касательной к кривой $f(X)$ в точке P задается в виде $f'(X_0)$. Следовательно, можно определить X_1 из соотношения

$$X_1 = X_0 - \frac{f(X_0)}{f'(X_0)}. \quad (10)$$

Этот процесс можно повторить для увеличения точности. В общем случае получаем соотношение для определения корня на $(n+1)$ -й итерации:

$$X_{n+1} = X_n - \frac{f(X_n)}{f'(X_n)}. \quad (11)$$

Будет ли процесс сходящимся и насколько быстро он сходится, зависит от функции $f(X)$ и начальной величины X_0 . Наилучшая сходимость будет в том случае, если величина $f(X_0)$ мала, а $f'(X_0)$ велика.

Рассмотрим конкретный пример. Задана функция

$$f(X) = X^3 + 3X^2 + 4X - 2.$$

Выберем произвольное решение $X_0 = 2^*$. Определим

$$f'(X) = 3X^2 - 6X + 4, \quad f(2) = 8 - 12 + 8 - 2 = 2, \quad f'(2) = 12 - 12 + 4 = 4.$$

Из уравнения (11) получим

$$X_1 = 2 - \frac{2}{4} = 1,5.$$

Продолжая этот процесс, получим:

$$f(1,5) = \left(\frac{3}{2}\right)^3 - 3\left(\frac{3}{2}\right)^2 + 4\left(\frac{3}{2}\right) - 2 = \frac{5}{8},$$

$$f'(1,5) = 3\left(\frac{3}{2}\right)^2 - 6\left(\frac{3}{2}\right) + 4 = \frac{7}{4},$$

$$X_2 = 1,5 - \frac{5}{14} = 1,143.$$

На следующей итерации получим

$$f(1,143) = 0,147, \quad f'(1,143) = 1,060,$$

$$X_3 = 1,143 - \frac{0,147}{1,060} = 1,004.$$

Мы можем продолжать вычисления дальше и определять на каждом шаге величину $f(X_i)$, чтобы определить, как быстро сходится решение (решение получается в точке X^* , в которой $f(X^*) = 0$), и, следовательно, можем получать решение с любой заданной точностью.

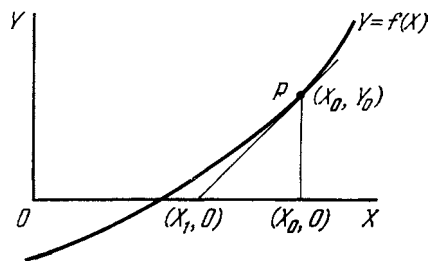


Рис. 7.5. Решение уравнения с помощью метода Ньютона.

* В нашем примере очевидно, что решением является $X = 1$. Величина отклонения от значения $X = 1$ определяет точность нашей аппроксимации.

ЛИТЕРАТУРА

1. Ackoff R. L., *The Design of Social Research*, University of Chicago Press, Chicago, 1953.
2. Alford L. P. and Bangs J. R., *Production Handbook*, Ronald Press, New York, pp. 99—106, 1944.
3. Arrow K. J., *Mathematical Models in the Social Sciences*, in Daniel Lerner and H. D. Laswell (eds.), *The Policy Sciences*, Stanford University Press, Stanford, 1951.
4. Avery F. B., *Economic Manufacturing Quantity*, *Industr. Mgmt.*, 63, no. 3, 169—170, 189 (Mar. 1922).
5. Churchman C. West, *Theory of Experimental Inference*, The Macmillan Co., New York, 1948.
6. Churchman C. West and Ackoff R. L., *Methods of Inquiry*, Educational Publishers, St. Louis, 1950.
7. Curtiss J. H., *Sampling Methods Applied to Differential and Difference Equations*, in *Seminar on Scientific Computation*, International Business Machines Corp., New York.
8. Dewey John, *Logic: The Theory of Inquiry*, Henry Holt & Co., New York, 1938.
9. Grant Eugene L., *Principles of Engineering Economy*, Ronald Press, New York, pp. 263—268, 272—273, 1938.
10. Kahn H., *Applications of Monte Carlo*, Project RAND, RM-1237-AEC, Santa Monica, Apr. 19, 1954.
11. Kahn H. and Marshall A. W., *Methods of Reducing Sample Size in Monte Carlo Computations*, *J. Opns. Res. Soc. Am.*, 1, no. 5, 263—278 (Nov. 1953).
12. Kimball D. A., *Industrial Economics*, McGraw-Hill Book Co, New York, pp. 283—287, 1929.
13. King Gilbert W., *The Monte Carlo Method as a Natural Mode of Expression in Operations Research*, *J. Opns. Res. Soc. Am.*, 1, no. 2, 46—51 (Feb. 1953).
14. Koepke C. A., *Plant Production Control*, John Wiley & Sons, New York, pp. 379—387, 1941.
15. Lehoczy P. N., *Lower Costs by Economic Lot Sizes*, *Mfg. Inds.*, 16, no. 4, 299—300 (Aug. 1928).
16. Littlefield P. H., *The Determination of the Economic Size of Production Orders* Massachusetts Institute of Technology, Cambridge, Course XV, Thesis No. 3, 1924.
17. Mellen G. H., *Practical Lot Quantity Formula*, *Mgmt. & Adm.*, 9, no. 6, 565—566 (June 1925), and 10, no. 3, 155 (Sept. 1925).
18. Neumann J. von and Morgenstern O., *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 2nd ed., 1947.
19. Raymond F. E., *Quantity and Economy in Manufacture*, McGraw-Hill Book Co., New York, 1931.
20. Rich R. P., *Simulation as an Aid to Model Building*, *J. Opns. Res. Soc. Am.*, 3, no. 1, 15—19 (Feb. 1955).
21. The RAND Corporation, *A Million Random Digits*, The Free Press, Glencoe, 1955.
22. Thrall R. M., Coombs C. H. and Davis R. L., eds., *Decision Processes*, John Wiley & Sons, New York, 1954.
23. U.S. Dept. of Commerce, National Bureau of Standards, *Monte Carlo Method*, Applied Mathematics Seminar 12, June 11, 1951.
24. Wilson E. Bright, Jr., *An Introduction to Scientific Research*, McGraw-Hill Book Co., New York, 1952.
25. Younger J. and Gesechelin J., *Work Routing, Scheduling, and Dispatching in Production*, Ronald Press, New York, 3rd ed., pp. 52—57, 1947.

Задачам управления запасами всегда было посвящено в теории исследования операций гораздо больше работ, чем любой другой проблеме экономики. Сейчас для этих задач все более расширяется и круг военных применений. Вот почему подавляющее большинство существующих моделей разработано именно для задач управления запасами.

Еще в 1915 году Харрисом было получено уравнение для определения при известном спросе оптимального размера партии, минимизирующего сумму затрат, связанных с заказом и хранением запасов. Совместными усилиями инженеров, экономистов и математиков к 1950 году список работ, посвященных этому вопросу, значительно расширился. Работавший в этой области Уайтн [33] сделал обзор всех результатов, полученных до 1952 года.

Однако большая часть методов и приемов, которыми сейчас постоянно пользуются в теории исследования операций при решении задач, связанных с управлением производством и запасами, появилась лишь в самые последние годы *. Эти исследования начались с попыток обобщить существовавшие методы и на те случаи, когда спрос неизвестен и его можно лишь приближенно оценить.

Одна из задач, которая возникает при учете неопределенности спроса, связана с созданием резервов для предотвращения нехватки запасов. Важные результаты при исследовании этой задачи получили Фрай [15], Эйзенхарт [13], Эрроу, Харрис и Маршак [2], Томпкинс [28, 29] и другие. Уайтн [33] исследовал, какое влияние оказывает необходимость создания запасов на размер партии. Дворецкий, Кифер и Вольфовиц [11, 12] определили условия, при которых можно определить оптимальный уровень запасов.

В своей книге Уайтн так оценивает эти исследования: «Работа Эрроу, Харриса и Маршака значительно расширила ранее полученные результаты. Через год были опубликованы работы Дворецкого, Кифера и Вольфовица. Это были, безусловно, во всех отношениях выдающиеся исследования — и с точки зрения изящности и общности изложения и по мощности примененных авторами математических методов. Дальнейшее обобщение этих работ позволило учитывать случайный характер сроков поставок с помощью вероятностного распределения, одновременное поступление нескольких заказов, зависимость спроса от времени и рассматривать также случаи неполностью известного вероятностного распределения спроса» (стр. 35—36).

Модели, которые описываются в главе 8, не являются столь универсальными и годятся лишь для вполне определенных задач управления запасами. На этих моделях виден постепенный переход от простейших случаев к более сложным. Это делается для того, чтобы главное внимание обратить на сам метод конструирования моделей и показать, помимо всего прочего, фундаментальную зависимость оптимальной процедуры планирования от отношения стоимости складирования к штрафу при нехватке запасов. Математические выкладки, относящиеся к этим моделям, приводятся в приложении в конце главы 8.

Берман и Кларк [8] разработали другой класс весьма распространенных моделей для иерархических (многоступенчатых, многобункерных.— *Прим. перев.*) систем

* Популярное изложение этих работ содержится в статье Ледермана, Литтауэра и Вайсса [20].

с запасами. Такие модели применяются, например, если несколько периферийных складов снабжается через один центральный склад.

Зависимость размера заказа от скидки за количество была исследована Уайтном [33]. Обобщенный метод, позволяющий учитывать многоступенчатую скидку за количество при наличии ограничений, излагается на примерах в главе 9 настоящей книги.

В главе 10 при рассмотрении моделей запасов учитываются различные ограничения, связанные с лимитированием производственных мощностей, времени или денежных средств. Поясняется, как следует учитывать эти ограничения при составлении моделей запасов и как получить при этом оптимальные решающие правила. Поскольку методы, которые используются при составлении моделей с подобными ограничениями, носят довольно частный характер, все математические выводы приводятся непосредственно в тексте этой главы. Результаты, которые приводятся в главах 8—10, относятся к так называемым статическим проблемам управления запасами. Однако имеется ряд работ, посвященных динамическим задачам, в которых мы сталкиваемся с тем, что принимаемые в данный момент решения влияют и на некоторый последующий период времени. Имеется несколько различных подходов к динамическим задачам.

Первый подход непосредственно использует основную идею следящей системы — принцип обратной связи. По существу, этот подход заключается в применении одной из форм обратной связи для регулирования производства или закупок в зависимости от изменяющегося спроса. Одна из возможных процедур регулирования, основанная на разработанных Винером автокорреляционных методах, была предложена и реализована в Технологическом институте Карнеги [19, 26].

Сходный метод, принадлежащий Вашану [30], при некоторых условиях обеспечивает минимизацию дисперсии отклонения запасов от нормы. Этот метод отличается от предыдущего рассмотрением дискретных распределений для спроса и запасов. В двух проектах по исследованию операций на промышленных объектах, выполненных в Кейсовском технологическом институте, в качестве устройства, играющего роль обратной связи, взята карта статистического контроля.

При динамическом подходе к проблеме управления запасами учитываются и расходы, которые возникают при изменении уровня производства. Существуют методы, позволяющие определить такой уровень производства, при котором сумма издержек, связанных со складированием, заказом, уплатой штрафов при нехватке, вместе с упомянутыми выше расходами при изменении уровня производства, оказывается минимальной. Однако пока еще отсутствуют способы, позволяющие одновременно оптимизировать и объем производства и размер заказов на отдельные виды продукции.

Чарнс, Купер и Фэрр [9] применили для определения уровней производства в случае значительных колебаний сезонного спроса метод линейного программирования (V часть настоящей книги). При этом спрос предполагался известным.

Разработав свой новый метод — динамическое программирование, Беллман [3—5, 7] сделал возможным как теоретически, так и вычислительно исследование задач управления производством и запасами с позиций вариационного исчисления. Беллман, Гликсберг и Гросс применили этот метод при выводе оптимальной стратегии в задаче управления запасами при различных предположениях относительно условий задачи [6]. В Технологическом институте Карнеги был разработан матричный метод — квадратичное программирование [18], который применим при определении уровней производства, когда функции стоимости представляют собою квадратичные формы.

Обзор последних результатов в теории управления производством и запасами содержится в работах Уайтна [32] и Симона и Хольта [27].

Прежде чем обратиться к моделям запасов и их исследованию, следует еще раз подчеркнуть, что скорее всего ни одна из описываемых ниже и вообще любая готовая модель не будет *как таковая* пригодной в какой-то конкретной ситуации. Однако исследователь, вероятно, сможет сам модифицировать готовые модели, если он разобрался, как они получаются, т. е. если он владеет методологией конструирования моделей.

Глава 8

ПРОСТЕЙШИЕ МОДЕЛИ УПРАВЛЕНИЯ ЗАПАСАМИ

ВВЕДЕНИЕ

В этой главе предполагается познакомить читателя с аналитическими методами, позволяющими строить математические модели управления запасами. Мы рассмотрим ряд специфических моделей, применимых в некоторых относительно простых задачах управления запасами, не делая здесь попыток разработать универсальную модель, пригодную во всех случаях. Ради практичности моделей, получаемых в результате такого подхода, приходится жертвовать их универсальностью. Но именно вопросы применения нас и будут интересовать в первую очередь.

Основной класс рассматриваемых здесь задач связан с оптимальным регулированием запасов. Эти задачи можно сформулировать следующим образом:

1. Моменты времени, в которые принимаются заказы на пополнение запасов, фиксированы. Остается определить объем заказов.

2. Необходимо определить и объем и время заказов.

Задача исследования состоит в отыскании оптимального решения этих задач. Под оптимальным здесь понимается решение, минимизирующее сумму всех расходов, связанных с созданием запасов. Эти расходы бывают трех типов:

1. Расходы, вызываемые оформлением и получением заказа при закупке или производстве. Это величина, не зависящая от размера партии и, следовательно, переменная для единицы продукции.

2. Стоимость хранения единицы продукции на складе. Сюда включаются затраты, связанные с организацией хранения, устареванием и порчей, расходы на страхование и налоги.

3. Расходы (штрафы), возникающие при истощении запасов, когда происходит задержка в обслуживании или спрос вообще невозможно удовлетворить.

Все затраты могут оставаться постоянными или изменяться как функции времени (например, в зависимости от сезона может быть различным штраф за задержку в обслуживании). Затраты могут зависеть также от объема запасов (размером партии может, например, определяться стоимость хранения единицы товара на складе).

В задачах управления запасами учитываются также характеристики спроса и возможности пополнения запасов.

Спрос может быть известным или неизвестным, постоянным или зависящим от времени. Величина, характеризующая спрос, может быть как дискретной (например, количество автомобилей), так и непрерывной.

Спрос на запасенные товары может возникать в определенные моменты времени (спрос на мороженое на стадионе) или существовать постоянно (спрос на мороженое в большом аэропорту).

Заказы на пополнение запасов в ряде случаев могут выполняться немедленно (например, при заказе молока в небольшом магазине). В других случаях выполнение заказа требует значительного времени. Заказы можно делать в любые или только в определенные моменты времени.

Объем поступающей на склад продукции может измеряться дискретной или непрерывной величиной и может быть как постоянным, так и переменным. И наконец, само поступление может быть дискретным или непрерывным и происходить равномерно или неравномерно.

Возможны и другие варианты, но даже перечисленные случаи дают нам при управлении запасами тысячи различных задач. Чтобы в них легче ориентироваться, перечисленные выше характеристики сведены в таблицу 8.1.

Таблица 8.1

Классификация задач управления запасами

1. Стоимость заказа при покупке или производстве		
2. Стоимость хранения единицы продукции		а. Постоянные б. Переменные
3. Штраф при нехватке		
4. Спрос	А. Известный Б. На основе опытных данных	а. Постоянный б. Переменный
5. Величина спроса	А. Дискретная Б. Непрерывная	
6. Распределение спроса по времени	А. Непрерывное Б. Дискретное	а. Равномерное б. Неравномерное
7. Время задержки поставки	А. Практически нулевое Б. Положительное	
8. Время выполнения поставки	А. Известное Б. На основе опытных данных	а. Постоянное б. Переменное
9. Заказ на пополнение	А. Дискретный Б. Непрерывный	а. Постоянный б. Переменный
10. Распределение заказов на пополнение запасов по времени	А. Непрерывное Б. Прерывное	а. Равномерное б. Неравномерное

Мы рассмотрим ниже модели лишь для очень ограниченного числа задач. Однако знакомство с методом их решения облегчит читателю поиски решения и в других случаях.

В главе VI приняты следующие обозначения:

q — объем заказа (при пополнении запасов);

q_i — объем заказа, производимого в начале i -го интервала;

q_0 — оптимальный размер заказа;

r — спрос за некоторый интервал времени;

r_i — спрос за i -й интервал времени;

S_i — уровень запасов к началу i -го интервала;

s_i — уровень запасов к концу i -го интервала. Отметим, что $s_i = S_i - r_i$ и $S_i = s_{i-1} + q_i$.

S_0 — оптимальный уровень запасов к началу некоторого интервала времени;

t — интервал времени;

t_s — интервал времени между двумя заказами;

t_{s0} — оптимальный интервал времени между заказами;

T — период времени, для которого ищется оптимальная стратегия;

R — полный спрос за время T ;

C_1 — стоимость хранения единицы продукции в единицу времени;

C_2 — величина штрафа за нехватку одной единицы продукции (в определенный момент времени);

C_s — стоимость заказа (при покупке или производстве);

Q — ожидаемые суммарные накладные расходы*. В этой главе они иногда называются ожидаемыми суммарными расходами, но в действительности, так как размер заказов при пополнении запасов не влияет на продажную цену каждого предмета, которая поэтому не учитывается в моделях I — VI, мы всегда имеем в виду ожидаемые суммарные *накладные* расходы.

Q_0 — минимум ожидаемых суммарных накладных расходов;

$P(r)$ — вероятность спроса на r единиц товара (r — целое число);

$f(r)$ — плотность вероятности величины r , где r непрерывно;

$P(r \leq S)$ — вероятность того, что спрос не превысит уровень S (r — целое число);

$F(r)$ — функция распределения величины r , где r непрерывно;

$F(S) = \int_0^S f(r) dr$ — вероятность того, что спрос не превысит уровень S (r — непрерывная величина).

МОДЕЛЬ I

Пусть некий предприниматель должен поставлять своим клиентам R изделий равномерно в течение интервала времени T . Таким образом, спрос фиксирован и известен. Нехватка товара не допускается, т. е. штраф при неудовлетворенном спросе бесконечно велик ($C_2 = \infty$). Переменные затраты производства складываются из следующих элементов: C_1 — стоимость хранения одного изделия (в единицу времени), C_2 — стоимость запуска в производство одной партии изделий.

Предприниматель должен решить, как часто ему следует организовывать выпуск партий и каким должен быть размер каждой партии.

Уравнение цен и его аналитическое решение. Только что описанная ситуация представлена графически на рис. 8.1. Пусть q — размер партии, t_s — интервал времени между запусками в производство партий, а R — полный спрос за все время планирования T .

Тогда R/q — число партий за время T и

$$t_s = \frac{T}{R/q} = \frac{Tq}{R}.$$

Если интервал t_s начинается, когда на складе имеется q изделий, и заканчивается при отсутствии запасов, тогда $q/2$ — средний запас в течение t_s^{**} , $\frac{q}{2} C_1 t_s$ — затраты на хранение в интервале t_s .

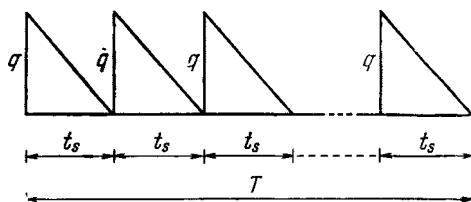


Рис. 8.1. Кривая запасов. Модель I.

* Эта величина в американской литературе обозначается обычно *TEC* (total expected cost). (Прим. перев.)

** Равенство $q/2 = q_{\text{ср}}$ следует рассматривать как приближенное. Точность его тем выше, чем больше R . Авторы и дальше часто не делают различия между дискретными и непрерывными величинами. (Прим. перев.)

Общая стоимость создания запасов в интервале t_s равна сумме стоимости хранения и стоимости запуска в производство

$$\frac{q}{2} C_1 t_s + C_s.$$

Для вычисления полной стоимости создания запасов за время T следует эту величину умножить на общее число партий за это время:

$$Q = \left(\frac{q}{2} C_1 t_s + C_s \right) \frac{R}{q}.$$

Подставляя сюда выражение для t_s , получаем

$$Q = \left(\frac{q}{2} C_1 \frac{Tq}{R} + C_s \right) \frac{R}{q},$$

или

$$Q = \frac{C_1 T q}{2} + \frac{C_s R}{q}. \quad (1)$$

Члены в правой части уравнения (1) представляют собой полную стоимость хранения и полную стоимость заказа в производстве всех партий. С увеличением размера партий первый член возрастает, а второй убывает. Решение задачи управления запасами и состоит в определении такого размера партии q_0 , при котором суммарная стоимость была бы наименьшей (рис. 8.2).

Решение этой задачи может быть получено аналитически. Оно приведено в приложении 1 в конце настоящей главы.

Найденное оптимальное значение q_0 размера партии

$$q_0 = \sqrt{2 \frac{RC_s}{TC_1}}. \quad (2)$$

Для оптимальных t_{s0} и Q_0 имеем

$$t_{s0} = \sqrt{2 \frac{T}{R} \frac{C_s}{C_1}}, \quad (3)$$

$$Q_0 = \sqrt{2RTC_1C_s}. \quad (4)$$

Пример I. Пусть предприниматель

должен поставлять своему заказчику 24 000 единиц продукции в год. Так как получаемая продукция исполь-

зуется непосредственно на сборочной линии и заказчик не имеет для нее специальных складов, поставщик должен ежедневно отгружать дневную норму. В случае нарушения поставок поставщик рискует потерять заказ. Поэтому нехватка продукции недопустима, т. е. штраф при нехватке можно считать бесконечным. Хранение единицы продукции в месяц стоит 0,1 долл. Стоимость запуска в производство одной партии продукции составляет 350 долл.

Требуется определить оптимальный размер партии q_0 , оптимальный период t_{s0} и вычислить минимум общих ожидаемых годовых затрат Q_0 . В данном случае $T = 12$ месяцев, $P = 24\,000$ единиц, $C_1 = 0,1$ долл./месяц, $C_s = 350$ долл./партия. Подстановка этих значений в уравнения (2), (3)

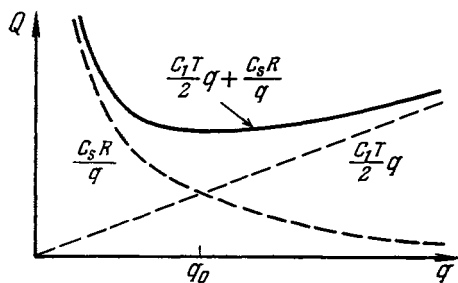


Рис. 8.2. Определение q_0 . Модель I.

и (4) дает нам

$$q_0 = \sqrt{2 \frac{24\,000 \cdot 350}{12 \cdot 0,1}} = 3740 \text{ единиц},$$

$$t_{s0} = \sqrt{2 \frac{12 \cdot 350}{24\,000 \cdot 0,1}} = 1,87 \text{ месяца} = 8,1 \text{ недели},$$

$$Q_0 = \sqrt{2 \cdot 24\,000 \cdot 12 \cdot 0,1 \cdot 350} = 4490 \text{ долл./год.}$$

МОДЕЛЬ II

Рассмотрим теперь случай, который отличается от предыдущего только тем, что превышение спроса над запасами уже допускается, т. е. штраф за нехватку конечный.

Уравнение цен и его аналитическое решение. Рассматриваемая ситуация изображена на рис. 8.3. В начале каждого интервала имеется уровень запасов. Из подобия треугольников находим

$$t_1 = \frac{S}{q} t_s, \quad t_2 = \frac{q-S}{q} t_s.$$

Средний запас в течение t_1 равен $S/2$ *. Поэтому затраты на хранение за все время t_1 составляют $\frac{S}{2} C_1 t_1$. Средняя нехватка (превышение спроса над уровнем запасов)

за время t_2 равна $(q-S)/2$, и штраф за время t_2 составляет $\frac{q-S}{2} C_2 t_2$.

Таким образом, ожидаемые суммарные расходы за все время T определяются следующим выражением:

$$Q(q, S) = \left(\frac{S}{2} C_1 t_1 + \frac{q-S}{2} C_2 t_2 + C_s \right) \frac{R}{q}.$$

Подставляя сюда найденные выше выражения для t_1 и t_2 и учитывая полученное в предыдущем разделе выражение для t_s , имеем

$$Q(q, S) = \frac{S^2 C_1 T}{2q} + \frac{(q-S)^2 C_2 T}{2q} + \frac{C_s R}{q}. \quad (5)$$

Из уравнения (5) можно найти оптимальные значения для q и S . Подробный вывод дается в приложении 2 в конце этой главы. Получаем

$$q_0 = \sqrt{2 \frac{R}{T} \frac{C_s}{C_1}} \sqrt{\frac{C_1 + C_2}{C_2}}, \quad (6)$$

$$S_0 = \sqrt{2 \frac{R}{T} \frac{C_s}{C_1}} \sqrt{\frac{C_2}{C_1 + C_2}}. \quad (7)$$

Этим значениям соответствуют

$$t_{s0} = \sqrt{2 \frac{R}{T} \frac{C_s}{C_1}} \sqrt{\frac{C_1 + C_2}{C_2}}, \quad (8)$$

$$Q_0 = \sqrt{2 R T C_1 C_s} \sqrt{\frac{C_2}{C_1 + C_2}}. \quad (9)$$

* См. сноску на стр. 165.

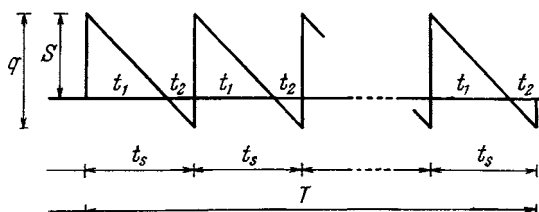


Рис. 8.3. Кривая запасов. Модель II.

При сравнении результатов, полученных для моделей I и II, можно заметить, что, во-первых, уравнения (2), (3) и (4) можно получить из уравнений (6), (8), (9), если в них устремить C_2 к бесконечности. Этот результат нельзя считать неожиданным, так как модель I есть частный случай модели II.

Во-вторых, если $C_2 \neq \infty$, то

$$\sqrt{2RTC_1C_s} \sqrt{\frac{C_2}{C_1+C_2}} < \sqrt{2RTC_1C_s}.$$

Следовательно, ожидаемые суммарные расходы в модели II меньше, чем в модели I.

Пример II. Пусть сохраняются все условия примера I, но только штраф C_2 за нехватку теперь равен 0,2 долл. за одно изделие в месяц. Из уравнений (6) — (9) получаем:

$$q_0 = \sqrt{2 \frac{24\,000 \cdot 350}{12 \cdot 0,1}} \sqrt{\frac{0,1+0,2}{0,2}} = 4578 \text{ единиц},$$

$$S_0 = \sqrt{2 \frac{24\,000 \cdot 350}{12 \cdot 0,1}} \sqrt{\frac{0,2}{0,1+0,2}} = 3056 \text{ единиц},$$

$$t_{s0} = \sqrt{2 \frac{12 \cdot 350}{24\,000 \cdot 0,1}} \sqrt{\frac{0,1+0,2}{0,2}} = 2,29 \text{ месяца} = 9,9 \text{ недели},$$

$$Q_0 = \sqrt{2 \cdot 24\,000 \cdot 12 \cdot 0,1 \cdot 350} \sqrt{\frac{0,2}{0,1+0,2}} = 3667 \text{ долл.}$$

При оптимальной стратегии ожидаемый дефицит к концу каждого периода составлял бы $4578 - 3056 = 1522$ изделия.

МОДЕЛЬ III

При построении этой модели штрафы, связанные с дефицитом запасов, так же как в модели II, считаются конечными. Кроме того, модель III имеет следующие особенности.

1. Спрос и пополнение запасов оцениваются на основе опытных данных.
2. Рассматривается производство и потребление дискретного продукта.
3. Распределения по времени спроса и заказов на пополнение дискретные и неравномерные.
4. Известно и постоянно время выполнения заказов*.

Пример III. Компания по производству электроэнергии собирается приобрести новый генератор для своей электростанции. Одна из основных деталей генератора весьма сложна и дорога, и целесообразно при заказе генератора заказать и несколько штук этих деталей в запас. Однако эта деталь подгоняется индивидуально для каждого генератора и ее уже нельзя будет использовать на другом агрегате.

Компания желает знать, сколько запасных частей ей следует заказывать для каждого генератора. При решении этого вопроса компания располагает следующей информацией. Стоимость одной детали, если ее заказывать вместе с генератором, составляет 500 долл. Отсутствие этой детали в запасе при поломке приводит к выходу генератора из строя, и простой

* В этой модели, так же как и в модели VI, расходы, связанные с запасными частями, учитываются только для неиспользованных деталей, т. е. в этой элементарной задаче управления запасами сравниваются расходы на приобретение запасных частей, которые могут оказаться лишними, и убытки, возникающие, если в нужный момент запасных частей не оказывается.

генератора и срочный заказ детали обходятся в 10 000 долл. Данные о частоте выхода этой детали из строя (по 100 генераторам) приведены в таблице 8.2.

Таблица 8.2

Потребовалось запасных деталей	0	1	2	3	4	5	6 и более
Число генераторов, в которых потребовалось указ. число деталей	90	5	2	1	1	1	0
Эмпирическая вероятность выхода из строя указ. числа деталей	0,90	0,05	0,02	0,01	0,01	0,01	0,00

Уравнение цен. Составим уравнение цен, т. е. математическую модель. Пусть в запасе имелось S деталей, а за все время эксплуатации генератора потребовалось r деталей. Тогда запасу в S деталей будут соответствовать следующие затраты: $(S - r) C_1$, если $r \leq S$, т. е. запас оказался чрезмерным, и $(r - S) C_2$, если $S > r$, т. е. запасных деталей не хватило.

Хотя нам неизвестно заранее, какое число деталей потребуется, но вероятности выхода из строя определенного числа деталей $P(r)$ мы знаем.

Чтобы вычислить ожидаемые при данном уровне запасов затраты, мы должны просуммировать значения расходов для каждого r , умноженные на соответствующие вероятности $P(r)$:

$$Q(S) = C_1 \sum_{r=0}^S P(r) (S-r) + C_2 \sum_{r=S+1}^{\infty} P(r) (r-S). \quad (10)$$

Численное решение. Применим уравнение (10) к примеру III и подсчитаем ожидаемый суммарный расход при различных разумных уровнях запасов (т. е. от 0 до 5, так как вероятность более пяти поломок равна нулю).

Будем в этом примере считать, что убыток от неиспользованных запасных деталей равен их покупной стоимости, так как расходы на хранение пренебрежимо малы. Таким образом, $C_1 = 50$ долл. Нехватка запасных деталей обходится в 10 000 долл. за одну штуку — такова плата за простой генератора и специальный заказ недостающей детали.

Произведя с помощью уравнения (10) расчеты для различных уровней запасов, получаем следующие значения для ожидаемых суммарных расходов:

$$Q(S=5) = 500 [0,90 (5-0) + 0,05 (5-1) + 0,02 (5-2) + 0,01 (5-3) + 0,01 (5-4) + 0,01 (5-5)] = \$2395,$$

$$Q(S=4) = 500 [0,90 (4-0) + 0,05 (4-1) + 0,02 (4-2) + 0,01 (4-3) + 0,01 (4-4)] + 10\,000 [0,01 (5-4)] = \$2000,$$

$$Q(S=3) = 500 [0,90 (3-0) + 0,05 (3-1) + 0,02 (3-2) + 0,01 (3-3)] + 10\,000 [0,01 (5-3) + 0,01 (4-3)] = \$1710,$$

$$Q(S=2) = 500 [0,90 (2-0) + 0,05 (2-1) + 0,02 (2-2)] + 10\,000 [0,01 (5-2) + 0,01 (4-2) + 0,01 (3-2)] = \$1525,$$

$$\begin{aligned}
 Q(S=1) &= 500 [0,90(1-0) + 0,05(1-1)] + \\
 &+ 10\,000 [0,01(5-1) + 0,01(4-1) + 0,01(3-1) + 0,02(2-1)] = \$1550, \\
 Q(S=0) &= 10\,000 [0,01(5-0) + 0,01(4-0) + 0,01(3-0) + \\
 &+ 0,02(2-0) + 0,05(1-0)] = \$2100.
 \end{aligned}$$

Расчеты показывают, что оптимальный уровень запасов равен 2.

Аналитическое решение. Эту задачу можно решить и аналитически. Решение приводится в приложении 3 в конце этой главы и дает следующий результат: оптимальный уровень запасов S_0 , при котором достигается минимум ожидаемых суммарных расходов, должен удовлетворять неравенствам

$$P(r \leq S_0 - 1) < \frac{C_2}{C_1 + C_2} < P(r \leq S_0). \quad (11)$$

Воспользуемся этими неравенствами для решения примера III.

1. Запишем данные примера III в виде таблицы 8.3.

2. Подсчитаем значение

$$\frac{C_2}{C_1 + C_2} = \frac{10\,000}{500 + 10\,000} = 0,952.$$

3. По таблице 8.3 найдем значение S , удовлетворяющее условию

$$P(r \leq S - 1) < 0,952 < P(r \leq S);$$

таким значением будет $S = 2$:

$$P(r \leq 1) < 0,952 < P(r \leq 2),$$

т. е. $0,950 < 0,952 < 0,970$.

Итак, оптимальный уровень запасов равен 2. Следует заметить, что если S_0 таково, что

$$P(r \leq S_0) = \frac{C_2}{C_1 + C_2},$$

тогда имеется два оптимума: S_0 и $S_0 + 1$. Если же

$$P(r \leq S_0 - 1) = \frac{C_2}{C_1 + C_2},$$

то тогда также будет два оптимума: $S_0 - 1$ и S_0 . Таким образом, оба эти равенства эквивалентны.

Подведем итоги.

1. По исходным данным составляется таблица, в которой каждому уровню запасов r соответствует $P(r)$ и $P(r \leq S)$.

2. Вычисляется отношение $C_2/C_1 + C_2$.

3. Определяется значение S , удовлетворяющее неравенству (или равенству)

$$P(r \leq S - 1) < \frac{C_2}{C_1 + C_2} < P(r \leq S).$$

Оценка потерь, возникающих при недостаточных запасах. Полученным аналитическим решением можно воспользоваться для оценки тех убытков, на которые предприниматель фактически рассчитывает,

Таблица 8.3

S	r	$P(r)$	$P(r \leq S)$
0	1	0,900	0,900
1	1	0,050	0,950
2	2	0,020	0,970
3	3	0,010	0,980
4	4	0,010	0,990
5	5	0,010	1,000
6 или более		0,000	1,000
		1,000	

допуская дефицит запасов. В следующем примере предположим, что нам неизвестна функция потерь *, а уровень запасов, который предприниматель стремится поддерживать, равен трем изделиям. Естественно возникает вопрос, для какого штрафа этот уровень запасов будет оптимальным. На этот вопрос можно ответить следующим образом:

$$P(r \leq S_0 - 1) < \frac{C_2}{C_1 + C_2} < P(r \leq S_0).$$

Подставляя $S_0 = 3$, получим

$$P(r \leq 2) < \frac{C_2}{C_1 + C_2} < P(r \leq 3),$$

$$0,970 < \frac{C_2}{\$500 + C_2} < 0,980.$$

Определим минимальное значение C_2

$$\frac{C_2}{\$500 + C_2} = 0,970,$$

т. е.

$$C_2 = \frac{0,970 \cdot \$500}{1 - 0,970} = \$16\,167.$$

Определим максимальное значение

$$\frac{C_2}{\$500 + C_2} = 0,980,$$

т. е.

$$C_2 = \frac{0,980 \cdot \$500}{1 - 0,980} = \$24\,500.$$

Следовательно, $\$16\,167 \leq C_2 \leq \$24\,500$, т. е. предприниматель считает, что размер штрафа заключен в пределах от 16 167 до 24 500 долл.

МОДЕЛЬ IV

Эта задача отличается от предыдущей лишь тем, что в ней рассматривается непрерывный (а не дискретный) уровень запасов. Следовательно, вероятность того, что спрос будет заключен в пределах от r_1 до r_2 , выра-

жается интегралом $\int_{r_1}^{r_2} f(r) dr$, а вероятность того, что спрос не превзойдет величины S , равна **

$$\int_0^S f(r) dr = F(S).$$

Уравнение цен. Уравнение цен в данной задаче подобно уравнению (10), составленному для модели III, если вместо $P(r)$ рассматривать $f(r) dr$, а суммирование заменить интегрированием:

$$Q(S) = C_1 \int_0^S (S-r) f(r) dr + C_2 \int_S^\infty (r-S) f(r) dr. \quad (12)$$

* То есть в данном случае зависимость штрафа от размера дефицита. (Прим.перев.)

** Отрицательные заказы (т. е. возврат) здесь не рассматриваются.

Из аналитического решения этого уравнения (см. приложение 4) следует, что оптимальное значение S_0 , при котором ожидаемые суммарные расходы оказываются минимальными, должно удовлетворять следующему условию:

$$F(S) = \int_0^S f(r) dr = \frac{C_2}{C_1 + C_2}. \quad (13)$$

Пример IV. Кондитерское предприятие торгует вразвес своими тортами. Каждый фунт торта, проданный в день изготовления, приносит 95 центов прибыли. Все непроданные торты можно со скидкой продать на следующий день. Убыток при этом равен 15 центам за каждый фунт. Иными словами, штраф при неудовлетворенном спросе составляет 95 центов/фунт, а стоимость создания запасов — 15 центов за каждый фунт непроданного товара. На основании прошлого опыта известно, что плотность распределения для спроса на торты представляет собой треугольное распределение, показанное на рис. 8.4, т. е. функция плотности вероятности для r в данном случае равна

$$f(r) = 0,02 - 0,0002r.$$

Рис. 8.4. Распределение дневного спроса для кондитерского предприятия. Пример IV.

Требуется определить оптимальную дневную выработку.

Решение. Имеем $C_1 = 15$ центов, $C_2 = 95$ центов. Отсюда

$$\frac{C_2}{C_1 + C_2} = \frac{95}{15 + 95} = 0,8636.$$

Чтобы определить оптимальную ежедневную выработку, необходимо подсчитать уровень запасов, который удовлетворяет соотношению

$$F(S) = \frac{C_2}{C_1 + C_2} = 0,8636, \quad \text{т. е.} \quad \int_0^S f(r) dr = 0,8636.$$

Это может быть сделано следующим образом:

$$\begin{aligned} \int_0^S f(r) dr &= \int_0^S (0,02 - 0,0002r) dr = \left(0,02r - \frac{0,0002r^2}{2} \right)_0^S = \\ &= 0,02S - 0,0001S^2 = 0,8636. \end{aligned}$$

Получаем

$$S = 100 \pm 36,93;$$

следовательно, имеется два решения:

1. $q_1 = 100 + 36,93 = 136,93$ фунта.
2. $q_2 = 100 - 36,93 = 63,07$ фунта.

Первое решение не подходит, так как по данной функции плотности вероятности видно, что дневной спрос не превышает 100 фунтов. Следовательно, ответом будет второе решение.

В данном частном случае, поскольку $f(r)$ линейно, тот же результат можно получить и без применения интегрального исчисления из простых геометрических соображений.

На рис. 8.5 снова приведен график $f(r)$. Площадь под прямой $f(r)$ равна 1. Нам надо найти такое значение S , при котором площадь под прямой $f(r)$ от 0 до S (площадь трапеции $OCDS$) равна 0,8636. Площадь этой трапеции равна сумме площадей прямоугольника A и треугольника B . Таким образом, сумма площади A и площади B должна равняться 0,8636.

$$S[f(S)] + \frac{1}{2} S[0,02 - f(S)] = 0,8636,$$

т. е.

$$S(0,02 - 0,0002S) + \frac{1}{2} S(0,02 - 0,02 + 0,0002S) = 0,8636,$$

или

$$0,02S - 0,0001S^2 = 0,8636,$$

и окончательно

$$S = 100 \pm 36,93.$$

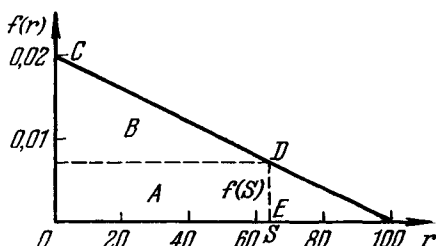


Рис. 8.5. График $f(r)$. Пример IV.

Интерпретация полученных результатов. Уравнение $F(S) = \frac{C_2}{C_1 + C_2}$ можно переписать в виде

$$\frac{F(S)}{1 - F(S)} = \frac{C_2}{C_1}$$

и получить интересную интерпретацию полученных результатов. При оптимальном уровне запасов отношение вероятности того, что спрос не превзойдет уровня запасов, к вероятности неудовлетворенного спроса равно отношению размера штрафа C_2 к стоимости создания запасов C_1 .

МОДЕЛЬ V

Рассматриваемый здесь случай отличается от предыдущих одним существенным обстоятельством: время, которое проходит, прежде чем заказ на пополнение запасов выполняется (время запаздывания поставки), значительно и должно приниматься во внимание.

Пример V. Владелец магазина ежедневно заказывает товар. Товар поступает к нему неделю спустя после заказа. Таким образом, время запаздывания поставки равно 7 дням. Пусть в какой-то день запас товара в магазине составляет 10 фунтов. В течение 6 предшествующих дней владелец магазина заказал 2, 4, 1, 10, 11 и 5 фунтов (в том порядке, как они будут поступать в следующие 6 дней). Для упрощения выкладок примем условия последнего примера: $C_1 = 15$ центов, $C_2 = 95$ центов, и распределение спроса на семидневный период (R')

$$f(R') = 0,02 - 0,0002R'.$$

Требуется определить размер заказа, который следует сделать на 7 дней вперед, т. е. найти q_7 .

Уравнение цен и его аналитическое решение. Прежде всего перечислим известные величины, которые рассматриваются здесь. k — число заказов, даваемых за время запаздывания поставки, S_0 — уровень запасов к моменту

очередного заказа, $q_1, q_2, \dots, q_{k-1}$ — размер заказов, которые сделаны ранее и должны поступить через 1, 2, $\dots$, $(k-1)$ день,

$$f(R') = f\left(\sum_{i=1}^k r_i\right),$$

где R' — потребность в товаре к моменту выполнения заказа (в данном случае через 7 дней). Задача состоит в определении такого размера заказа q_k , который минимизирует ожидаемые суммарные расходы за k дней (т. е. за время запаздывания поставки).

Выведем уравнение цен, которое относится ко всему времени, проходящему до получения заказанного товара, т. е. ко всем k периодам между заказами. Учесть в этом уравнении все k периодов необходимо, так как $q_1, q_2, \dots, q_{k-1}$ заказы уже произведены и на $k-1$ период вперед все определено. Поэтому минимизация суммарных ожидаемых затрат за все k периодов равнозначна минимизации ожидаемых затрат в k -м периоде. Запас s_k в конце k -го периода можно определить выражением

$$S_k = S_0 + \sum_{i=1}^{k-1} q_i + q_k - \sum_{i=1}^k r_i.$$

Пусть

$$S' = S_0 + \sum_{i=1}^{k-1} q_i + q_k, \quad R' = \sum_{i=1}^k r_i.$$

Подставляя в уравнение (12) S' вместо S и R' вместо R , получим

$$Q(S') = C_1 \int_0^{S'} (S' - R') f(R') dR' + C_2 \int_{S'}^{\infty} (R' - S') f(R') dR'. \quad (14)$$

Так как уравнение (14), по существу, аналогично уравнению (12), получаем из уравнения (13), что оптимальное значение S' должно удовлетворять уравнению

$$F(S'_0) = \frac{C_2}{C_1 + C_2}. \quad (15)$$

Раз оптимальное значение S' (т. е. S'_0) найдено, можно определить и оптимальную величину q_k :

$$q_{k0} = S'_0 - \left(S_0 + \sum_{i=1}^{k-1} q_i\right). \quad (16)$$

Подробное обоснование этого решения дается в конце этой главы в приложении 5.

Решение для примера V. Найдем прежде всего оптимальное значение S' . Оно должно удовлетворять условию

$$F(S') = \frac{C_2}{C_1 + C_2} = \frac{95}{15 + 95} = 0,8636.$$

Для распределения $f(R') = 0,02 - 0,0002R'$ величина S' определяется из уравнения

$$F(S') = \int_0^{S'} (0,02 - 0,0002R') dR' = 0,8636.$$

Выше в этой главе было найдено, что отсюда $S' = 63,07$, или 63 фунта. Так как

$$S' = S_0 + \sum_{i=1}^6 q_i + q_7,$$

мы можем определить оптимальную величину q_7 . Имеем

$$63 = 10 + 2 + 4 + 1 + 10 + 11 + 5 + q_7,$$

$$q_7 = 63 - 10 - 33 = 20 \text{ фунтов.}$$

Таким образом, наиболее экономичный размер заказа равен 20 фунтам.

МОДЕЛЬ VI

Задачи, которые здесь рассматриваются, подобны задачам, подходящим к модели III. Отличие состоит в том, что здесь расходование запасов происходит непрерывно *. При этом предполагается, что скорость расходования

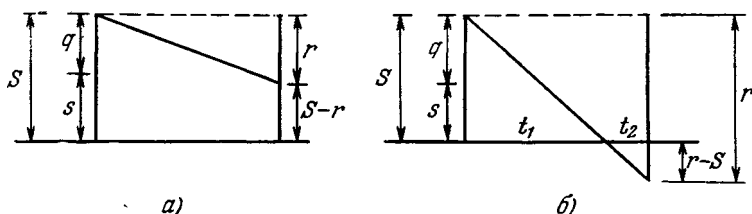
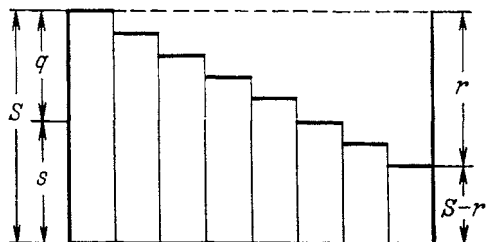


Рис. 8.6. Кривые запасов. Модель VI.

практически постоянна. Такую ситуацию можно представить графически (рис. 8.6). Рис. 8.6, а соответствует случаю $r \leq S$. Рис. 8.6, б имеет место при $r > S$, т. е. когда спрос превышает запасы. Область ниже горизонтальной оси соответствует дефициту запасов.

Пример VI. Предприниматель хочет определить для некоторого изделия оптимальный уровень запасов. Изделия равномерно расходуются, однако общая потребность в них непостоянна. Заказы самого предпринимателя выполняются практически немедленно. Эти заказы делаются в начале каждого месяца. Изучение спроса позволило определить вероятности различных значений месячного спроса. Эти вероятности приведены в таблице 8.4. Наконец, затраты на хранение одной детали на складе в течение

* Авторы вновь не делают различия между дискретной и непрерывной задачами. График на рис. 8.6, а будет в действительности иметь вид



илишь при большом S приближается к прямой линии. Так же обстоит дело и с рис. 8.6, б. (Прим. перев.)

месяца составляют 1 долл., а нехватка одной детали обходится в 20 долл. в месяц.

Таблица 8.4

Число изделий, требующихся в месяц	0	1	2	3	4	5	6 и более
Вероятность	0,1	0,2	0,2	0,3	0,1	0,1	0,0

Уравнение цен и его аналитическое решение. Рассмотрим сначала затраты для ситуации, изображенной на рис. 8.6, а. Средний уровень запасов здесь при фиксированном $r \leq S$ составляет за период между заказами

$$\frac{1}{2} [S + (S - r)] = S - \frac{r}{2}.$$

Так как $P(r)$ есть вероятность того, что потребуется ровно r изделий, суммарные ожидаемые расходы определяются как

$$C_1 \sum_{r=0}^S P(r) \left[S - \frac{r}{2} \right].$$

Теперь рассмотрим рис. 8.6, б, где $r > S$. Сначала учтем ту часть периода, в течение которой запасы оказываются достаточными. Доля $t_1 / (t_2 + t_1)$ периода, для которой это справедливо, равна S/r , как следует из подобия треугольников. Средний уровень запасов при этом равен $S/2$. Тогда для каждого значения r затраты в эту часть периода составляют

$$C_1 \left(\frac{S}{2} \right) \left(\frac{S}{r} \right) = C_1 \left(\frac{S^2}{2r} \right)^*.$$

Часть периода, когда спрос не удовлетворяется, равна $(r - S)/r$. Средний дефицит $(r - S)/2$. Тогда для каждого значения r штраф за нехватку запасов составляет

$$C_2 \left(\frac{r - S}{r} \right) \left(\frac{r - S}{2} \right) = C_2 \frac{(r - S)^2}{2r}.$$

Объединяя эти выражения, мы получаем следующее уравнение цен:

$$Q(S) = C_1 \sum_{r=0}^S P(r) \left(S - \frac{r}{2} \right) + C_1 \sum_{r=S+1}^{\infty} P(r) \frac{S^2}{2r} + C_2 \sum_{r=S+1}^{\infty} P(r) \frac{(r - S)^2}{2r}. \quad (17)$$

Пользуясь методом, который приводится в приложении 6 в конце главы, можно найти, что оптимальное значение S удовлетворяет следующему

* В действительности эти затраты определяются по формуле $C_1 \frac{(S+1)S}{2r} t$.

При $S = 2$ результаты отличаются в 1,5, а при $S = 1$ в 2 раза. Однако формулой авторов можно пользоваться при большом S : при $S = 10$ разница составляет 10%. (Прим. перев.)

условию:

$$\left[P[r \leq (S-1)] + \left(S - \frac{1}{2}\right) \sum_{r=S}^{\infty} \frac{P(r)}{r} \right] < \frac{C_2}{C_1 + C_2} < \left[P(r \leq S) + \left(S + \frac{1}{2}\right) \sum_{r=S+1}^{\infty} \frac{P(r)}{r} \right]. \quad (18)$$

Для упрощения представим (18) в следующем виде:

$$L(S-1) < \frac{C_2}{C_1 + C_2} < L(S). \quad (19)$$

Решение примера. Для облегчения вычислений удобно сначала подготовить рабочий бланк, который приведен в таблице 8.5. Далее вычислим

$$\frac{C_2}{C_1 + C_2} = \frac{20}{1 + 20} = \frac{20}{21} = 0,9524.$$

Выберем теперь из таблицы 8.5 значение S , которое удовлетворяет условию

$$L(S-1) < \frac{C_2}{C_1 + C_2} < L(S).$$

Этому условию удовлетворяет $S = 3$.

Действительно $0,8625 < 0,9524 < 0,9575$.

Таблица 8.5

Рабочий бланк

S	r	$P(r)$	$\frac{P(r)}{r}$	$\sum_{r=S+1}^{\infty} \frac{P(r)}{r}$	$\left(S + \frac{1}{2}\right) \sum_{r=S+1}^{\infty} \frac{P(r)}{r}$	$P(r \leq S)$	$L(S) = P(r \leq S) + \left(S + \frac{1}{2}\right) \sum_{r=S+1}^{\infty} \frac{P(r)}{r}$
0	0	0,1	∞	0,445	0,2225	0,1	0,3225
1	1	0,2	0,200	0,245	0,3675	0,3	0,6675
2	2	0,2	0,100	0,145	0,3625	0,5	0,8625
3	3	0,3	0,100	0,045	0,1575	0,8	0,9575
4	4	0,1	0,025	0,020	0,0900	0,9	0,9900
5	5	0,1	0,020	0,000	0,0000	1,0	1,0000
>5	>5	0,0	0,000	0,000	0,0000	1,0	1,0000

Ожидаемый суммарный расход, соответствующий уровню запасов, равному 3 изделиям, можно вычислить с помощью уравнения (17):

$$\begin{aligned} Q(3) &= \$1 [0,1 \times 3 + 0,2 \times 2,5 + 0,2 \times 2,0 + 0,3 \times 1,5] + \\ &+ \$1 \left[0,1 \times \frac{9}{8} + 0,1 \times \frac{9}{10} \right] + \$20 \left[0,1 \times \frac{1}{8} + 0,1 \times \frac{4}{10} \right] = \\ &= 1,65 + 0,2025 + 1,05 = 1,05 = \$2,9025. \end{aligned}$$

Применение модели VI в реальной задаче. До сих пор рассматривались весьма простые и, быть может, «нежизненные» примеры. Поэтому будет, вероятно, интересно применить одну из предложенных моделей в реальной задаче.

Возьмем для этого модель VI. Пусть некоторая компания занимается производством одной довольно дешевой и простой детали, применяемой в различных автомобилях, тракторах, самолетах и т. д. Технологический процесс ее производства разделен на две стадии. На первой стадии производится черновая обработка заготовки и термообработка, на второй стадии

деталь штампуется и проходит окончательную обработку. Это разделение объясняется тем, что число типоразмеров выпускаемой детали велико и многие типоразмеры изготавливаются из одинаковых полуфабрикатов, т. е. на второй стадии производится гораздо больше различных типоразмеров, чем на первой. Кроме того, окончательная обработка обходится дороже и требует большего времени. Задержка на второй стадии может принести предприятию значительный убыток, поэтому запасы полуфабрикатов поддерживаются такими, чтобы при поступлении заказа не возникало задержек из-за их отсутствия. Правило, которого придерживается компания при регулировании запасов, заключается в поддержании такого уровня запасов полуфабрикатов, чтобы удовлетворить 95% заказов, которые могут поступить за время между получением очередных партий полуфабрикатов. При этом, очевидно, компания считает, что убыток от нехватки одной детали обходится ей в 19 раз дороже стоимости хранения этой детали *. Задача заключается в определении как интервала времени между запусками в производство очередных партий полуфабрикатов, так и в нахождении размера этих партий. Из практических соображений желательно, чтобы время между партиями составляло 1, 2, 3 или более месяцев, а, конечно, не $2\frac{1}{2}$ или, скажем, $5\frac{1}{4}$ недели.

Упростим данные задачи, чтобы упростить вычисления и не затемнять саму схему расчета. Прежде всего, примем элементы расходов равными

$$C_1 = \$100, \quad C_2 = 19 \quad C_1 = \$1900, \quad C_s = \$350.$$

Предположим, что распределение месячного спроса соответствует таблице 8.6.

Таблица 8.6

Месячный спрос	r	0	1	2	3	4	5 и более
Вероятность	$P(r)$	0,1	0,2	0,4	0,2	0,1	0,0

Определим оптимальный размер партии в предположении, что выпуск организуется ежемесячно. Начнем с составления таблицы для вычисления (см. табл. 8.7).

Таблица 8.7

s	r	$P(r)$	$\frac{P(r)}{r}$	$\sum_{s+1}^{\infty} \frac{P(r)}{r}$	$(s + \frac{1}{2}) \sum_{s+1}^{\infty} \frac{P(r)}{r}$	$P(r \leq s)$	$L(s)$
0	0	0,1	∞	0,492	0,246	0,1	0,346
1	1	0,2	0,200	0,292	0,438	0,3	0,738
2	2	0,4	0,200	0,092	0,230	0,7	0,930
3	3	0,2	0,067	0,025	0,088	0,9	0,988
4	4	0,1	0,025	0,000	0,000	0,0	1,000

* Хотя вывод этого факта здесь не приводится, он полностью аналогичен выкладкам, выполненным для модели IV, т. е. условие $C_2 = 19C_1$ сразу же следует из $C_2/C_1 + \frac{1}{2} = 0,95$.

Затем подсчитываем

$$\frac{C_2}{C_1 + C_2} = \frac{19 \cdot 100}{100 + 19 \cdot 100} = \frac{1900}{2000} = 0,95.$$

Далее по таблице 8.7 выбираем значение S , удовлетворяющее условию

$$L(S-1) < \frac{C_2}{C_1 + C_2} < L(S).$$

Этому условию удовлетворяет $S=3$. Действительно, $0,930 < 0,95 < 0,988$. Общий ожидаемый месячный расход (связанный только с запасами) при $S=3$ подсчитывается с помощью уравнения (17):

$$\begin{aligned} Q(S=3) &= \$100 [0,1 \cdot 3 + 0,2 \cdot 2,5 + 0,4 \cdot 2,0 + 0,2 \cdot 1,5] + \\ &+ \$100 \left[0,1 \cdot \frac{9}{8} \right] + 1900 \left[0,1 \cdot \frac{1}{8} \right] = \$100 \cdot 1,9 + \$100 \cdot 0,1125 + \$1900 \cdot 0,0125 = \\ &= \$190 + \$11,25 + \$23,75 = \$225. \end{aligned}$$

Для определения суммарных ожидаемых накладных расходов* мы должны добавить стоимость заказа партии полуфабрикатов

$$\$225 + \$350 = \$575.$$

Таким образом, при ежемесячном заказе минимальный общий расход составляет в год

$$\$12 \cdot \$575 = \$6900.$$

Теперь мы хотим определить годовой расход для случая заказа партии 1 раз каждые 2 месяца. Для этого, во-первых, требуется установить закон распределения спроса в пределах двухмесячного периода. Вероятность отсутствия спроса в течение двух месяцев составляет $0,1 \cdot 0,1 = 0,01$, т. е. равна произведению вероятностей отсутствия спроса в течение одного месяца. Подобным образом мы можем найти, и весь закон распределения (P_2 — вероятность, связанная с двухмесячным периодом, P_1 — вероятность, связанная с месячным периодом)**:

$$P_2(1) = P_1(0)P_1(1) + P_1(1)P_1(0) = 0,2 \cdot 0,1 + 0,1 \cdot 0,2 = 0,04,$$

$$\begin{aligned} P_2(2) &= P_1(0)P_1(2) + P_1(1)P_1(1) + P_1(2)P_1(0) = \\ &= 0,1 \cdot 0,4 + 0,2 \cdot 0,2 + 0,4 \cdot 0,1 = 0,12, \end{aligned}$$

$$\begin{aligned} P_2(3) &= P_1(0)P_1(3) + P_1(1)P_1(2) + P_1(2)P_1(1) + P_1(3)P_1(0) = \\ &= 0,1 \cdot 0,2 + 0,2 \cdot 0,4 + 0,4 \cdot 0,2 + 0,2 \cdot 0,1 = 0,20, \end{aligned}$$

$$\begin{aligned} P_2(4) &= P_1(0)P_1(4) + P_1(1)P_1(3) + P_1(2)P_1(2) + P_1(3)P_1(1) + \\ &+ P_1(4)P_1(0) = 0,1 \cdot 0,1 + 0,2 \cdot 0,2 + 0,4 \cdot 0,4 + 0,2 \cdot 0,2 + 0,1 \cdot 0,1 = 0,26, \end{aligned}$$

$$\begin{aligned} P_2(5) &= P_1(1)P_1(4) + P_1(2)P_1(3) + P_1(3)P_1(2) + P_1(4)P_1(1) = \\ &= 0,2 \cdot 0,1 + 0,4 \cdot 0,2 + 0,2 \cdot 0,4 + 0,1 \cdot 0,2 = 0,20, \end{aligned}$$

$$\begin{aligned} P_2(6) &= P_1(2)P_1(4) + P_1(3)P_1(3) + P_1(4)P_1(2) = \\ &= 0,4 \cdot 0,1 + 0,2 \cdot 0,2 + 0,1 \cdot 0,4 = 0,12, \end{aligned}$$

$$P_2(7) = P_1(3)P_1(4) + P_1(4)P_1(3) = 0,2 \cdot 0,1 + 0,1 \cdot 0,2 = 0,04,$$

$$P_2(8) = P_1(4)P_1(4) = 0,1 \cdot 0,1 = 0,01.$$

* Заметим, что себестоимость при производстве (без стоимости заказа) не зависит от размера партии. Поэтому она не включается в накладные расходы производства.

** Здесь предполагается, что закон распределения спроса не меняется от месяца к месяцу.

Как и ранее, подготовим расчетный бланк (см. табл. 8.8).

Таблица 8.8

s	r	$P(r)$	$\frac{P(r)}{r}$	$\sum_{s+1}^{\infty} \frac{P(r)}{r}$	$(s + \frac{1}{2}) \sum_{s+1}^{\infty} \frac{P(r)}{r}$	$P(r \leq s)$	$L(s)$
0	0	0,01		0,299	0,150	0,01	0,160
1	1	0,04	0,040	0,259	0,389	0,05	0,439
2	2	0,12	0,060	0,199	0,498	0,07	0,668
3	3	0,20	0,067	0,132	0,462	0,37	0,832
4	4	0,26	0,065	0,067	0,302	0,63	0,932
5	5	0,20	0,040	0,027	0,149	0,83	0,979
6	6	0,12	0,020	0,007	0,046	0,95	0,996
7	7	0,04	0,006	0,001	0,008	0,99	0,998
8	8	0,01	0,001	0,000	0,000	1,00	1,000

Таким образом, если заказ производится каждые 2 месяца, минимальный суммарный ожидаемый расход составляет за год

$$\begin{aligned}
 & 6 \left\{ \$350 + \$200 [0,01 \cdot 5 + 0,04 \cdot 4,5 + 0,12 \cdot 4,0 + 0,20 \cdot 3,5 + \right. \\
 & \quad \left. + 0,26 \cdot 3,0 + 0,20 \cdot 2,5] + \$200 \left[0,12 \cdot \frac{25}{12} + 0,04 \cdot \frac{25}{14} + 0,01 \cdot \frac{25}{16} \right] + \right. \\
 & \quad \left. + \$3800 \left[0,12 \cdot \frac{1}{12} + 0,04 \cdot \frac{4}{14} + 0,01 \cdot \frac{9}{16} \right] \right\} = \\
 & = 6 \{ \$350 + \$200 \cdot 2,69 + \$200 \cdot 0,337 + \$3800 \cdot 0,027 \} = 6 \cdot \$1058 = \$6348.
 \end{aligned}$$

Теперь можно произвести расчет для трехмесячного цикла между партиями. При этом (читатель может проверить эти цифры) минимум суммарных расходов за год составляет 7109 долл.

Таким образом, мы можем следующим образом подвести итоги:

Цикл производства в месяцах	1	2	3
Минимум суммарных расходов в год	\$6900	\$6348	\$7109

В данном случае дальнейшие вычисления показали бы, что при увеличивающемся цикле производства годовые расходы возрастают. Таким образом ясно, что достаточно каждый второй месяц производить такое количество деталей, чтобы начальный запас состоял из пяти деталей.

Приложение 1

В этом приложении мы определим значение q (обозначается q_0), минимизирующее общие ожидаемые расходы $Q(q)$, где

$$Q(q) = \frac{1}{2} C_1 T q + C_s \frac{R}{q}.$$

После дифференцирования по q находим

$$\frac{dQ}{dq} = \frac{1}{2} C_1 T - \frac{C_s R}{q^2}.$$

Приравнявая производную нулю, имеем *

$$q_0 = \sqrt{\frac{2RC_s}{TC_1}}.$$

Поэтому

$$t_{so} = \frac{Tq_0}{R} = \frac{T}{R} \sqrt{2 \frac{RC_s}{TC_1}} = \sqrt{2 \frac{TC_s}{RC_1}},$$

$$Q_0 = \frac{C_1 T}{2} q_0 + \frac{C_s R}{q_0} = \frac{C_1 T}{2} \sqrt{2 \frac{RC_s}{TC_1}} + \frac{C_s R}{\sqrt{2 \frac{RC_s}{TC_1}}}.$$

Таким образом,

$$Q_0 = \sqrt{2RTC_1C_s}.$$

(Заметим, что для этой модели при оптимальном q стоимость заказов равна расходам на хранение запасов.)

Приложение 2

В этом приложении мы определим значение q и S , при которых полные ожидаемые расходы будут минимальными

$$Q(q, S) = \frac{S^2 C_1 T}{2q} + \frac{(q-S)^2 C_2 T}{2q} + \frac{C_s R}{q}. \quad (20)$$

После дифференцирования имеем

$$\frac{\partial Q}{\partial S} = \frac{SC_1 T}{q} - \frac{(q-S) C_2 T}{q},$$

$$\frac{\partial Q}{\partial q} = -\frac{S^2 C_1 T}{2q^2} + \frac{4q(q-S) - 2(q-S)^2}{4q^2} C_2 T - \frac{C_s R}{q^2}.$$

Приравнявая эти частные производные нулю ** и упрощая полученные выражения, получаем

$$\left. \begin{aligned} S &= q \frac{C_2}{C_1 + C_2}, \\ q^2 C_2 - (C_1 + C_2) S^2 &= \frac{2C_s R}{T}. \end{aligned} \right\} \quad (21)$$

Решая эту систему уравнений относительно S и q , находим

$$q_0 = \sqrt{2 \frac{RC_s}{TC_1}} \sqrt{\frac{C_1 + C_2}{C_2}}, \quad (22)$$

$$S_0 = \sqrt{2 \frac{RC_s}{TC_1}} \sqrt{\frac{C_2}{C_1 + C_2}}, \quad (23)$$

* Так как для $q = q_0$ $\frac{d^2 Q}{dq^2} = 2C_s R / q^3 > 0$, имеем при $q = q_0$ минимум.

** Равенство нулю частных производных — необходимое, но недостаточное условие экстремума. Необходимые и достаточные условия максимума и минимума изложены в [24, стр. 281].

и, следовательно,

$$t_{s0} = \frac{Tq_0}{R} = \sqrt{2 \frac{TC_s}{RC_1}} \sqrt{\frac{C_1 + C_2}{C_2}}. \quad (24)$$

Чтобы получить Q_0 , заметим, что

$$\frac{S_0^2}{q_0} = q_0 \left(\frac{C_2}{C_1 + C_2} \right)^2.$$

Подставляя (23) и (24) в (20), после упрощения получаем

$$Q_0 = \sqrt{2RTC_1C_3} \sqrt{\frac{C_2}{C_1 + C_2}}. \quad (25)$$

Заметим, что для оптимальных значений q_0 и S_0 затраты на хранение, как и раньше, равны стоимости заказов. Кроме того, отношение расходов на хранение излишков к убыткам, возникающим при дефиците, обратно пропорционально отношению C_1 к C_2 .

Приложение 3

Определим величину S , при которой ожидаемые суммарные расходы Q минимальны

$$Q(S) = C_1 \sum_{r=0}^S (S-r) P(r) + C_2 \sum_{r=S+2}^{\infty} (r-S) P(r). \quad (26)$$

Положим в этом уравнении $S = S+1$. Тогда

$$\begin{aligned} Q(S+1) &= C_1 \sum_{r=0}^{S+1} (S+1-r) P(r) + C_2 \sum_{r=S+1}^{\infty} (r-S-1) P(r) = \\ &= C_1 \sum_{r=0}^S (S+1-r) P(r) + C_1 [(S+1) - (S+1)] P(S+1) + \\ &+ C_2 \sum_{r=S+1}^{\infty} (r-S-1) P(r) - C_2 [(S+1) - (S+1)] P(S+1) = \\ &= C_1 \sum_{r=0}^S (S-r) P(r) + C_1 \sum_{r=0}^S P(r) + C_2 \sum_{r=S+1}^{\infty} (r-S) P(r) - C_2 \sum_{r=S+1}^{\infty} P(r). \end{aligned}$$

Так как

$$\sum_{r=0}^{\infty} P(r) = 1$$

или

$$\sum_{r=S+1}^{\infty} P(r) = 1 - \sum_{r=0}^S P(r),$$

можем написать

$$Q(S+1) = Q(S) + (C_1 + C_2) P(r \leq S) - C_2 \quad (27)$$

Точно так же

$$Q(S-1) = Q(S) - (C_1 + C_2) P(r \leq S-1) + C_2. \quad (28)$$

Теперь рассмотрим S_0 , при котором

$$\left. \begin{aligned} (C_1 + C_2) P(r \leq S_0) - C_2 &> 0, \\ -(C_1 + C_2) P(r \leq S_0 - 1) + C_2 &> 0. \end{aligned} \right\} \quad (29)$$

Для любого целого S' , большего S_0 , и любого целого S'' , меньшего S_0 , неравенства (29) будут выполняться, так как при возрастании $P(r \leq S_0)$ величина S_0 не возрастает. Следовательно, если неравенства (29) удовлетворяются*, то

$$Q(S'') > Q(S_0) \quad \text{при любом } S'' < S_0$$

и

$$Q(S') > Q(S_0) \quad \text{при любом } S' > S_0.$$

Таким образом, мы нашли значение S , при котором ожидаемые суммарные расходы минимальны; таким S будет S_0 , удовлетворяющее неравенствам (29). Эти неравенства можно записать иначе:

$$P(r \leq S_0 - 1) < \frac{C_2}{C_1 + C_2} < P(r \leq S_0). \quad (30)$$

Заметим, что если S_0 таково, что

$$P(r \leq S_0 - 1) < \frac{C_2}{C_1 + C_2} = P(r \leq S_0),$$

из уравнения (27) получим

$$Q(S + 1) = Q(S).$$

В этом случае оптимальным значением S будет либо S_0 , либо $S_0 + 1$. Если же S_0 таково**, то

$$P(r \leq S_0 - 1) = \frac{C_2}{C_1 + C_2} < P(r \leq S_0),$$

из уравнения (28) получим

$$Q(S_0 - 1) = Q(S_0),$$

и в этом случае оптимальным значением S будет либо $S_0 - 1$, либо S_0 .

Приложение 4

Определим значение S , при котором ожидаемые суммарные расходы Q минимальны. Имеем

$$Q(S) = C_1 \int_0^S (S - r) f(r) dr + C_2 \int_S^\infty (r - S) f(r) dr. \quad (31)$$

Если некоторая функция $g(x)$ записывается в виде

$$g(x) = \int_{h(x)}^{k(x)} f(x, y) dy,$$

то

$$\frac{dg(x)}{dx} = \int_{h(x)}^{k(x)} \frac{\partial f(x, y)}{\partial x} dy + f[x, k(x)] \frac{dk(x)}{dx} - f[x, h(x)] \frac{dh(x)}{dx}. \quad (32)$$

* Это показывает, что в S_0 имеется абсолютный, а не относительный (локальный) минимум.

** Следует отметить, что эти два случая эквивалентны.

Таким образом, из уравнения (31) получаем

$$\begin{aligned} \frac{dQ}{dS} &= C_1 \int_0^S f_1(r) dr - C_2 \int_S^\infty f(r) dr = C_1 F(S) - C_2 [1 - F(S)] = \\ &= (C_1 + C_2) F(S) - C_2. \end{aligned} \quad (33)$$

Ожидаемые суммарные расходы будут иметь относительный экстремум (максимум или минимум) при S_0 , если

$$\left. \frac{dQ}{dS} \right|_{S=S_0} = 0.$$

Из уравнения (33) имеем поэтому необходимое условие экстремума:

$$(C_1 + C_2) F(S) - C_2 = 0.$$

Отсюда

$$F(S_0) = \frac{C_2}{C_1 + C_2}. \quad (34)$$

Далее,

$$\left. \frac{d^2Q}{dS^2} \right|_{S=S_0} = C_1 f(S_0) + C_2 f(S_0) = (C_1 + C_2) f(S_0).$$

Так как C_1 и C_2 не равны нулю, а $f(S) \geq 0$, то

$$\left. \frac{d^2Q}{dS^2} \right|_{S=S_0} \geq 0.$$

Если имеет место неравенство, то при S_0 получаем минимум. Равенство возможно лишь при $f(S_0) = 0$. Функция $f(r)$ непрерывна, и $f(r) \geq 0$; поэтому, если $f(S_0) = 0$, функция $f(r)$ в S_0 имеет минимум, равный нулю. Отсюда следует, что при $S = S_0$ функция $Q(S)$ всегда имеет минимум, т. е. при $S = S_0$ ожидаемые суммарные расходы минимальны.

Приложение 5

Приведем подробное решение уравнения модели V.

Пусть

$$\begin{aligned} s_1 &= s_0 + q_1 - r_1, \\ s_2 &= s_1 + q_2 - r_2 = s_0 + (q_1 + q_2) - (r_1 + r_2), \\ s_3 &= s_2 + q_3 - r_3 = s_0 + (q_1 + q_2 + q_3) - (r_1 + r_2 + r_3), \\ &\vdots \\ s_k &= s_{k-1} + q_k - r_k. \end{aligned}$$

Тогда

$$s_k = s_0 + (q_1 + q_2 + \dots + q_k) - (r_1 + r_2 + \dots + r_k),$$

т. е.

$$s_k = s_0 + \sum_{i=1}^{k-1} q_i + q_k - \sum_{i=1}^k r_i.$$

Пусть

$$S' = s_0 + \sum_{i=1}^{k-1} q_i + q_k, \quad R' = \sum_{i=1}^k r_i.$$

Имеем тогда $dS' = dq_k$. Нетрудно также увидеть, что

$$\begin{aligned} s_k &> 0, \quad \text{когда } R' < S, \\ s_k &< 0, \quad \text{когда } R' > S'. \end{aligned}$$

Следовательно, так как

$$Q(S') = C_1 \int_0^{S'} (S' - R') f(R') dR' + C_2 \int_{S'}^{\infty} (R' - S') f_1(R') dR',$$

получим

$$Q(q_k) = C_1 \int_0^{S'} (S' - R') f(R') dR' + C_2 \int_{S'}^{\infty} (R' - S') f(R') dR'.$$

Вычислим первую производную

$$\frac{dQ}{dq_k} = \frac{dQ}{dS'} \frac{dS'}{dq_k} = C_1 F(S') - C_2 [1 - F(S')].$$

Приравнявая это выражение нулю, получаем

$$F(S'_0) = \frac{C_2}{C_1 + C_2}$$

и, наконец, определив оптимальное S'_0 , найдем q_{k0} :

$$q_{k0} = S'_0 - s_0 - \sum_{i=1}^{k-1} q_i.$$

Приложение 6

Определим величину S , при которой ожидаемые суммарные расходы Q будут минимальны. Имеем

$$Q(S) = C_1 \sum_0^S \left(S - \frac{r}{2}\right) P(r) + C_1 \sum_{S+1}^{\infty} \frac{S^2}{2r} P(r) + C_2 \sum_{S+1}^{\infty} \frac{(r-S)^2}{2r} P(r). \quad (35)$$

Полагая в этом уравнении $S = S+1$, получим

$$\begin{aligned} Q(S+1) = C_1 \sum_0^{S+1} \left(S+1 - \frac{r}{2}\right) P(r) + C_1 \sum_{S+2}^{\infty} \frac{(S+1)^2}{2r} P(r) + \\ + C_2 \sum_{S+2}^{\infty} \frac{(r-S-1)^2}{2r} P(r). \quad (36) \end{aligned}$$

Далее,

$$\begin{aligned} C_1 \sum_0^{S+1} \left(S+1 - \frac{r}{2}\right) P(r) &= C_1 \sum_0^S \left(S+1 - \frac{r}{2}\right) P(r) + \\ &+ C_1 \left(S+1 - \frac{S+1}{2}\right) P(S+1) = \\ &= C_1 \sum_0^S \left(S - \frac{r}{2}\right) P(r) + C_1 \sum_0^S P(r) + C_1 \left(\frac{S+1}{2}\right) P(S+1). \end{aligned}$$

Точно так же

$$\begin{aligned}
 C_1 \sum_{s+2}^{\infty} \frac{(S+1)^2}{2r} P(r) &= \\
 &= C_1 \sum_{s+1}^{\infty} \frac{S^2}{2r} P(r) + C_1 S \sum_{s+1}^{\infty} \frac{P(r)}{r} + \frac{C_1}{2} \sum_{s+1}^{\infty} \frac{P(r)}{r} - C_1 \frac{S+1}{2} P(S+1), \\
 C_2 \sum_{s+2}^{\infty} \frac{(r-S-1)^2}{2r} P(r) &= \\
 &= C_2 \sum_{s+1}^{\infty} \frac{(r-S)^2}{2r} P(r) - C_2 \sum_{s+1}^{\infty} P(r) + S C_2 \sum_{s+1}^{\infty} \frac{P(r)}{r} + \frac{1}{2} C_2 \sum_{s+1}^{\infty} \frac{P(r)}{r}.
 \end{aligned}$$

Из уравнений (35) и (36) следует

$$Q(S+1) = Q(S) + (C_1 + C_2) \left[P(r \leq S) + \left(S + \frac{1}{2} \right) \sum_{s+1}^{\infty} \frac{P(r)}{r} \right] - C_2. \quad (37)$$

Пусть теперь

$$L(S) = P(r \leq S) + \left(S + \frac{1}{2} \right) \sum_{s+1}^{\infty} \frac{P(r)}{r}. \quad (38)$$

Тогда из уравнения (37)

$$Q(S+1) = Q(S) + (C_1 + C_2) L(S) - C_2. \quad (39)$$

Точно так же, полагая в уравнении (39) $S = S-1$, получаем

$$Q(S-1) = Q(S) - (C_1 + C_2) L(S-1) + C_2. \quad (40)$$

$L(S)$ — невозрастающая функция S . Это можно показать следующим образом:

$$\begin{aligned}
 L(S+1) &= P(r \leq S+1) + \left(S+1 + \frac{1}{2} \right) \sum_{s+2}^{\infty} \frac{P(r)}{r} = \\
 &= P(r \leq S) + P(S+1) + \left(S + \frac{1}{2} \right) \sum_{s+1}^{\infty} \frac{P(r)}{r} - \frac{\left(S + \frac{1}{2} \right)}{(S+1)} P(S+1) + \\
 &\quad + \sum_{s+1}^{\infty} \frac{P(r)}{r} - \frac{P(S+1)}{(S+1)} = L(S) = \frac{1}{2} \frac{P(S+1)}{S+1} + \sum_{s+1}^{\infty} \frac{P(r)}{r},
 \end{aligned}$$

т. е.

$$L(S+1) = L(S) + \sum_{s+2}^{\infty} \frac{P(r)}{r} + \frac{1}{2} \frac{P(S+1)}{S+1}.$$

Так как

$$\sum_{s+2}^{\infty} \frac{P(r)}{r} + \frac{1}{2} \frac{P(S+1)}{S+1} \geq 0,$$

имеем

$$L(S+1) \geq L(S). \quad (41)$$

Теперь рассмотрим S_0 такое, что

$$\left. \begin{aligned} (C_1 + C_2) L(S_0) - C_2 &> 0, \\ -(C_1 + C_2) L(S_0 - 1) + C_2 &> 0. \end{aligned} \right\} \quad (42)$$

Для любых $S' > S_0$ и $S'' < S_0$ неравенства (42) выполняются, поскольку $L(S)$ не возрастает; следовательно,

$$Q(S'') > Q(S_0), \quad \text{если } S'' < S_0;$$

$$Q(S') > Q(S_0), \quad \text{если } S' > S_0.$$

Следовательно, ожидаемые суммарные расходы минимальны, если в качестве S взять S_0 , которое удовлетворяет неравенствам

$$L(S_0 - 1) < \frac{C_2}{C_1 + C_2} < L(S_0), \quad (43)$$

где

$$L(S) = P(r \leq S) + \left(S + \frac{1}{2}\right) \sum_{s+1}^{\infty} \frac{P(r)}{r}.$$

Заметим, наконец, что, так же как в приложении 3,

$$\frac{C_2}{C_1 + C_2} = L(S_0)$$

означает, что оптимальным значением S является либо S_0 , либо $S_0 + 1$, и

$$\frac{C_2}{C_1 + C_2} = L(S_0 - 1)$$

означает, что оптимальным будет либо $S_0 - 1$, либо S_0 .

ЛИТЕРАТУРА

1. Ackoff R. L., Production and Inventory Control in a Chemical Process, J. Opns. Res. Soc. Am., 3, no. 3, 319—338 (Aug. 1955).
2. Arrow K., Harris T. and Marschak J., Optimal Inventory Policy, Econometrica, 19, no. 3, 250—272 (July 1951).
3. Bellman R., Some Applications of the Theory of Dynamic Programming, J. Opns. Res. Soc. Am., 2, no. 3, 275—288 (Aug. 1954).
4. Bellman R., Some Problems in the Theory of Dynamic Programming, Econometrica, 22, no. 1, 37—48 (Jan. 1954).
5. Bellman R., The Theory of Dynamic Programming, Bull. Amer. Math. Soc., no. 6, 503—516 (Nov. 1954).
6. Bellman R., Glicksberg I. and Gross O., On the Optimal Inventory Equation, Mgmt. Sci., 2, no. 1, 83—104 (Oct. 1955).

7. Bellman R., Glicksberg I. and Gross O., The Theory of Dynamic Programming as Applied to a Smoothing Problem, *J. Soc. Ind. Appl. Math.*, 2, no. 2, 82—88 (June 1954).
8. Bergman E. B. and Clark A. J., An Optimal Inventory Policy for a Military Organization, RAND Report D-647, Mar. 30, 1955.
9. Charnes A., Cooper W. W. and Farr D., Linear Programming and Profit Preference Scheduling for a Manufacturing Firm, *J. Opns. Res. Soc. Am.*, 1, no. 3, 114—129 (May 1953).
10. Dannerstedt G., Production Scheduling for an Arbitrary Number of Periods Given the Sales Forecast in the Form of a Probability Distribution, *J. Opns. Res. Soc. Am.*, 3, no. 3, 300—318 (Aug. 1955).
11. Dvoretzky A., Kiefer J. and Wolfowitz J., On the Optimal Character of the (A, S) Policy in Inventory Theory, *Econometrica*, 21, no. 4, 586—596 (Oct. 1953).
12. Dvoretzky A., Kiefer J. and Wolfowitz J., The Inventory Problem, *Econometrica*, 20, no. 2, 187—222 (Apr. 1952) and no. 3, 450—466 (July 1952).
13. Eisenhart C., Some Inventory Problems, National Bureau of Standards, Techniques of Statistical Inference, A2-2C, Lecture 1, Jan. 6, 1948 (holographed notes).
14. Feeney G. J., A Basis for Strategic Decisions on Inventory Control Operations, *Mgmt. Sci.*, 2, no. 1, 69—82 (Oct. 1955).
15. Fry T. C., Probability and Its Engineering Uses, D. Van Nostrand and Co., New York, 1928 (see in particular pp. 229—232).
16. Hare V. C. and Hugli W. C., Applications of Operations Research to Production Scheduling and Inventory Control, II, Proceedings of the Conference on «What is Operations Research Accomplishing in Industry?», Case Institute of Technology, Cleveland, 1955.
17. Hoffman A. J. and Jacobs W., Smooth Patterns of Production, *Mgmt. Sci.*, 1, no. 1, 92—95 (Oct. 1954).
18. Holt C. C., Modigliani F. and Simon H. A., A Linear Decision Rule for Production and Employment Scheduling, *Mgmt. Sci.*, 2, no. 1, 1—30 (Oct. 1955).
19. Holt C. C. and Simon H. A., Optimal Decision Rules for Production and Inventory Control, Proceedings of the Conference on Production and Inventory Control, Case Institute of Technology, Cleveland, 1954.
20. Laderman J., Littauer S. B. and Weiss L., The Inventory Problem, *J. Amer. statist. Ass.*, 48, no. 264, 717—732 (Dec. 1953).
21. Magee J. F., Production Scheduling to Meet a Sales Forecast, Notes from M.I.T. Summer Course on Operations Research, Cambridge, pp. 134—138, 1953.
22. Magee J. F., Studies in Operations Research, I: Application of Linear Programming to Production Scheduling, Arthur D. Little, Inc., Cambridge, Mass. (unpublished).
23. Magee J. F., Guides to Inventory Policy. No. 1. Functions and Lot Size, *Harv. Busin. Rev.*, 34, no. 1, 49—60 (Jan.—Feb. 1956).
24. Pipes L. O. — A., Applied Mathematics for Engineers and Physicists, McGraw-Hill Book Co., New York, 1946.
25. Raymond F. E., Quantity and Economy in Manufacture, McGraw-Hill Book Co., New York, 1931.
26. Simon H. A., On the Application of Servomechanism Theory in the Study of Production Control, *Econometrica*, 20, no. 2, 247—268 (Apr. 1952).
27. Simon H. A. and Holt C. C., The Control of Inventory and Production Rates — A Survey, *J. Opns. Res. Soc. Am.*, 2, no. 3, 289—301 (Aug. 1954).
28. Tompkins C. B., Determination of a Safety Allowance, Logistics Papers, Engineering Research Associates, Issue no. 2, Appendix I to Bimonthly Progress Report No. 18.
29. Tompkins C. B., Lead Time and Optimal Allowances — an Extreme Example, Conference on Mathematical Problems in Logistics, George Washington University, Appendix I to Quarterly Progress Report No. 1, Dec. 1949—Feb. 1950.
30. Vassian H. J., Application of Discrete Variable Servo Theory to Inventory Control, *J. Opns. Res. Soc. Am.*, 3, no. 3, 272—282 (Aug. 1955).
31. Whitin T. M., Inventory Control and Price Theory, *Mgmt. Sci.*, 2, no. 1, 61—68 (Oct. 1955).
32. Whitin T. M., Inventory Control Research: A Survey, *Mgmt. Sci.*, 1, no. 1, 32—40 (Oct. 1954).
33. Whitin T. M., The Theory of Inventory Management, Princeton University Press, Princeton, 1953.

Глава 9

МОДЕЛИ УПРАВЛЕНИЯ ЗАПАСАМИ ПРИ НЕСКОЛЬКИХ УРОВНЯХ ЦЕН

В задачах управления запасами, которыми мы будем заниматься в этой главе, себестоимость и покупные цены уже не остаются постоянными. Этот случай, например, имеет место при закупке товара, на который установлена «скидка за количество»*. Хотя в этом направлении можно обобщить любую из описанных в главе 8 моделей, мы продемонстрируем его лишь на наиболее часто применяющейся на практике модели I.

Несколько изменяя условия модели I (см. главу 8), представим себе предпринимателя, который должен продавать или поставлять своим клиентам R единиц своей продукции равномерно в течение времени T . Таким образом, спрос является известным и фиксированным. Нехватка продукции считается недопустимой, и, следовательно, штраф за нехватку бесконечен (в обозначениях главы 8 $C_2 = \infty$). Введем обозначения для переменных элементов расходов, связанных с производством и приобретением продукции **::

k_1 — затраты на единицу продукции при ее производстве или приобретении;

P — стоимость складирования (выражаемая в процентах от стоимости единицы продукции);

C_s — стоимость оформления заказа при производстве партии изделий или сумма накладных расходов при приобретении партии товара;

W — ожидаемые суммарные расходы***;

W_0 — минимальные (оптимальные) ожидаемые суммарные расходы, как и в главе 8;

T — период времени, для которого ищется оптимальная стратегия;

R — полный спрос за время T ;

t_s — интервал времени между двумя заказами;

q — объем заказа (при пополнении запасов);

q_0 — оптимальный размер заказа (наиболее экономичный размер партии или наиболее экономичный объем закупки).

Переменные затраты при производстве или приобретении естественнее всего рассматривать применительно к случаю приобретения какого-либо товара, на который установлена «скидка за количество». Поэтому в дальнейшем мы будем касаться только покупки изделий при пополнении запасов. Этим без какой-либо потери общности мы упростим рассмотрение задач, которые приводятся ниже.

В этих задачах требуется определить, (1) как часто следует производить закупку товара и (2) каким должен быть объем каждой закупки.

* Скидка за количество (quantity discount)— снижение продажной цены при оптовых закупках. (Прим. перев.)

** В главе 8, где затраты на производство или приобретение продукции предполагались неизменными, их можно было не учитывать, а рассматривать лишь расходы, связанные со складированием продукции (holding charges). При вычислении в главе 8 мы так и поступали. Однако теперь эти переменные затраты на производство или приобретение продукции следует учитывать непосредственно, и для этого требуется введение новых обозначений.

*** Для обозначения этих расходов авторы, чтобы подчеркнуть отличие от главы 8, употребляют термин *ТЕС* (total expected cost). (Прим. перев.)

Здесь уместно отметить, что накладные расходы при приобретении партии товара, обозначаемые C_s , не обязательно ограничиваются одними лишь издержками оформления заказа. От объема закупки может зависеть, например, закупочная цена товара (receiving cost), расходы на приемочный контроль (cost of receiving inspection) и т. п. В тех случаях, когда на эти расходы влияет размер закупки, это следует учитывать при расчете.

Основные уравнения цен. Для каждого определенного значения покупной цены k_1 описанную выше ситуацию можно представить графически, как показано, например, на рис. 8.1. Так же как в главе 8, R/q — число заказов за период времени T , $t_s = Tq/R$, $1/2 q$ — средний запас в течение интервала t_s . Таким образом, для каждой партии (или закупки) запасы в расчете на всю партию составляют

$$\frac{1}{2} q t_s = \frac{1}{2} q (Tq/R) = \frac{1}{2} T q^2 / R,$$

а в расчете на одно изделие запасы составляют

$$\frac{1}{2} q t_s / q = \frac{1}{2} T q / R.$$

Для каждого заказа расходы будут определяться величинами: C_s — стоимость оформления заказа при покупке; qk_1 — покупная стоимость q изделий (цена одного изделия равна k_1);

$C_s \left(\frac{Tq}{2R} \right) P$ — дополнительные расходы при оформлении заказа на покупку изделий, образующих запас в течение периода t_s ;

$qk_1 \left(\frac{Tq}{2R} \right) P$ — накладные расходы при покупке изделий, образующих запас в течение периода t_s .

Таким образом, суммарные расходы в течение периода t_s составляют

$$C_s + qk_1 + C_s \frac{Tq}{2R} P + qk_1 \frac{Tq}{2R} P,$$

а суммарные расходы для всего периода T равны

$$W = \left(C_s + qk_1 + C_s \frac{Tq}{2R} P + qk_1 \frac{Tq}{2R} P \right) \frac{R}{q},$$

т. е.

$$W = \frac{C_s R}{q} + k_1 R + \frac{C_s T P}{2} + k_1 \frac{T P}{2} q. \quad (1)$$

Минимум суммарных ожидаемых расходов можно найти, приравняв нулю первую производную W по переменной q . Имеем

$$\frac{dW}{dq} = -\frac{C_s R}{q^2} + \frac{1}{2} k_1 T P.$$

Полагая $dW/dq = 0$, получаем *

$$q_0 = \sqrt{\frac{C_s R 2}{k_1 T P}}. \quad (2)$$

* Из сравнения (2) с уравнением (2) главы 8 (модель I) видно, что закупочная стоимость, равная здесь k_1 , составляет существенную часть стоимости складирования одного изделия. Это нетрудно понять, так как стоимость оформления заказа не зависит от размера партии.

Подстановка в уравнение (1) дает оптимальный полный расход W_0 при покупной цене k_1 :

$$W_0 = \frac{C_s R}{\sqrt{(2C_s R)/(k_1 T P)}} + k_1 R + \frac{C_s T P}{2} + k_1 \frac{T P}{2} \sqrt{\frac{2C_s R}{k_1 T P}} = \\ = \sqrt{\frac{k_1 T P C_s R}{2}} + k_1 R + \frac{1}{2} C_s T R + \sqrt{\frac{k_1 T P C_s R}{2}}.$$

Отсюда

$$W_0 = \sqrt{2k_1 T P C_s R} + k_1 R + \frac{1}{2} C_s T P. \quad (3)$$

Рассмотрим теперь первое обобщение этой задачи — случай, когда «скидка за количество» определяет два уровня покупной цены.

МОДЕЛЬ СОЗДАНИЯ ЗАПАСОВ ПРИ ДВУХ УРОВНЯХ ЦЕН

Здесь мы рассмотрим типичную ситуацию, возникающую при закупке товара, когда «скидка за количество» определяет два уровня цен. Эта ситуация может быть представлена следующим образом:

Об- ласть	Объем закупки	Покуп- ная цена
R_1 R_2	$1 \leq q_1 < b$ $q_2 \geq b$	k_{11} k_{12}

Здесь b — такой объем закупки, при достижении которого цена снижается. Таким образом, при любом размере закупки q_1 в пределах R_1 ожидаемые полные расходы W_1 будут равны

$$W_1 = \frac{C_s R}{q_1} + k_{11} R + \\ + \frac{C_s T P}{2} + k_{12} \frac{T R}{2} q_1, \quad (4)$$

а для любого размера закупки q_2 в R_2 ожидаемые полные расходы составляют

$$W_2 = \frac{C_s R}{q_2} + k_{12} R + \\ + \frac{C_s T R}{2} + k_{12} \frac{T P}{2} q_2. \quad (5)$$

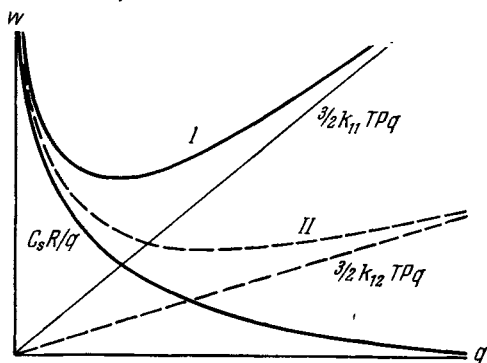


Рис. 9.1. Кривые закупок при двух уровнях цен.

Эту ситуацию можно следующим образом представить графически. На рис. 9.1 в выражении для W мы пока не учитываем члены $k_1 R + \frac{1}{2} C_s T P$, учет этих членов дает нам рис. 9.4. Так как $k_{12} < k_{11}$, то

$$\left(k_{12} R + \frac{1}{2} C_s T P\right) < \left(k_{11} R + \frac{1}{2} C_s T P\right).$$

Из рис. 9.1, а также из уравнений (4) и (5) очевидно, что минимум расходов для кривой II, которая соответствует k_{12} , ниже, чем минимум

расходов для кривой I , которая соответствует k_{11}^* . Поэтому, если обозначить $q_{1,0}$ и $q_{2,0}$ значения q_1 и q_2 , при которых расходы оказываются минимальными, мы получим следующие правила для принятия решения.

1. Подсчитывается $q_{2,0}$. Если $q_{2,0} \geq b$ (рис. 9.2), оптимальным размером закупки будет $q_{2,0}$.

2. Если $q_{2,0} < b$ (рис. 9.3), то при таком объеме закупки снижения цены не происходит. Кроме того, так как минимум затрат достигается при объеме закупки, меньшем чем b , т. е. $q_{2,0} < b$, следовательно, ожидаемые суммарные расходы будут в области R_2 моно-

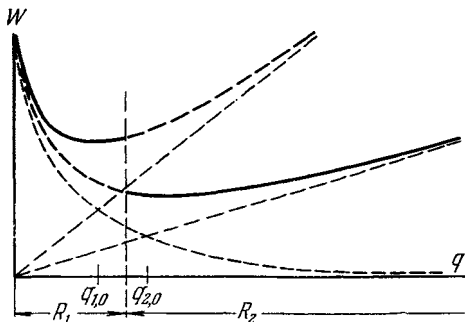


Рис. 9.2. Кривые закупок при $q_{2,0} \geq b$.

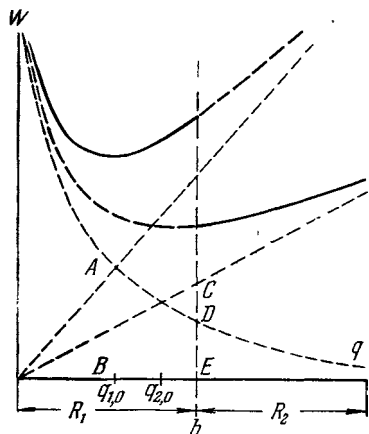


Рис. 9.3. Кривые закупок при $q_{2,0} < b$.

тонно возрастать и минимум расходов будет соответствовать размеру закупки $q = b$. Таким образом, чтобы узнать оптимальный объем заказа, нам требуется только сравнить суммарные ожидаемые расходы при заказе $q_{1,0}$ и при заказе $q = b$. Эти уравнения цен получаются из (1) и имеют вид

$$W_0(q_{1,0}) = \frac{C_s R}{q_{1,0}} + k_{11} R + \frac{C_s TP}{2} + k_{11} \frac{TP}{2} q_{1,0}, \quad (1a)$$

$$W(b) = \frac{C_s R}{b} + k_{12} R + \frac{C_s TP}{2} + k_{12} \frac{TP}{2} b. \quad (16)$$

Почленно сравнивая $W(b)$ и $W_0(q_{1,0})$, находим, так как $q_{1,0} < b$ и $k_{12} < k_{11}$,

$$\left(\frac{C_s R}{b} + k_{12} R + \frac{C_s TP}{2} \right) < \left(\frac{C_s R}{q_{1,0}} + k_{11} R + \frac{1}{2} C_s TP \right).$$

Рис. 9.4. Кривые закупок при $q_0 = b$.

Однако $\frac{1}{2} k_{12} TP b$ может быть и больше и меньше соответствующего члена $\frac{1}{2} k_{11} TP q_{1,0}$. Поэтому необходимо сравнить суммарные расходы, как показано ниже.

* Вообще при n уровнях цен имеют место неравенства $q_{1,0} < q_{2,0} < \dots < q_{n,0}$. Они легко получаются из уравнения (3) при $k_{1,n} < k_{1,n-1} < \dots < k_{11}$.

Из рис. 9.3 и 9.4 видно, что мы сравниваем

$$\left[2\overline{AB} + \left(k_{11}R + \frac{1}{2}C_sTP \right) \right]$$

и

$$\left[\overline{CE} + \overline{DE} + \left(k_{12}R + \frac{1}{2}C_sTP \right) \right].$$

Покажем теперь на трех примерах применение этих правил.

Пример I. Владелец машиностроительного завода требуется в течение года закупать 2400 отливок, т. е. предполагается, что спрос известный и фиксированный. На эти отливки существует «скидка за количество», которая устанавливает два уровня цен. Необходимо определить оптимальный размер закупки q_0 .

Таким образом, дано: T — 12 месяцев, R — 2400 изделий, P — 2%, C_s — 350 долл. за один заказ (в том числе оплата заявок и т. д.), b — 500, k_{11} — 10 долл. ($1 \leq q_1 < 500$), k_{12} — 9,25 долл. ($q_2 \geq 500$).

Вычисляя $q_{2,0}$ согласно (2), получаем

$$q_{2,0} = \sqrt{\frac{350 \cdot 2400 \cdot 2}{9,25 \cdot 12 \cdot 0,02}} = 870.$$

Так как $q_{2,0} = 870$ (т. е. больше, чем $b = 500$), оптимальным объемом закупки будет $q_{2,0} = 870$. На рис. 9.5 графически представлена данная ситуация.

Пример II. Иная ситуация возникает, если стоимость оформления заказа равна всего 100 долл., т. е. $C_s = 100$ долл.

В этом случае

$$q_{2,0} = \sqrt{\frac{100 \cdot 2400 \cdot 2}{9,25 \cdot 12 \cdot 0,02}} = 465.$$

Так как $q_{2,0} = 465 < 500$, нам придется также вычислить $q_{1,0}$

$$q_{1,0} = \sqrt{\frac{100 \cdot 2400 \cdot 2}{10 \cdot 12 \cdot 0,02}} = 447.$$

Теперь надо сравнить $W(447)$ и $W(500)$, т. е. сравнить оптимальные расходы, соответствующие цене k_{11} (в данном случае при $q_{1,0} = 447$), и расходы, связанные с приобретением наименьшего числа изделий, при котором нам обеспечивается скидка (в данном случае $b = 500$). Выражения для $W(b)$ и W_0 следуют из уравнений (1) и (3). Получаем

$$W_0(447) = \sqrt{2 \cdot 10 \cdot 12 \cdot 0,02 \cdot 100 \cdot 2400} + 10 \cdot 2400 + \frac{1}{2} \cdot 100 \cdot 12 \cdot 0,02.$$

Отсюда $W_0(447) = 25085$ долл.,

$$W(500) = \frac{100 \cdot 2400}{500} + 9,25 \cdot 2400 + \frac{1}{2} \cdot 100 \cdot 12 \cdot 0,02 + \frac{1}{2} \cdot 9,25 \cdot 12 \cdot 0,02 \cdot 500.$$

Отсюда $W(500) = 23\,247$ долл.

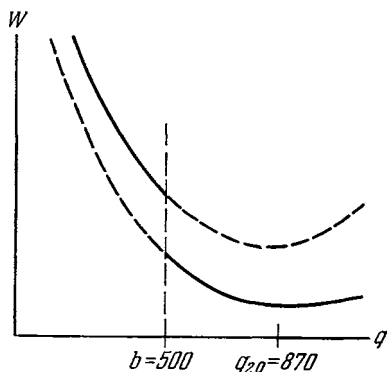
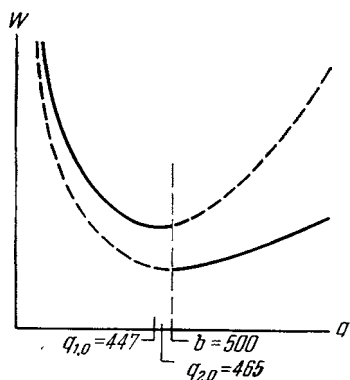
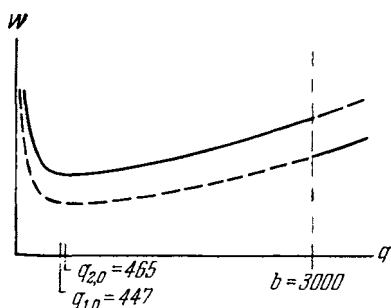


Рис. 9.5. $q_0 = 870$.

Так как $W(500) < W_0(447)$, оптимальный объем заказа определяется здесь условием скидки и $q_0 = 500$. Эта ситуация изображена на рис. 9.6.

Рис. 9.6. $q_0 = 500$.Рис. 9.7. $q_0 = 447$.

Пример III. Предположим теперь, что в условиях примера II изменилось условие скидки и снижение цены происходит при $q = 3000$. Тогда, как и раньше,

$$q_{2,0} = 465 < 3000$$

и $q_{1,0}$, как уже было найдено, равно 447. Соответственно $W_0(447) = 25\,085$ долл. Сравниваем $W_0(447)$ и $W(3000)$

$$W(3000) = \frac{100 \cdot 2400}{3000} + 9,25 \cdot 2400 + \frac{1}{2} \cdot 100 \cdot 12 \cdot 0,02 + \frac{1}{2} \cdot 9,25 \cdot 12 \cdot 0,02 \cdot 3000,$$

т. е. $W(3000) = 25\,622$ долл. Здесь $W(3000) > W_0(447)$, поэтому оптимальный размер заказа $q_0 = 447$. Эта ситуация представлена на рис. 9.7.

МОДЕЛИ ПОПОЛНЕНИЯ ЗАПАСОВ ПРИ ТРЕХ УРОВНЯХ ЦЕН

В этом разделе мы сделаем еще один шаг в обобщении моделей управления запасами и рассмотрим ситуацию, которая возникает, если условия скидки определяют три уровня цен. Представим это следующим образом:

Об- ласть	Объем закупки	Покуп- ная цена
R_1	$1 \leq q_1 < b_1$	k_{11}
R_2	$b_1 \leq q_2 < b_2$	k_{12}
R_3	$b_2 \leq q_3$	k_{13}

Здесь b_1 и b_2 — размер закупки, при превышении которого предоставляется скидка. Все рассуждения, относящиеся к предыдущему случаю, остаются справедливыми и здесь и дают нам следующие правила принятия решения.

1. Вычисляется $q_{3,0}$. Если $q_{3,0} \geq b_2$, оптимальный размер заказа равен $q_{3,0}$.

2. Если $q_{3,0} < b_2$, вычисляется $q_{2,0}$. Так как $q_{3,0} < b_2$, следовательно, $q_{2,0} < b_2^*$, причем либо $q_{2,0} < b_1$, либо $b_1 \leq q_{2,0} < b_2$. Если $q_{3,0} < b_2$

* Заметим, что $q_{1,0} < q_{2,0} < \dots < q_{n,0}$.

и $b_1 \leq q_{2,0} < b_2$, мы поступаем так же, как и в случае двух уровней цен, т. е. сравниваем $W_0(q_{2,0})$ и $W(b_2)$, чтобы найти оптимальный объем закупки.

3. Если $q_{3,0} < b_2$ и $q_{2,0} < b_1$, вычисляем $q_{1,0}$ которое обязательно будет удовлетворять неравенству $q_{1,0} < b_1$. В этом случае, чтобы найти оптимальный размер заказа, сравниваем $W_0(q_{1,0})$, $W(b_1)$ и $W(b_2)$.

Таблица 9.1

Символы (определяются в начале главы)	Единица измерения	Пример IV	Пример V	Пример VI	Пример VII	Пример VIII
T	месяцы	12	12	12	12	12
R	штуки	2400	2400	2400	2400	2400
C_s	доллары	350	100	100	100	100
P	проценты	2	2	2	2	2
k_{11}	доллары	10	10	10	10	10
b_1	штуки	500	500	400	500	3000
k_{12}	доллары	9,25	9,25	9,25	9,25	9,25
b_2	штуки	750	750	3000	1500	5000
k_{13}	доллары	8,75	8,75	8,75	9,0	8,75

Проиллюстрируем применение этих правил на пяти примерах. Читатель, вероятно, заметит, что эти примеры исчерпывают все пять возможных ситуаций, которые могут возникнуть в задаче определения оптимального размера заказа в случае трех уровней цен. В таблице 9.1 приводятся данные для всех этих пяти примеров.

Пример IV. Вычисляя $q_{3,0}$, получаем

$$q_{3,0} = \sqrt{\frac{350 \cdot 2400 \cdot 2}{8,75 \cdot 12 \cdot 0,02}} = 894 > 750.$$

Следовательно, оптимальный размер заказа $q_0 = 894$. Этот случай представлен на рис. 9.8.

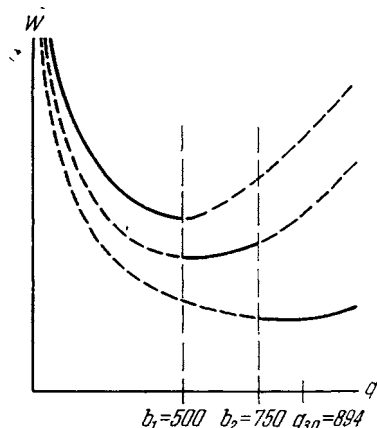
Пример V. Вычисляя q , получаем

$$q_{3,0} = \sqrt{\frac{100 \cdot 2400 \cdot 2}{8,75 \cdot 12 \cdot 0,02}} = 478.$$

Так как $q_{3,0} = 478 < 750$, мы теперь вычисляем * $q_{2,0}$. Это уже было сделано в примере II, откуда $q_{2,0} = 465$.

Поскольку $q_{2,0} = 465 < 500$, нам следует затем вычислить $q_{1,0}$ которое, также из примера II, $q_{1,0} = 447$.

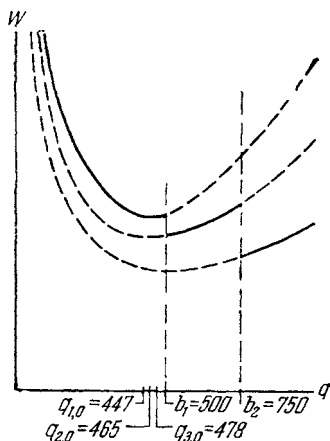
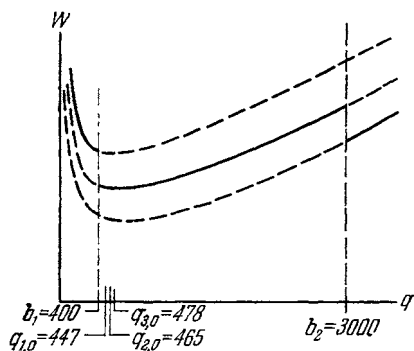
Теперь остается сравнить $W_0(447)$, $W(500)$ и $W(750)$. Из примера II мы уже имеем $W_0(447) = 25\,085$ долл., $W(500) = 23\,247$ долл., а из уравнения (1) находим $W(750) = 22\,119,5$ долл.


 Рис. 9.8. $q_0 = 894$.

* Читатель мог бы заметить, что из неравенства $q_{n-1,0} < q_{n,0}$ следует, что $q_{2,0}$ в данном случае будет меньше 478. Так как $478 < 500$, то нам фактически не нужно вычислять $q_{2,0}$, так как это значение здесь уже не может быть решением.

Мы получили, таким образом, $W(750) < W(500) < W(447)$; следовательно, экономичный размер заказа равен $q_0 = 750$. Этот случай представлен на рис. 9.9.

Пример VI. Так же как и в примере V, $q_{3,0} = 478$, $q_{2,0} = 465$. Так как $q_{2,0}$ оказалось в диапазоне $R_2(400, 3000)$, нам не требуется вычислять

Рис. 9.9. $q_0 = 750$.Рис. 9.10. $q_0 = 465$.

$q_{1,0}$ и достаточно сравнить $W_0(465)$ и $W(3000)$:

$$W(3000) = 24\,242 \text{ долл.}, W_0(465) = 23\,244 \text{ долл.}$$

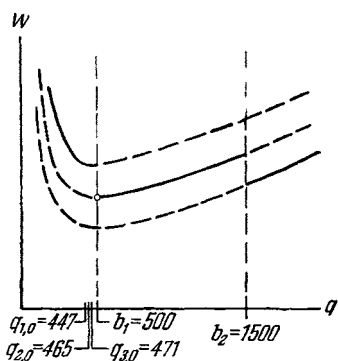
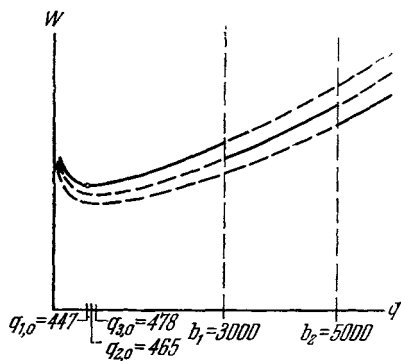
Таким образом, в данном случае экономичный размер заказа равен $q_0 = 465$. Этот случай представлен на рис. 9.10.

Пример VII. Здесь

$$q_{3,0} = \sqrt{\frac{100 \cdot 2400 \cdot 2}{9 \cdot 12 \cdot 0,02}} = 471 < 1500.$$

Из примера II имеем $q_{2,0} = 465 < 500$, $q_{1,0} = 447$.

Сравним $W(150)$, $W(500)$ и $W_0(447)$ $W(1500) = 23\,392$ долл.

Рис. 9.11. $q_0 = 500$ Рис. 9.12. $q_0 = 447$.

Из примера II $W(500) = 23\,247$ долл., $W_0(447) = 25\,085$ долл.

Таким образом, здесь оптимальный размер заказа равен $q_0 = 500$. Этот случай представлен на рис. 9.11.

Пример VIII. Так же как в примере V,

$$q_{3,0} = 478, \quad q_{2,0} = 465, \quad q_{1,0} = 447.$$

Здесь мы должны сравнить $W_0(447)$, $W(3000)$ и $W(5000)$.

Из примера II имеем $W_0(447) = 25\,085$ долл. Для $q = 3000$ при цене 9,25 долл и для $q = 5000$ при цене 8,75 долл. получаем

$$W(3000) = 25\,622 \text{ долл.}, \quad W(5000) = 26\,310 \text{ долл.}$$

Сравнение $W_0(447)$, $W(3000)$ и $W(5000)$ дает нам размер наиболее экономичной закупки, равный $q_0 = 447$. Этот случай представлен на рис. 9.12.

МОДЕЛИ ПОПОЛНЕНИЯ ЗАПАСОВ ПРИ ЧИСЛЕ УРОВНЕЙ ЦЕН СВЫШЕ ТРЕХ

После того как мы рассмотрели правила для определения наиболее экономичного размера заказа при двух и трех уровнях покупных цен, нетрудно обобщить теперь эти правила на случай любого числа уровней покупных цен.

Обозначим области, в которых цены остаются постоянными, через $R_1, R_2, \dots, R_n$, размеры заказов, при которых цены снижаются, через

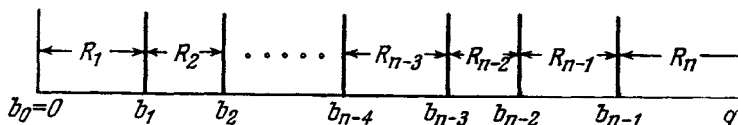


Рис. 9.13. Общая процедура решения при n уровнях цен.

$b_1, b_2, \dots, b_{n-1}$ и экономичные размеры заказов, соответствующие определенной цене, через $q_{1,0}, q_{2,0}, \dots, q_{n,0}$. Рассмотрим рис. 9.13. В этом случае мы получаем следующие правила решения.

1. Вычисляем $q_{n,0}$. Если $q_{n,0} \geq b_{n-1}$, оптимальный размер заказа будет равен $q_{n,0}$.

2. Если $q_{n,0} < b_{n-1}$, вычисляем $q_{n-1,0}$. Если окажется, что $q_{n-1,0} \geq b_{n-2}$ (т. е. $b_{n-2} \leq q_{n-1,0} < b_{n-1}$), поступаем так же, как в случае двух уровней цен, т. е. для определения оптимального размера заказа сравниваем $W_0(q_{n-1,0})$ и $W(b_{n-1})$.

3. Если $q_{n-1,0} < b_{n-2}$, вычисляем $q_{n-2,0}$. Если при этом $q_{n-2,0} \geq b_{n-3}$, поступаем далее так же, как в случае трех уровней цен, т. е. для определения оптимального размера заказа сравниваем $W_0(q_{n-2,0})$, $W(b_{n-2})$ и $W(b_{n-1})$.

4. Если $q_{n-2,0} < b_{n-3}$, тогда вычисляем $q_{n-3,0}$. Если $q_{n-3,0} \geq b_{n-4}$, сравниваем $W_0(q_{n-3,0})$ с $W(b_{n-3})$, $W(b_{n-2})$ и $W(b_{n-1})$.

5. И вообще продолжаем этот процесс до тех пор, пока $q_{n-j,0} \geq b_{n-(j+1)}$ [$j = 0, 1, \dots, (n-1)$], и затем сравниваем $W_0(q_{n-j,0})$ с $W(b_{n-j})$, $W(b_{n-j+1})$, $\dots$, $W(b_{n-1})$.

Этот процесс имеет конечное число шагов, не превышающее n , где n — число уровней покупной цены.

ПРИМЕНЕНИЕ РАССМОТРЕННЫХ МОДЕЛЕЙ

Рассмотренные выше модели управления запасами в несколько измененном виде были использованы компанией, которая описывалась в главе 4. Эта компания в небольшом городке на Среднем Западе США занимается производством мощных двигателей.

* Полагаем $b_0 = 0$.

Значительную часть изделий, необходимых для производства, компания закупает на стороне. Эти покупные изделия можно разделить на две категории: (1) изделия, для которых поставщиком уже установлены условия «скидки за количество», и (2) изделия, для которых условия скидки не установлены (например, изделия, производство которых существует недавно).

Для изделий второй категории потребитель обычно представляет на рассмотрение поставщика предполагаемую цену. Предполагаемые цены анализируются поставщиком, и затем устанавливаются отпускные цены, причем «скидки за количество» часто учитывают требования потребителя.

Применение моделей управления запасами с несколькими уровнями цен при заказе изделий первой категории довольно очевидно. При заказе изделий второй категории можно получить дополнительную экономию средств, если с помощью моделей запасов сначала установить желаемые условия снижения цен. Проиллюстрируем сказанное на примере IV, который был рассмотрен выше.

В примере IV экономичный размер заказа оказался равным 894 изделиям. Этот размер заказа относится к области заказов, превышающих 750 изделий, и уровень цены в этой области равен 8,75 долл. Естественно, что установленная по требованию заказчика скидка при 894 изделиях давала бы более низкую цену.

Если, например, условия скидки для заказов, превышающих 894 изделия, определяли бы цену 8,5 долл., суммарные ожидаемые расходы были бы равны

$W(894 \text{ при цене } \$8,5) =$

$$= \left[350 + 894(8,50) + 350 \frac{(12)(894)}{2(2400)}(0,02) + 894(8,50) \frac{(12)(894)}{(2)(2400)}(0,02) \right] \frac{2400}{894},$$

т. е. $W(894 \text{ при цене } \$8,5) = \$22\,293,5$. Суммарные ожидаемые расходы при цене 8,75 долл. составляют

$$W(894 \text{ при цене } \$8,75) = \$22\,920.$$

Таким образом, экономия, которую давало бы введение по требованию потребителя еще одного уровня цен, составляла бы 627 долл.*.

Здесь следует заметить, что при этом «наилучшего» решения мы не получаем. Действительно, если для примера IV подсчитать экономичный размер заказа, исходя из цены 8,5 долл. ($q \geq 894$), мы получим

$$q_0 = \sqrt{\frac{(350)(2400)(2)}{(8,50)(12)(0,02)}} = 907.$$

Однако разница между оптимальным суммарным расходом и «почти оптимальным» обычно бывает невелика. Возвращаясь к нашему примеру, находим, что эта разница

$$\begin{aligned} W(q=894 \text{ при цене } \$8,5) - W_0(q_0=907, \text{ при цене } \$8,5) = \\ = \$22\,293,5 - \$22\,293 = \$0,5 \end{aligned}$$

действительно пренебрежимо мала.

* Заметим, что помимо очевидной экономии $894(8,75 - 8,5) = 223,5$ долл. мы получаем также уменьшение стоимости складирования.

Понятно, что такая процедура шаг за шагом приведет к оптимальному ответу. Возвращаясь к нашему примеру, «улучшение» ответа можно получить, если вновь пересмотреть цены и установить снижение цен при 907 изделиях. Однако, как это уже произошло в нашем примере, в конце концов наступает момент, когда вычисления уже не будут окупаться достигаемой экономией.

Г л а в а 10

МОДЕЛИ УПРАВЛЕНИЯ ЗАПАСАМИ ПРИ НАЛИЧИИ ОГРАНИЧЕНИЙ

ВВЕДЕНИЕ

В главе 8 мы получили целый ряд моделей управления запасами для случаев, когда можно не учитывать никаких ограничений, связанных с возможностями производства, размерами хранилищ, временем или денежными средствами. Когда такие ограничения имеют место и запасы состоят более чем из одного вида продукции, приходится распределять имеющиеся ограниченные ресурсы между этими видами продукции. Следовательно, модель должна помочь нам определить, какое количество каждого вида изделий надо производить (или закупать) при заданных ограничениях. Такие модели являются «смешанными» в том смысле, что они описывают как распределение (или планирование производства), так и само управление запасами. Таким образом, настоящая глава представляет собой переход к главе 11, в которой рассматриваются более сложные, но зато «чистые» задачи распределения.

ОСНОВНЫЕ ФОНДЫ (ПОСТОЯННЫЕ РЕСУРСЫ)

Распределение, как мы выяснили при рассмотрении этого вопроса в главе 7, заключается в разделении имеющихся в наличии материалов, машин, рабочей силы и других ресурсов (например, рабочих площадей) между различными работами, которые требуется произвести. Эти материалы, машины, люди и иные элементы производства представляют собой «общие фонды» («актив») предприятия, поскольку их использование влечет за собой получение некоторой прибыли. В тех случаях, когда эти фонды длительное время сохраняются, предприятием принимается, что они «фиксированы». Однако следует ясно представлять себе, что разница между «фиксированными» активами или основными фондами и оборотными фондами (например, запасом готовой продукции) лишь относительна в смысле трудностей их прибыльной реализации.

Основные фонды составляют большую часть общих фондов (активов) современных промышленных предприятий. Средства, затраченные на их приобретение, обычно возмещаются (посредством амортизационных отчислений и снижения их учетной стоимости) из доходов в течение определенного числа лет. Эти ежегодные отчисления увеличивают стоимость производства всех видов продукции. Однако нельзя допустить, чтобы обычная бухгалтерская процедура по распределению постоянных расходов влияла на распределение и использование основных средств предприятия. Наоборот, их распределение и использование должно производиться таким образом, чтобы максимизировать доходы фирмы.

Допустим, что общая стоимость рассматриваемых основных фондов (основных средств) уже установлена и на ее величину мы никак не можем

влиять. Однако ценность этих фондов возрастает, так как они приносят прибыль, и определение ее весьма важно для решения вопросов о привлечении дополнительных основных фондов. В настоящей главе нахождение оптимального размера основных средств (постоянных ресурсов) и их оптимальное распределение будет проиллюстрировано при решении задачи об определении экономического размера партии при рассмотрении различных условий: (1) ограничения на постоянные ресурсы отсутствуют, (2) ограничен один постоянный ресурс, (3) ограничены два постоянных ресурса. В каждом из этих случаев будет проведено вычисление минимальных издержек производства и стоимости постоянных ресурсов.

В главе 2 и при выводе соответствующего уравнения для определения экономического размера партии в главе 7 мы рассматривали проблему оптимизации независимо от постоянных ресурсов. Рассмотрение этой проблемы здесь мы начнем с построения очень похожей модели, отличие которой от построенной ранее будет заключаться в том, что уравнение цен состоит из двух частей. Предполагается, что получение заказанной продукции происходит равномерно и непрерывно. Ниже мы будем заниматься проблемой распределения основных фондов (постоянные ресурсы).

Экономически выгодный размер партии при отсутствии ограничений. Компания занимается выпуском продукции двух видов: X_1 и X_2 , спрос на которые известен и остается постоянным. Требуется производство каждого вида продукции организовывать отдельными партиями, а не непрерывно, и при этом желательно иметь возможно меньшие суммарные издержки производства. Издержки производства складываются из (1) прямых затрат на материалы и изготовление, (2) расходов, связанных с заказом каждой партии, и (3) расходов на складирование полученной продукции. Следуя предложенному в главе 7 методу, мы получим выражение для суммарных издержек как функцию размера партии для каждого вида продукции.

Пусть L_i — месячная потребность в продукции вида i (предполагается известной и постоянной), C_{i1} — расходы, связанные с заказом партии продукции вида i , C_{i2} — затраты на материалы и изготовление единицы продукции вида i , N_i — число единиц продукции вида i в каждой партии, P — стоимость месячного хранения на складе в процентах от стоимости запасов.

Издержки производства одной партии продукции вида i складываются из расходов, связанных с заказом, и прямых затрат на изготовление $C_{i1} + N_i C_{i2}$. Очевидно, что если спрос на продукцию все время один и тот же, одной партии хватит на N_i/L_i месяцев. Например, если в месяц продается 300 единиц продукции, то партии в 200 единиц хватит на 200/300, или на 2/3 месяца, а партии в 600 единиц продукции хватит на 2 месяца. Следовательно, в среднем в месяц будет заказано L_i/N_i партий. Если партии хватает на 2/3 месяца, то в течение одного месяца будет заказано в среднем 3/2, или 1,5 партии. Если партии хватает на 2 месяца, в среднем за месяц будет заказано 0,5 партии.

Таким образом, сумма ежемесячных издержек производства будет выражаться как

$$\frac{L_i}{N_i} (C_{i1} + N_i C_{i2}), \quad \text{т. е.} \quad \frac{L_i C_{i1}}{N_i} + L_i C_{i2}.$$

Накладные расходы, связанные с запасами, состоят главным образом из налогов, расходов на страхование, процентов на вложенный капитал (предположим здесь, что расходы на складирование являются частью постоянных расходов). Размер вложений в запасы определяется как произ-

ведение среднего числа складываемых единиц продукции на стоимость складирования одной единицы. Так как мы предполагаем, что продажа (т. е. отпуск продукции со склада) производится непрерывно, запасы каждого вида продукции будут изменяться линейно от максимального количества до минимального количества, равного 0 (см. рис. 10.1). Средний уровень запасов на складе будет равен $N_i/2$. Расходы, отнесенные к единице запасов, будут составлять $(C_{i1}/N_i) + C_{i2}$, т. е. средние расходы, связанные с заказом, плюс прямые издержки производства.

Можно записать, таким образом, среднюю стоимость запасов продукции вида i как

$$\frac{N_i}{2} \left(\frac{C_{i1}}{N_i} + C_{i2} \right), \text{ т. е. } \frac{C_{i1} + N_i C_{i2}}{2}.$$

При бухгалтерских расчетах может потребоваться учитывать в стоимости запасов и накладные расходы, однако их величина не зависит от размера партии, и поэтому их рассмотрение нецелесообразно.

Если выразить ежемесячные накладные расходы как процент от средних расходов на складирование, выражение для месячных издержек, связанных с продукцией вида i , дополнится, и, таким образом,

$$K_i = \frac{L_i C_{i1}}{N_i} + L_i C_{i2} + P \frac{C_{i1} + N_i C_{i2}}{2}. \quad (1)$$

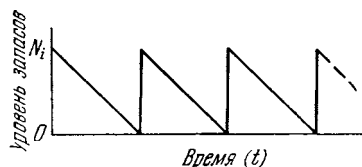


Рис. 10.1. Зависимость уровня запасов от времени.

Поэтому суммарные издержки на все виды продукции, без учета постоянных расходов, определяются как

$$K = \sum_i K_i = \sum_i \frac{L_i C_{i1}}{N_i} + \sum_i L_i C_{i2} + \frac{P}{2} \sum_i N_i C_{i2} + \frac{P}{2} \sum_i C_{i1}. \quad (2)$$

Мы можем предположить, что всевозможные расходы учитываются либо в фиксированных расходах, либо в оплате заказа, либо в прямых издержках производства, либо в накладных расходах. Наша задача состоит в том, чтобы найти величину N_i , при которой минимизируется общая сумма расходов, или, что то же самое, минимизируется величина K . Чтобы можно было воспользоваться методами дифференциального исчисления, предположим, что величина N_i может изменяться непрерывно. Тогда из уравнения (2) мы получим, что

$$\frac{\partial K}{\partial N_i} = -\frac{L_i C_{i1}}{N_i^2} + \frac{P C_{i2}}{2} \quad (3)$$

для всех N_i . Приравняв производную нулю и решая полученное уравнение относительно N_i , находим оптимальное значение N_i , которое мы будем обозначать N_i^* . Имеем

$$N_i^* = \sqrt{\frac{2 L_i C_{i1}}{P C_{i2}}}. \quad (4)$$

При этом значении N_i мы добиваемся минимума суммарных затрат*.

* Если задана функция $f(X_1, X_2, \dots, X_n)$, то условие равенства нулю первых частных производных является лишь необходимым (но не достаточным) для получения максимального или минимального значения. Однако в данном случае нетрудно показать, что значение N_i^* есть именно минимальное значение.

Приводимый ниже численный пример иллюстрирует методику вычисления величины N_i^* . Допустим, относительно продукции вида X_1 и X_2 нам известно следующее:

Продукция	L_i	c_{i1}	c_{i2}
X_1	200	100 долл.	12 долл.
X_2	400	25 долл.	7 долл.

и $P = 0,005$. Тогда

$$N_1^* = \sqrt{\frac{2 \cdot 200 \cdot 100}{0,005 \cdot 12}} = \sqrt{\frac{400 \cdot 100}{0,060}} = \frac{200}{\sqrt{0,060}} = 816,$$

$$N_2^* = \sqrt{\frac{2 \cdot 400 \cdot 25}{0,005 \cdot 7}} = \sqrt{\frac{400 \cdot 100}{0,070}} = \frac{200}{\sqrt{0,070}} = 756.$$

Можно графически представить зависимость издержек производства для каждого вида продукции от размера партии. На рис. 10.2 показано

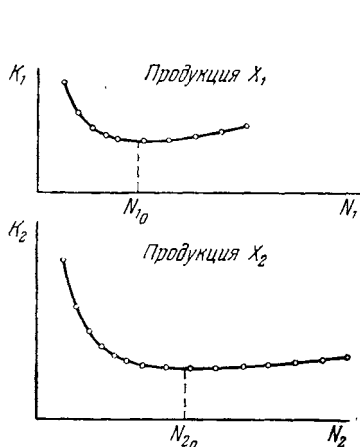


Рис. 10.2. Зависимость издержек производства от размера партии.

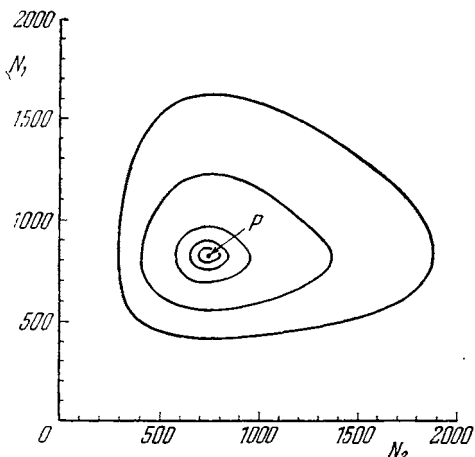


Рис. 10.3. Кривые равной стоимости для выбора размера партий при двух видах продукции.

отношение минимальных издержек к издержкам, получаемым при других размерах партий. Суммарные расходы фирмы при оптимальных размерах партий для продукции X_1 и X_2 составляют

$$K = \frac{200 \cdot 100}{816} + \frac{400 \cdot 25}{756} + 200 \cdot 12 + 400 \cdot 7 + \frac{0,005}{2} [816 \cdot 12 + 756 \cdot 7] + \frac{0,005}{2} \cdot 100 + 25 = \$5275,76.$$

Оптимальный размер партий можно изобразить на двумерном графике, как показано на рис. 10.3, на котором точка P представляет пару чисел 816 и 756, т. е. минимальные суммарные расходы. Замкнутые кривые вокруг точки P соответствуют парам размеров партий, при которых получаются одинаковые суммарные издержки. Эти линии «равных издержек» представляют по мере удаления от точки P все большие и большие издержки.

ЭКОНОМИЧЕСКИ ВЫГОДНЫЙ РАЗМЕР ПАРТИИ ПРИ НАЛИЧИИ ЛИНЕЙНЫХ ОГРАНИЧЕНИЙ

Предположим, что размера партии, при котором получаются минимальные издержки, не удастся достигнуть, так как наши ресурсы ограничены. Допустим, например, что ограничены размеры складских помещений. Как мы уже отмечали выше, средний уровень запасов каждого вида продукции равен половине размера партии. Таким образом, если единица продукции вида i требует для хранения W_i кубических футов, средний объем, занимаемый продукцией вида i , равен $0,5 W_i N_i$. Таким образом, в среднем требуемая суммарная емкость склада равна $0,5 \sum_i W_i N_i$.

На практике часто встречается ситуация, когда объем хранилища невозможно использовать полностью. Следовательно, фактически требующийся объем превышает величину $0,5 \sum_i W_i N_i$. В частности, если продукция хранится в баках или бункерах, каждый из которых предназначен для своего вида продукции, требующийся объем хранилища будет равен $\sum_i W_i N_i$. Однако в данном при-

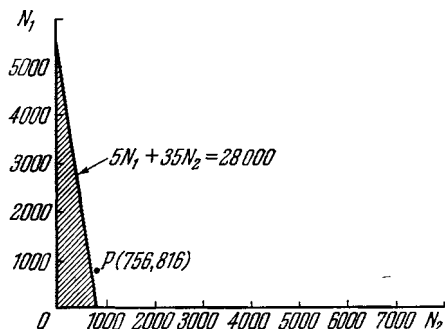


Рис. 10.4. Размеры двух партий различных видов продукции, при которых не превышает емкость склада.

мере мы все же примем объем равным $0,5 \sum_i W_i N_i$. В дальнейшем будет показано, что с минимальными изменениями полученные результаты можно применять и в иных ситуациях. Таким образом, если S есть величина полезного объема склада (за вычетом проходов и т. п.), то мы потребуем, чтобы

$$\sum_i (W_i N_i / 2) \leq S.$$

В частности, если $W_1 = 5$ куб. фут., $W_2 = 35$ куб. фут., $S = 14\,000$ куб. фут., то для нашего примера, рассмотренного выше, мы потребуем, чтобы

$$5 \frac{N_1}{2} + 35 \frac{N_2}{2} \leq 14\,000,$$

или, что то же самое, чтобы

$$5N_1 + 35N_2 \leq 28\,000.$$

Это неравенство вместе с требованием, чтобы размер партий был неотрицательной величиной, определяет область, которая показана на рис. 10.4 в виде заштрихованной площади. Границы этой области задаются уравнениями

$$N_1 = 0, \quad N_2 = 0 \quad \text{и} \quad 5N_1 + 35N_2 = 28\,000.$$

Точка $P(756, 816)$ находится вне заштрихованной области, так как

$$\frac{5 \cdot 816}{2} + \frac{35 \cdot 756}{2} = 15\,270 > 14\,000.$$

Поэтому мы должны найти новую точку, которая лежит в заштрихованной области и соответствует размерам партий, минимизирующих суммарные издержки производства при заданных ограничениях на емкость склада.

Мы будем поступать следующим образом *. Определим величину λ такую, что **

$$\left. \begin{aligned} \lambda < 0, \quad \text{когда} \quad S - \frac{1}{2} \sum W_i N_i = 0, \\ \lambda = 0, \quad \text{когда} \quad S - \frac{1}{2} \sum W_i N_i > 0. \end{aligned} \right\} \quad (5)$$

Тогда $\lambda \left(S - \frac{1}{2} \sum W_i N_i \right)$ тождественно равно нулю. Следовательно, это выражение можно добавить к уравнению (2), не изменяя значения K . Таким образом,

$$K = \sum \frac{L_i C_{i1}}{N_i} + \sum L_i C_{i2} + \frac{P}{2} \sum C_{i1} + \frac{P}{2} \sum N_i C_{i2} + \lambda \left(S - \frac{1}{2} \sum W_i N_i \right). \quad (6)$$

Хотя K не изменилось по величине, частная производная от K по N_i изменяется:

$$\frac{\partial K}{\partial N_i} = -\frac{L_i C_{i1}}{N_i^2} + \frac{P C_{i2}}{2} - \frac{\lambda W_i}{2} \quad \text{для каждого } N_i. \quad (7)$$

Приравнявая производные нулю и находя из полученных уравнений оптимальные N_i , имеем

$$N_i^* = \sqrt{\frac{2 L_i C_{i1}}{P C_{i2} - \lambda W_i}}. \quad (8)$$

Для каждого вида продукции известны все величины L_i , C_{i1} , C_{i2} , W_i и P , кроме пока еще неизвестной величины λ . Однако, если придать величине λ произвольное значение, можно будет вычислить N_i , а следовательно, и $0,5 \sum W_i N_i$ (т. е. средний объем требующихся складских помещений). Если величина $0,5 \sum W_i N_i$ превышает S , значит, размер партии слишком велик. В этом случае последовательно уменьшаем λ и производим новые подсчеты до тех пор, пока не будет достигнуто равенство $0,5 \sum W_i N_i = S$. Если $0,5 \sum W_i N_i$ не превышает S при всех отрицательных λ , принимаем λ равной нулю. Полученные значения N_i обеспечивают фирме минимальные издержки при существующих размерах хранилища.

Если предположить, что можно арендовать складское помещение за D долл. за куб. фут *** в месяц, можно записать уравнение стоимости, в котором учитывается арендная плата за складские помещения:

$$K = \sum \frac{L_i C_{i1}}{N_i} + \sum L_i C_{i2} + \frac{P}{2} \sum C_{i1} + \frac{P}{2} \sum N_i C_{i2} + D \sum \frac{1}{2} W_i N_i. \quad (9)$$

* Далее следует, по существу, изложение метода множителей Лагранжа, применение которого было предложено в неопубликованной работе Бекмана [1]. Другой подход дается в работе Клейна [3].

** Заметим, что $S - \frac{1}{2} \sum W_i N_i < 0$ недопустимо и, следовательно, этот случай можно не рассматривать.

*** Арендная плата за куб. футы складского помещения не входит в величину P . В величину P включены проценты на капитал, налоги, страховые сборы и прочие обложения, которые пропорциональны размеру хранимой партии продукции. D представляет собой расходы на куб. фут складского помещения и складывается из арендной платы, оплаты отопления, освещения и т. п.

Тогда

$$\left. \begin{aligned} \frac{\partial K}{\partial N_i} &= -\frac{L_i C_{i1}}{N_i^2} + \frac{P}{2} C_{i2} + \frac{DW_i}{2}, \\ N_i^* &= \sqrt{\frac{2L_i C_{i1}}{P C_{i2} + DW_i}}. \end{aligned} \right\} \quad (10)$$

Сравнивая уравнения (8) и (10), мы видим, что $-\lambda$ есть величина арендной платы, или условная арендная плата при размере собственных складских помещений S . Поэтому, если $-\lambda$ оказывается больше, чем S , для фирмы становится выгодно арендовать дополнительные хранилища.

Таблица 10.1 *

λ	N_1^*	N_2^*	$\frac{1}{2}(5N_1 + 35N_2)$	λ	N_1^*	N_2^*	$\frac{1}{2}(5N_1 + 35N_2)$
-0,0000	816	756	15.270,0	-0,0005	800	617	12.797,5
-0,0001	813	721	14.650,0	-0,0007	794	580	12.135,0
-0,0002	810	690	14.100,0	-0,0010	784	535	11.322,5
-0,0003	806	663	13.617,5				

* Величины N_i^* подсчитывались по уравнению (8).

Чтобы определить N_1 и N_2 в нашем примере, мы будем до тех пор брать последовательные значения λ , пока потребность в объеме хранилища не дойдет до 14 000 куб. футов. Это показано в таблице 10.1. Мы видим из таблицы, что потребность в объеме равна допустимым 14 000 куб. футам, когда λ примерно равно $-0,0002$. Оптимальные размеры партий, которые соответствуют этой величине, составляют для продукции X_1 и X_2 соответственно 810 и 690 изделий.

Для этих оптимальных размеров партий мы следующим образом можем вычислить месячные расходы фирмы:

$$\begin{aligned} K &= \left[\frac{200 \cdot 100}{810} + \frac{400 \cdot 25}{690} \right] + [200 \cdot 12 + 400 \cdot 7] + \\ &+ \frac{0,005}{2} [810 \cdot 12 + 690 \cdot 7] + \frac{0,005}{2} (100 + 25) = \\ &= (24,69 + 14,49) + 5200 + 36,69 = \$5275,87 \end{aligned}$$

при ограниченном объеме хранилища. Увеличение стоимости по сравнению со случаем, когда ограничения отсутствовали ($K = 5.275,76$ долл.), оказалось пренебрежимо малым, ибо имеющаяся мощность хранилища почти равна той, которая требовалась при отсутствии ограничений. Таким образом, в данном примере оптимальный размер партии почти совпадает с оптимальным, полученным выше.

Рассмотрение примера позволяет разобраться в методике определения оптимального размера партии, когда некоторые переменные ограничены. Мы находим, что условная арендная плата при данном объеме хранилища равна 0,0002 долл. за куб. фут в месяц. Арендовать дополнительные помещения, чтобы можно было при определении размера партии не считаться с ограничениями, экономически целесообразно лишь тогда, когда арендная плата за один куб. фут окажется меньше, чем величина λ , т. е. в данном примере меньше чем 0,0002 долл. за куб. фут.

Хотя ограничения, которые мы рассматривали здесь, касались размеров складских помещений, аналогичные рассуждения имели бы место и при ограничении общей стоимости запасов, размера партии или любой другой линейной функции от размера партии.

ЭКОНОМИЧЕСКИ ВЫГОДНЫЙ РАЗМЕР ПАРТИИ ПРИ НЕЛИНЕЙНЫХ ОГРАНИЧЕНИЯХ

Ограничения, связанные с объемом хранилища, которыми мы занимались в предыдущем параграфе, являются линейными ограничениями, наложенными на размер партии. Это условие определяет прямую, которая, как

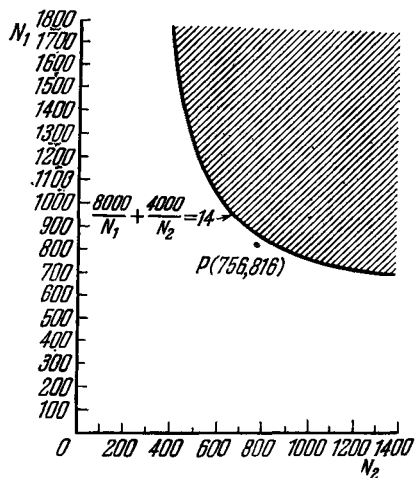


Рис. 10.5. Размеры двух партий, при которых не превышает допустимое время переналадки оборудования.

показано на рис. 10.4, ограничивает размер партии в пределах заштрихованной площади. Другое ограничение, которое обычно влияет на размер партии, связано с лимитированием рабочего времени. Помимо того, что сам заказ партии изделий требует некоторых расходов, на наладку оборудования уходит какое-то время, в течение которого оборудование простаивает. Частые остановки требуют много времени, сокращая время на само производство. Очевидно, что чем меньше партии, тем больше времени уходит на переналадку.

Время, которое теряется на наладку, можно выразить как функцию размера партии. Используя введенные нами обозначения, среднее число заказов на производство продукции типа X_i запишем как L_i/N_i .

Пусть t_i равно времени, которое требуется на подготовку к производству продукции типа X_i . Тогда ожидаемое время, необходимое ежемесячно на подготовительные работы, выражается как $\sum [(L_i t_i)/N_i]$. Таким образом, если T — время, которое можно выделить на подготовку производства (после учета времени, необходимого для производства единиц каждого вида продукции), мы потребуем, чтобы

$$T - \sum \frac{L_i t_i}{N_i} \geq 0.$$

Предположим, что в дополнение к условиям нашего примера нам задано $t_1 = 40$ час. (для продукции типа X_1), $t_2 = 10$ час. (для продукции типа X_2), $T = 14$ час. (время, которое отведено на наладку). Тогда

$$14 - \left[\frac{200 \cdot 40}{N_1} + \frac{400 \cdot 10}{N_2} \right] \geq 0, \quad \text{т. е.} \quad \frac{8000}{N_1} + \frac{4000}{N_2} \leq 14.$$

Ограничение, связанное со временем, выражается как нелинейная функция от размера партии. Это ограничение показано графически на рис. 10.5. Читатель, быть может, заметил, что посредством весьма несложного преобразования можно превратить данную функцию в линейную. Однако ниже мы рассмотрим одновременное действие двух ограничений: линейного — из-за ограниченной емкости хранилища и нелинейного — из-за лимитиро-

ванного времени. Так как оба эти ограничения не могут одновременно быть линейными, мы и в данном случае не будем пользоваться линеаризацией.

Заштрихованная площадь на рис. 10.5 соответствует парам N_1 и N_2 , при которых не нарушаются временные ограничения. Так же как и раньше, мы обнаруживаем, что точка P оказывается вне заштрихованной области. Не учитывая ограничений, связанных с объемом хранилища, которыми мы занимались ранее, постараемся найти, при каких размерах партий, соответствующих заштрихованной области, будут достигнуты минимальные издержки. Проведем следующие операции.

Определим величину μ такую, что

$$\left. \begin{aligned} \mu < 0, \quad \text{когда} \quad T - \sum \frac{L_i t_i}{N_i} = 0; \\ \mu = 0, \quad \text{когда} \quad T - \sum \frac{L_i t_i}{N_i} > 0. \end{aligned} \right\} \quad (11)$$

При этом выражение $\mu \left(T - \sum \frac{L_i t_i}{N_i} \right)$ всегда равно нулю, и, следовательно, может быть добавлено к уравнению стоимости, не изменяя его решения. Таким образом, как и прежде,

$$K = \sum \frac{L_i C_{i1}}{N_i} + \sum L_i C_{i2} + \frac{P}{2} \left(\sum C_{i1} + \sum N_i C_{i2} \right) + \mu \left(T - \sum \frac{L_i t_i}{N_i} \right). \quad (12)$$

Чтобы определить минимальное K , найдем частные производные относительно N_i и приравняем их нулю:

$$\frac{\partial K}{\partial N_i} = -\frac{L_i C_{i1}}{N_i^2} + \frac{P}{2} C_{i2} + \mu \frac{L_i t_i}{N_i^2}.$$

Отсюда

$$N_i^* = \sqrt{\frac{2L_i(C_{i1} - \mu t_i)}{P C_{i2}}} \quad \text{для всех } N_i. \quad (13)^*$$

Для каждого вида продукции известны величины L_i , t_i , C_{i1} , C_{i2} и P , и требуется найти значение μ , при котором $[T - \sum (L_i t_i)/N_i] = 0$.

Таблица 10.2

μ	N_1^*	N_2^*	$\frac{8000}{N_1} + \frac{4000}{N_2} =$ = Общие затраты времени на подготовку производства	μ	N_1^*	N_2^*	$\frac{8000}{N_1} + \frac{4000}{N_2} =$ = Общие затраты времени на подготовку производства
0	816	756	9,80+5,29=15,09	-0,3	864	798	9,26+5,01=14,27
-0,1	832	769	9,62+5,20=14,82	-0,4	879	814	9,10+4,91=14,01
-0,2	848	784	9,43+5,10=14,53	-0,5	894	826	8,94+4,84=13,78

* Величины N_i^* подсчитывались по уравнению (13).

В данном примере мы будем вычислять N_1^* , N_2^* и $[(L_1 t_1)/N_1 + (L_2 t_2)/N_2]$ для различных значений μ . Результаты приведены в таблице 10.2. Заметим, что суммарные потери времени на подготовку производства составляют

* Очевидно, что эти условия вновь только необходимые, но не достаточные. Однако здесь, как и раньше, уравнение (13) дает минимум.

при $u = -0,4$ примерно 14 часов. Этому значению соответствуют размеры оптимальных партий при производстве продукции типа X_1 и X_2 соответственно в 879 и 814 изделий.

Издержки производства при этих размерах партий

$$K = \left[\frac{200 \cdot 100}{879} + \frac{400 \cdot 25}{814} \right] + [200 \cdot 12 + 400 \cdot 7] + \\ + \frac{0,005}{2} [879 \cdot 12 + 814 \cdot 7 + 100 + 25] = \\ = (22,75 + 12,29) + (5200) + 0,0025 (16\,371) = \$5275,97.$$

Изменение издержек и в этом примере получается незначительным. Причины этого такие же, как и прежде.

ЭКОНОМИЧЕСКИ ВЫГОДНЫЙ РАЗМЕР ПАРТИЙ ПРИ НАЛИЧИИ ДВУХ ОГРАНИЧЕНИЙ

Предположим теперь, что приходится одновременно следить за соблюдением ограничений, связанных с заданным объемом склада и с лимитированным временем на настройку. Из-за ограниченности хранилища приходится

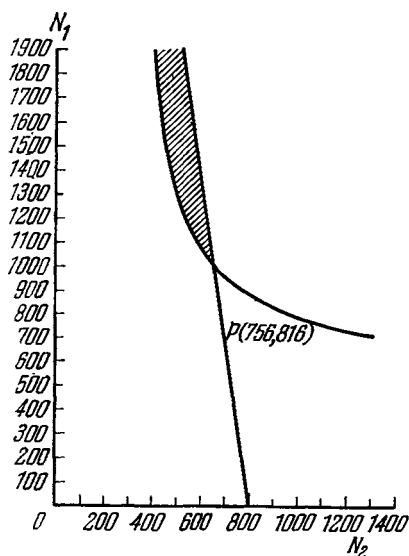


Рис. 10.6. Размеры двух партий, удовлетворяющие одновременно ограничениям на емкость склада и времени переналадки.

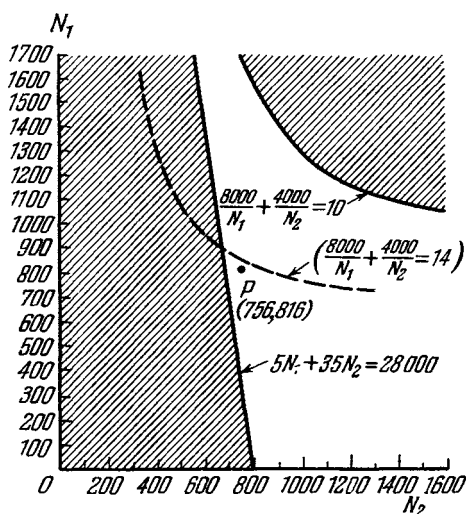


Рис. 10.7. Пример отсутствия решения в задаче с двумя ограничениями.

уменьшать размер партии, что приводит к увеличению времени на подготовительные работы. Точно так же при учете одних лишь временных ограничений происходит увеличение размера партий, а это приводит к увеличению требующейся емкости складских помещений. Потребуем теперь, чтобы размер партий изменялся таким образом, чтобы удовлетворялось и то и другое ограничение. Обоим ограничениям будут удовлетворять лишь те пары величины N_1 и N_2 , которые попадают в заштрихованную область, показанную на рис. 10.6.

Следует особенно подчеркнуть, что решение этой задачи существует не всегда, т. е. кривые, показанные на рис. 10.6, могут и не пересекаться.

Может оказаться, что увеличение спроса на какой-либо вид продукции приводит к увеличению времени на его производство, т. е. к сокращению подготовительного времени. Такая ситуация показана на рис. 10.7. Кривая временных ограничений сдвигается вправо. В результате сокращения времени на подготовку приходится увеличивать размер партий, а это связано с увеличением запасов. Поэтому увеличение производства невозможно без дополнительного оборудования или увеличения емкости складов. Последнее может означать или расширение объема хранилища, или просто введение амортизации для увеличения капиталовложений на организацию хранения.

В этом примере можно, как и раньше, составить уравнение с дополнительными членами, тождественно равными нулю. Получим

$$K = \sum \frac{L_i C_{i1}}{N_i} + \sum L_i C_{i2} + \frac{P}{2} \left(\sum N_i C_{i2} + \sum C_{i1} \right) + \lambda \left(S - \sum \frac{W_i N_i}{2} \right) + \mu \left(T - \sum \frac{L_i t_i}{N_i} \right). \quad (14)$$

Вычисляя частные производные, приравнявая их к нулю и решая уравнения относительно N_i , мы получим

$$N_i^* = \sqrt{\frac{2L_i(C_{i1} - \mu t_i)}{PC_{i2} - \lambda W_i}}. \quad (15)$$

Пользуясь этим уравнением, мы можем подсчитать N_1^* и N_2^* для различных значений μ и λ . Полученные результаты приведены в таблицах 10.3 и 10.4.

Таблица 10.3

Оптимальный размер партий (N_1^*) для продукции вида X_1
при выбранных значениях μ и λ

$\begin{matrix} -\lambda \\ -\mu \end{matrix}$	0,000	0,001	0,002	0,003	0,004	0,005	0,006
0	816	784	756	730	707	686	667
1	966	928	894	864	837	812	789
2	1095	1052	1014	980	949	920	894
3	1211	1164	1121	1083	1049	1017	989
4	1317	1265	1219	1178	1140	1106	1075
5	1414	1359	1309	1265	1225	1188	1155
6	1506	1446	1394	1346	1304	1265	1229

Таблица 10.4

Оптимальный размер партий (N_2^*) для продукции вида X_2
при выбранных значениях μ и λ

$\begin{matrix} -\lambda \\ -\mu \end{matrix}$	0,000	0,001	0,002	0,003	0,004	0,005	0,006
0	756	535	436	378	338	309	286
1	894	632	516	447	400	365	338
2	1014	717	586	507	454	414	383
3	1121	793	647	561	501	458	424
4	1219	862	704	609	545	498	461
5	1309	926	756	655	586	535	495
6	1394	986	805	697	623	569	527

На рис. 10.8 нанесены точки, соответствующие парам, составленным из размеров партий. Следуя от любой из этих точек пересечения вправо по линии, можно определить соответствующие значения μ , а следуя по линии вниз, определить значение λ . Пары, которые находятся в заштрихованном сегменте на рис. 10.8, удовлетворяют сразу обоим ограничениям. Из рисунка видно, что мы не попадем в заштрихованную область, если $\mu = 0$ или $\lambda = 0$. Следовательно, обе величины должны быть отрицательными, и поэтому

$$T - \sum \frac{L_i t_i}{N_i} = 0, \quad S - \sum \frac{W_i N_i}{2} = 0.$$

Эти условия выполняются на пересечении линий, ограничивающих заштрихованную область. Из рис. 10.8 видно, что эта точка получается приблизительно при $N_1 = 1015$ и $N_2 = 655$.

Проверяя наши результаты, убеждаемся, что при этом требуется емкость хранилища

$$\frac{1015}{2} \cdot 5 + \frac{655}{2} \times 35 = 14\,000 \text{ куб. футов,}$$

а время на подготовку производства составит

$$\frac{8000}{1015} + \frac{4000}{655} =$$

$$= 7,88 + 6,11 = 13,99 \text{ часа.}$$

Рис. 10.8. Размеры двух партий для некоторых значений μ и λ .

Следовательно, размер партий, минимизирующих издержки и все еще не приводящий к нарушению ограничений, равняется

$$N_1^* = 1015, \quad N_2^* = 655.$$

Из уравнения стоимости для этих размеров партий можно найти средние значения месячных издержек. Имеем

$$\begin{aligned} K &= \frac{200 \times 100}{1015} + \frac{400 \times 25}{655} + (200 \times 12 + 400 \times 7) + \\ &\quad + \frac{0,005}{2} (1015 \cdot 12 + 655 \cdot 7 + 125) = \\ &= 19,70 + 15,27 + 5200 + 42,23 = \$5277,20. \end{aligned}$$

По значениям N_1^* и N_2^* можно вычислить с помощью уравнения (15) неявные величины λ и μ . Получаем

$$(1015)^2 = \frac{2 \cdot 200 (100 - 40\mu)}{0,005 \times 12 - 5\lambda}, \quad (655)^2 = \frac{2 \cdot 400 (25 - 10\mu)}{0,005 \times 7 - 35\lambda},$$

или

$$\begin{aligned} (1015)^2 (0,060 - 5\lambda) &= 40\,000 - 16\,000\mu, \\ (655)^2 (0,035 - 35\lambda) &= 20\,000 - 8000\mu. \end{aligned}$$

Решая эти два уравнения относительно λ и μ , находим

$$\lambda = -0,001278, \quad \mu = -1,77576.$$

Подсчитанные значения λ и μ заключены соответственно между $-0,001$ и $-0,002$ и между -1 и -2 . Это видно и из рис. 10.8.

Можно подытожить результаты, полученные при рассмотрении всех четырех примеров (см. табл. 10.5), и сделать следующие основные выводы о влиянии ограничений на размер партии, издержки и величину активов.

Таблица 10.5

Сравнение четырех примеров

Условия	N_1^*	N_2^*	K , долл	λ	μ	$\frac{1}{2} \sum N_i W_i$	$\sum \frac{L_i t_i}{N_i}$
Ограничения отсутствуют	816	756	5275,76	0	0	15 270	15,09
Ограниченная емкость склада	810	690	5275,87	$-0,0002$	0	14 100	15,67
Ограниченное время на подготовку производства	879	814	5275,97	0	$-0,40$	16 443	14,02
Ограничения обоих типов	1015	655	5277,20	$-0,0013$	1,78	14 000	13,99

1. По сравнению со случаем, когда ограничения отсутствуют, при ограниченном объеме хранилища размеры партий уменьшаются, в то время как при ограничении подготовительного времени размеры партий возрастают. Комбинированные ограничения приводят к увеличению размера партии для одного вида изделий и к уменьшению для другого.

2. Любое ограничение приводит к увеличению расходов. При комбинированном ограничении возрастание расходов еще значительнее.

3. Стоимость активов (складов и машинного оборудования) при необходимости удовлетворять сразу двум ограничениям больше, чем при одном ограничении.

Интересно отметить, что мы рассмотрели случай, когда ни один из активов не допускает рассмотрения оптимальных размеров партий без учета ограничений. Однако можно получить решение и рассчитывая на аренду дополнительных помещений или приобретение оборудования, причем издержки производства не будут существенно отличаться от минимальных издержек при отсутствии ограничений. Неявная стоимость хранилищ составляет 0,13 центов за куб. фут в месяц, а неявная стоимость машинного времени — 1,78 долл. в час на месяц. На эти величины можно опираться при определении целесообразности аренды дополнительных помещений и/или дополнительного оборудования.

Ряд теоретических проблем возникает при использовании данного метода. Эти проблемы становятся весьма серьезными, когда число ограничений или видов продукции возрастает. Одна из этих проблем играет при решении особенно существенную роль (вопрос о самом существовании решения. — *Прим. перев.*). Выше было показано, что при некоторых ограничениях решения вообще не существует. В частности, мы показали, что ограничения машинного времени и емкости склада могут быть такими, что ни при какой комбинации размеров партий не удастся выпускать требуемое количество продукции и одновременно не нарушать оба ограничения (см. рис. 10.7).

Если известно, что решение существует, возникает другая теоретическая проблема, связанная с отысканием этого оптимального решения. Метод проб и ошибок, который применялся в наших примерах, можно улучшить, применяя такие методы, в которых условия очередного испытания определяются результатами предыдущего расчета.

Один из методов, который, возможно, окажется перспективным, изложен в статье Крокетта и Чернова [2]. В статьях Бекмана [1], Куна и Такера [4] и Слэтера [5] рассматриваются необходимые и достаточные условия оптимальности решения. Эти условия могут оказаться ключом к более совершенным методам нахождения решений.

Трудные теоретические проблемы, ждущие своего решения, не умаляют практической ценности методов, описанных в настоящей главе. Они могут оказаться полезными во многих задачах, когда число ограничений не очень велико.

ЛИТЕРАТУРА

1. Beckmann Martin, A Lagrangian Multiplier Rule in Linear Activitz Analysis and Some of Its Applications, Cowles Commission Discussion Paper: Economics, no. 2054 (unpublished), Nov. 5, 1942.
2. Crockett Jean Bronfenbrenner and Chernoff Herman, Gradient Methods of Maximization, Pac. J. Math., 5 (1955).
3. Klein Bertram, Direct Use of Extremal Principles in Solving Certain Optimizing Problems Involving Inequalities, J. Oper. Res. Soc. Amer., 3, no. 2, 168—175 (May 1955).
4. Kuhn H. W. and Tucker A. W., Nonlinear Programming, in Jerzy Neyman (ed.), Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, pp. 481—492, 1951.
5. Slater Morton, Lagrange Multipliers Revisited, Cowles Commission Discussion Paper: Mathematics, no. 403 (unpublished), Nov. 7, 1950.

Модели распределения используются для решения задач, в которых (а) нужно выполнить несколько операций, причем возможны различные способы их выполнения, и (б) не хватает ресурсов или оборудования для выполнения каждой из этих операций оптимальным образом. Задача заключается в том, чтобы выполнить операции с учетом ресурсов наиболее эффективным путем. Эти задачи разделяются на два типа:

1. Задан объем работ, которые необходимо выполнить. Имеются определенные ресурсы, т. е. фиксированы производственные мощности и количество материалов для выполнения работ, что представляет собой наложенные ограничения. Задача заключается в том, чтобы, используя эти ограниченные мощности машин и наличные материалы, выполнить все работы наиболее рациональным образом.

2. Заданы определенные материалы или оборудование, а иногда заданы и материалы и оборудование. Задача заключается в том, чтобы определить, какая работа дает максимальный доход при использовании имеющегося оборудования и материалов.

Основным аппаратом для решения задач распределения является математическое программирование, и в частности линейное программирование, и анализ производственной деятельности.

Важнейший вклад в эти методы, характеризуемый Орденом [36] как «научные основы», был внесен Купмансом* и его коллегами из комиссии Коулса [28, 29]. К этим исследованиям тесно примыкают работы в области межотраслевого баланса, начиная с трудов Леонтьева.

После появления симплексного метода, разработанного Данцигом [12], стало возможным практическое использование линейного программирования. Чарнс и Лемке [2, 6] видоизменили метод с целью устранения ошибок вследствие «зацикливания». Данциг и другие [11, 14—16] разработали обобщенный, или «модифицированный», симплексный метод получения решения в расходящихся (или «вырожденных») случаях. Следует также упомянуть о кратком изложении симплексного метода, опубликованном Чарнсом, Купером и Гендерсоном [4]. Симплексный метод рассмотрен в 11-й главе.

Важный подкласс распределительных задач составляют задачи, в которых потребности и запасы представляют собой однородный продукт (например, товарные вагоны или тонны стали). Транспортный метод для решения этого подкласса задач (в которых выполняются линейные ограничения) был разработан Хичкоком [27] в 1941 году и позднее независимо от него Купмансом в 1947 году. Совсем недавно Двайер [18] разработал новый метод решения транспортной задачи Хичкока, в котором используется прием сокращения размерности матриц. Метод Двайера значительно сокращает объем вычислений по сравнению с методом Хичкока и Купманса. Форд и Фулкерсон [21] недавно опубликовали метод решения транспортной задачи, основанный на теории потоков. Этот метод также упрощает расчетный аппарат. Видал [39] предложил графический способ решения транспортных задач, в которых потребности полностью соответствуют запасам. Транспортный метод приводится в 11-й главе.

* Первым математиком, заложившим основы линейного программирования, является советский математик Л. В. Канторович. К сожалению, иностранные авторы редко упоминают об этом. (Прим. перев.)

Теорема «двойственности» в линейном программировании была впервые доказана Гэйлом, Куном и Таккером [23], дополнительные разработки были проведены Рубинём, фон Нейманом, Орде́ном и Данцигом. В этой теореме доказывается, что каждой задаче максимизации в линейном программировании соответствует симметричная задача минимизации, в которую входят те же самые данные, что и в исходную задачу. Эта теорема облегчает решение некоторых практических задач и дает математикам возможность решать задачи из других разделов математики, которые на первый взгляд кажутся несходными. Например, Гофман использовал теорему двойственности для доказательства теоремы Дилворта о частично упорядоченных множествах в топологии. Фулкерсон и Данциг [22] использовали теорему для определения максимального потока в сети. Теорема двойственности приведена в приложении 11В.

Другой частной задачей линейного программирования является задача о назначении, в которой для каждой операции необходим один (и только один) вид оборудования или определенный материал. Хотя исследования этой задачи и ее модификаций были проведены Кёнигом еще в 1916 году, однако интерес к этой задаче появился только после недавних работ Вото и Ордена [40], Флада [19], Куна [31] и других. Работа Куна основана на некоторых идеях из теории матриц, разработанной Эгервари в 1931 году. Флад [20] разработал упрощенный вариант метода Куна — Эгервари, который был опубликован, и некоторые другие алгоритмы решения, которые еще не опубликованы. Задача о назначении приводится в 12-й главе.

Были разработаны программы для решения задач с помощью симплексного и транспортного методов, а также программы решения задачи о назначении на машине IBM. Чарнс и Купер разработали и использовали различные приемы вычислений, которые учитывают специальные свойства заданной матрицы и дают возможность решать задачи большой размерности. Данциг [13] разработал упрощенные численные методы решения задач, матрицы которых являются треугольными.

Чарнс и Лемке [7] рассматривали задачи, не ограниченные требованиями линейности. Они разработали «расширенный симплексный метод» получения решений с определенной степенью точности для задач, включающих определенные нелинейные функции (нелинейные выпуклые полиномы).

Много работ посвящено использованию методов линейного программирования для решения промышленных и военных задач. В 13-й главе рассматриваются трудности, связанные с решением некоторых задач с помощью методов линейного программирования.

Глава 11

ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

ВВЕДЕНИЕ

Линейное программирование является одним из методов решения общих задач оптимизации, в которых учитывается большое число переменных, подчиненных определенным ограничениям. При решении этих задач необходимо получить оптимальное (экстремальное) значение определенного критерия эффективности, например прибылей, затрат, количества произведенных продуктов или других показателей, при условии, что удовлетворяются поставленные ограничения. Эти ограничения в свою очередь носят различный характер и объясняются условиями производства, управления, сбыта, хранения, наличием сырья или законодательными положениями.

Например, на металлургическом предприятии определенные стальные изделия получают из различного сырья путем горячей прокатки, холодной прокатки, отжига, нормализации и различных операций по резке заготовок. Чтобы получить оптимальное решение, нужно перебрать все возможные комбинации (и даже перестановки) этих операций и материалов. При этом определяется оптимальная величина прибылей (или затрат *, или количества изготовленной продукции) при наличии ограничений, учитывающих характер процесса, например: (1) ограничения по производственной мощности действующего оборудования; (2) ограничения, учитывающие минимальное количество материалов, необходимых для получения каждого продукта; (3) ограничения по качеству и количеству выпущенной продукции; (4) требования заказчика; (5) ограничения, учитывающие наличный запас топлива; (6) ограничения, учитывающие наличный запас сырья.

РЕШЕНИЕ ЗАДАЧ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

В математике давно были известны общие методы решения задач, учитывающих большое количество взаимосвязанных переменных. Некоторые из этих методов, например методы решения систем уравнений, были первоначально разработаны математиками и впоследствии нашли широкое применение в таких областях, как экономика, техника, физика, биология и статистика. Кроме того, сфера применения этих методов значительно расширилась благодаря успехам в области вычислительной техники, особенно в связи с созданием быстродействующих электронных вычислительных машин. В задачах линейного программирования ограничения обычно формулируются в выражениях «не больше чем», «не меньше чем», «по крайней мере», «самое большее» и, следовательно, обычно представляются в виде неравенств или систем неравенств. Поэтому методы решения систем уравнений обычно неприменимы к решению задач линейного программирования **.

* Во многих случаях задачи максимизации прибыли и минимизации затрат существенно различны. И только в тех случаях, когда количество выпускаемой продукции постоянно или рыночные цены прямо пропорциональны затратам, задачи совпадают.

** В 10-й главе уже рассматривался один метод решения определенных задач оптимизации, в которых ограничения заданы в виде неравенств, т. е. модифицированный метод множителей Лагранжа. Обычный метод множителей Лагранжа неприменим к задачам, представленным в этой главе, потому что он не гарантирует неотрицательности переменных. Модифицированный метод множителей Лагранжа, рассмотренный в гл. 10, применим практически только для решения задач небольшого масштаба. Методы, приведенные в этой главе, эффективны для решения задач большого масштаба и гарантируют неотрицательность переменных, но они применимы только к линейным задачам.

Однако за последние несколько лет были разработаны итеративные методы решения системы неравенств и выбора наилучшего, или оптимального (одного или нескольких), решения из множества возможных решений на основании заданной функции цели. Наиболее важными методами решения таких задач являются метод решения транспортной задачи и симплексный метод. Прежде чем перейти к рассмотрению этих методов, рассмотрим некоторые задачи, которые сводятся к задачам линейного программирования.

КОГДА МОЖНО ПОЛЬЗОВАТЬСЯ ЛИНЕЙНЫМ ПРОГРАММИРОВАНИЕМ?

Линейное программирование можно использовать для решения задач оптимизации, в которых выполняются следующие условия:

1. Необходимо наличие линейной * функции цели (например, прибыли, затраты или количество произведенных продуктов), экстремум которой необходимо отыскать.

2. Необходимо указать ограничения, в пределах которых может достигаться экстремум функции цели **; эти ограничения следует задавать в виде системы линейных равенств или неравенств. Например, если требуется максимизировать прибыль, полученную при производстве определенных продуктов, с учетом ограничений по мощности действующего оборудования, то необходимы следующие данные:

1. Прибыль в расчете на деталь, произведенную на каждой машине.

2. Машинное время, необходимое для обработки детали на каждой машине.

3. Общее время работы каждой машины по изготовлению данной продукции.

4. Планируемое количество единиц изделий каждого типа.

Используя эти данные, следует выразить суммарную прибыль в виде линейной функции, а ограничения, наложенные на мощности оборудования, в виде системы линейных неравенств. Если задача поставлена правильно, то можно использовать методы линейного программирования для ее решения.

Примеры задач линейного программирования.

Задача I. Первыми задачами линейного программирования были транспортные задачи. Требуется минимизировать транспортные издержки на перевозку однородного продукта из n пунктов производства, расположенных в различных районах, в m разных пунктов потребления. Заданы стоимости перевозок и хранения груза, а также требования заказчика.

Один из вариантов этой задачи будет далее поставлен и решен в качестве иллюстрации одного из методов линейного программирования.

Задача II. Вторая задача связана с составлением смеси авиационного горючего [5]. Задаются цены и спецификации определенных сортов промышленного авиационного горючего в виде минимально возможного

* Требование линейности существенно для применения методов, изложенных в этой главе. Линейность означает, например, что для получения 10 изделий потребуется в 10 раз больше средств, чем для получения одного изделия, или на изготовление 10 изделий уйдет в 10 раз больше времени, чем на изготовление одного изделия, и т. д. Если же такое допущение неверно или нельзя получить линейную функцию за счет преобразования переменных, то методы линейного программирования неприменимы. Следует, правда, отметить, что были предприняты попытки решить задачи с кусочно-линейной функцией, а также задачи нелинейного программирования, особенно квадратичного.

** Ограничения накладываются обычно не на значения функции цели, а на значения ее аргументов. (Прим. ред.)

октанового числа, максимального давления паров и максимально допустимой концентрации тетраэтила свинца. Заданы также цены и химические свойства конечных продуктов и предельное количество исходных материалов, которые используются для получения различных сортов горючего. Задача заключается в определении такой комбинации исходных материалов, чтобы (а) получить максимальный доход и (б) исключить необходимость создания дополнительных емкостей для хранения.

Задача III [24]. Требуется заключить контракты с производителями на комплектную поставку различных изделий в заданном количестве по каждому изделию. Изделия считаются комплектными в том смысле, что они могут быть изготовлены одним и тем же производителем, а его возможности производства одного вида изделий зависят от того, сколько изделий других видов он должен изготовить по контракту. Вследствие ограниченных возможностей производителя увеличение производства изделий одного вида вызывает уменьшение производства изделий другого вида. Производитель может предложить поставить изделия одного вида или всех видов, требующихся заказчику. Поставщик должен указать:

- (а) цену каждого изделия, которое он может поставить;
- (б) максимальное и минимальное количество изделий каждого вида, которые он может поставить;
- (в) максимальный и минимальный объем всего контракта, который он намерен заключить. Эти условия можно выразить различными способами, например:

1. Максимальная (и минимальная) общая стоимость контракта в долларах.
2. Максимальное (и минимальное) общее количество изделий всех видов.
3. Частичная поставка каждого определенного изделия в процентах от количества, указанного в пункте (б), не больше 100% при условии, что суммарное количество всех изделий не превосходит 100%. Подразумевается, что ограничения, наложенные на каждое изделие, пропорциональны сложности его изготовления.
4. Поставщик может предложить поставить изделия нескольких видов, но контракт заключается на изделия только одного вида.

Задача состоит в заключении таких контрактов с различными поставщиками, чтобы удовлетворялись следующие условия: (а) общие затраты на приобретение всех необходимых изделий должны быть минимальными; (б) количество единиц каждого изделия, изготовленного всеми производителями, должно быть не меньше потребности в этом изделии.

Другие задачи. Для иллюстрации второго из двух рассматриваемых методов линейного программирования ниже приведена также четвертая задача. Кроме того, с помощью линейного программирования можно решать ряд других задач, в том числе (1) задачи о назначении персонала [37], (2) задачи определения оптимальной системы севооборота [26], (3) задачи об оптимальном распределении произведенной продукции [1, 9], (4) задачи построения оптимальных планов воздушных налетов на противника, (5) задачи проектирования систем вооружения, (6) задачи выбора оптимальной стратегии снабжения.

МЕТОД РЕШЕНИЯ ТРАНСПОРТНЫХ ЗАДАЧ

Чтобы практически разобрать методы линейного программирования, не затрагивая сложных математических сторон этих методов, рассмотрим аналитические методы решения задач на конкретных примерах. Рассмотрим

прежде всего метод решения транспортной задачи * на примере транспортной задачи на железной дороге, в которой отыскивается оптимальное распределение товарного порожняка. Из-за недостатка места рассматривается упрощенная задача. Кроме того, подобрана такая задача, что результаты, полученные аналитическим путем, можно проверить с помощью непосредственных вычислений.

Задача заключается в определении такого плана перегона порожних товарных вагонов из пунктов отправления (с «излишком» вагонов) в пункты назначения (с «недостатком» вагонов), чтобы суммарные затраты на перегон были минимальными при условии, что выполняются все наложенные ограничения.

В таблицах 11.1 и 11.2 приведены условия задачи и удельные (на один товарный вагон) затраты на перегон.

Таблица 11.1

Условия задачи

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Избыток
S_1	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	9
S_2	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	4
S_3	X_{31}	X_{32}	X_{33}	X_{34}	X_{35}	8
Потребность	3	5	4	6	3	21

Таблица 11.2

Удельные затраты

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5
S_1	C_{11} —10	C_{12} —20	C_{13} —5	C_{14} —9	C_{15} —10
S_2	C_{21} —2	C_{22} —10	C_{23} —8	C_{24} —30	C_{25} —6
S_3	C_{31} —1	C_{32} —20	C_{33} —7	C_{34} —10	C_{35} —4

* Транспортный метод применим к решению подкласса задач линейного программирования, в которых потребности и запасы выражаются в одних и тех же единицах.

В таблице 11.1 указано, что в пунктах S_1 , S_2 и S_3 имеются излишки порожних товарных вагонов в количествах 9, 4 и 8 вагонов соответственно, в пунктах D_1 , D_2 , D_3 , D_4 и D_5 не хватает соответственно 3, 5, 4, 6 и 3 вагонов. (Для простоты предполагается, что задача сбалансирована, т. е. суммарное количество излишних вагонов равно суммарному количеству недостающих вагонов *.)

В таблице 11.2 приведены удельные затраты C_{ij} ** на перегон одного порожнего вагона из пункта отправления i в пункт назначения j (согласно общепринятым правилам i означает строку и j означает столбец). Таким образом, требуется определить такие значения X_{ij} ($i = 1, 2, 3$; $j = 1, 2, 3, 4, 5$) таблицы 11.1, чтобы (1) удовлетворить поставленным ограничениям и (2) минимизировать суммарные затраты ***.

Получение исходного опорного плана. Первым шагом при решении транспортной задачи является получение допустимого плана, т. е. возможного решения, которое удовлетворяет поставленным ограничениям. Если при этом получены минимальные суммарные затраты, то такой план называется оптимальным планом или решением, в данном случае это минимальный план. Исходный план можно легко построить, используя алгоритм, разработанный Данцигом [10] и названный Чарнсом и Купером [8] «правилом северо-западного угла».

«Правило северо-западного угла» формулируется следующим образом:

1. Начать с северо-западного угла исходной таблицы неизвестных (табл. 11.1) и сравнить количество единиц продукта, имеющегося в пункте отправления S_1 , с количеством единиц продукта, потребного в пункте назначения D_1 .

(а) Если $D_1 < S_1$, т. е. если количество продукта, потребного в пункте D_1 , меньше количества продукта, имеющегося в пункте S_1 , то положить $X_{11} = D_1$ и перейти к переменной X_{12} (т. е. к следующей переменной по горизонтали).

(б) Если $D_1 = S_1$, то положить $X_{11} = D_1$ и перейти к переменной X_{22} (т. е. по диагонали).

(в) Если $D_1 > S_1$, положить $X_{11} = S_1$ и перейти к переменной X_{21} (т. е. по вертикали).

2. Продолжать этот процесс шаг за шагом от северо-западного до юго-восточного угла таблицы.

Таким образом, в нашем примере (см. табл. 11.3) процесс определения плана происходит следующим образом:

1. Положим $X_{11} = 3$, т. е. минимальной величине из S_1 (9) и D_1 (3).

2. Переходим к X_{12} (правило (а)). Сравниваем количество вагонов, оставшихся в пункте S_1 (т. е. 6), с количеством вагонов, требующихся в пункте D_2 (5), и соответственно полагаем $X_{12} = 5$.

3. Переходим к X_{13} (правило (а)). Сравниваем оставшееся количество вагонов в S_1 (1) и количество вагонов, требующееся в D_3 (4). Положим $X_{13} = 1$.

* Это не обязательное условие, так как любая транспортная задача может быть сведена к сбалансированной путем введения фиктивных пунктов отправления или назначения (см. [38, гл. 9]).

** Затраты взяты с отрицательными знаками, так как транспортные издержки представляют собой отрицательную прибыль.

*** Обозначение X_{ij} ($i = 1, 2, 3$; $j = 1, 2, 3, 4, 5$) относится к 15 возможным X_{ij} , которые получаются при значениях $i = 1, 2$ и 3 и $j = 1, 2, 3, 4$ и 5. Эти 15 X_{ij} представлены в таблице 11.1.

4. Переходим к X_{23} (правило (в)). Определяем $X_{23} = 3$.

5. Продолжая этот процесс, получим $X_{24} = 1$, $X_{34} = 5$ и наконец в юго-восточном углу получим $X_{35} = 3$.

Таблица 11.3

Первый опорный план

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Суммарные запасы
S_1	3	5	1			9
S_2			3	1		4
S_3				5	3	8
Суммарные потребности	3	5	4	6	3	21

Исходный план задачи, полученный с помощью правила северо-западного угла, приведен в таблице 11.3 в виде возможных значений X_{ij} , выделенных жирным шрифтом. Легко проверить, что этот набор значений X_{ij} является планом, так как удовлетворяются ограничения по каждой строке и по каждому столбцу. Суммарные затраты всех перевозок для этого плана можно получить, умножая выделенные значения X_{ij} таблицы 11.3 на соответствующие коэффициенты C_{ij} из таблицы 11.2 и суммируя полученные произведения. При этом получим *:

$$\text{Суммарные затраты} = \sum_{j=1}^5 \sum_{i=1}^3 C_{ij} X_{ij} = \sum_{i=1}^3 \sum_{j=1}^5 C_{ij} X_{ij}.$$

В нашем примере суммарные затраты, соответствующие исходному плану, определяются следующим образом:

$$\begin{aligned} \text{Суммарные затраты} &= X_{11}C_{11} + X_{12}C_{12} + X_{13}C_{13} + X_{23}C_{23} + X_{24}C_{24} + \\ &+ X_{34}C_{34} + X_{35}C_{35} = (3)(-10) + (5)(-20) + (1)(-5) + (3)(-8) + \\ &+ (1)(-30) + (5)(-10) + (3)(-4) = -\$251. \end{aligned}$$

Знак минус означает, что получаем убытки, а не прибыль.

Проверка плана на оптимальность. Когда построен план и определена соответствующая ему суммарная стоимость, необходимо определить, является ли этот план оптимальным. Чтобы определить, является ли полученный план оптимальным, необходимо «оценить» различные варианты, т. е. нужно оценить стоимости выигрыша (проигрыша), связанные с неиспользованием клеток, в которых нет выделенных чисел. Такая оценка

* Следует отметить, что все невыделенные значения X_{ij} принимаются равными нулю.

дается в таблице 11.4 (только для невыделенных чисел) и относится к плану, приведенному в таблице 11.3. Эта оценка получается так *:

1. Для любой клетки, в которой нет выделенных чисел, проведем контур следующим образом. Соединим данную клетку с ближайшей выделенной жирным шрифтом клеткой в той же строке, причем выбирается такая клетка, в одном столбце с которой расположена другая выделенная клетка. Например, в таблице 11.4, если начать с клетки S_3D_1 (строка 3, столбец 1), то

Таблица 11.4

Первый опорный план (с оценками): $C = 251$

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	3 ← — 5 — — 1 — — 18 — — 11					9
S_2	— 11 — — 13 — — 3 ← — — 1 — — 18					4
S_3	— — 8 — — 17 — — 19 — — 5 — — 3					8
Итого	3	5	4	6	3	21

число **5** в S_3D_4 (строка 3, столбец 4) удовлетворяет поставленным требованиям, т. е. это ближайшее выделенное число в третьей строке, причем другое выделенное число **1** в S_2D_4 находится в том же столбце (столбец 4). (Ясно, что выделенное число **3** в клетке S_3D_5 не удовлетворяет этим требованиям.)

2. Выполнить горизонтальное и вертикальное перемещения указанным в таблице образом. В рассматриваемом примере — от S_3D_1 к S_3D_4 и S_2D_4 (см. табл. 11.4).

3. Прodelав горизонтальное и вертикальное перемещения, повторить шаги 1 и 2 указанного процесса. Например, в нашем случае соответственно получим клетки S_2D_3 и S_1D_3 (перемещение от **1** в S_2D_4 в **1** в S_1D_3 через **3** в S_2D_3).

4. Продолжаем описанный процесс, передвигаясь от одной выделенной клетки до другой с помощью горизонтальных и вертикальных перемещений до тех пор, пока с помощью горизонтального перемещения не попадаем в тот столбец, в котором находится исходная оцениваемая клетка. (При этом процессе следует использовать минимально возможное количество шагов.)

В нашем примере процесс заканчивается перемещением из **1** S_1D_3 в **3** S_1D_1 .

5. Наконец, перемещение в клетку, для которой ищется оценка (S_3D_1). При этом заканчивается контур, построенный для оценки определенной клетки. (Следует отметить, что для получения оценки определенной клетки не столько важно начать контур в этой клетке, сколько закончить его.)

6. Образовать сумму из удельных затрат пройденных клеток,

* Другой метод оценки этих клеток см. в приложении 11А.

приписав им чередующиеся знаки плюс и минус (затраты приведены в табл. 11.2). Полученное число соответствует оценке невыделенных клеток таблицы 11.4. Например, для оценки клетки S_3D_1 получаем:

Контур табл. 11.4	S_3D_4	S_2D_4	S_2D_3	S_1D_3	S_1D_1	S_3D_1
Удельные затраты табл. 11.2	-10	-30	-8	-5	-10	-1
Оценки (S_3D_1)	$+(-10)$	$-(-30)$	$+(-8)$	$-(-5)$	$+(-10)$	$-(-1) = +8$

Таким образом, получаем +8 в клетке S_3D_1 таблицы 11.4.

7. Продолжаем указанный процесс до тех пор, пока не будут получены оценки всех невыделенных клеток.

Итеративный процесс получения оптимального плана. Закончив оценки таблицы, можно определить, является ли исходный план оптимальным или нет. Исходный план будет оптимальным, если все невыделенные числа (оценки) неотрицательны. Если одно или больше невыделенных чисел отрицательны, то возможно дальнейшее улучшение плана в смысле улучшения значения функции цели * (см., например, отрицательные числа в S_1D_4 , S_2D_2 и т. д. в табл. 11.4).

Улучшение плана достигается с помощью процесса последовательного приближения.

(1) Из отрицательных чисел (оценок) выбрать наибольшее по величине **, например $-N$.

(2) Проследить еще раз контур, соответствующий этому числу.

(3) Из выделенных чисел этого контура, которым приписан знак плюс в образованной сумме, выбрать наименьшее по величине число **, например m .

(4) Теперь можно образовать новую таблицу, в которой наибольшее по величине отрицательное число $-N$ заменено наименьшим по величине положительным числом m .

(5) Выделить это число m и образовать новый план, в который помимо этого числа входят все другие выделенные числа предыдущего плана (за исключением числа m); больше никаких чисел не добавляется ***. Так, в таблице 11.4 наибольшим по величине отрицательным числом **** является (-18) , которое стоит в клетках S_1D_4 и S_2D_5 (т. е. $-N = -18$). В этом случае можно выбрать любую клетку, содержащую это число.

* На этом этапе решения нужно тщательно следить за тем, чтобы не путать выделенные числа X_{ij} , полученные в исходном плане, с «оценками», которые проставляются в той же таблице.

** Если таких чисел несколько, то можно произвольно выбрать любое из них.

*** Улучшение функции цели от одного плана к другому составляет в этом случае mN .

**** Практически не обязательно выбирать наибольшее по величине отрицательное число. Иногда даже выгоднее взять первое полученное отрицательное число. Поскольку изменение функции цели от одного плана к другому составляет mN , то из таблицы 11.4 видно, что выбор S_2D_1 , S_2D_2 , S_2D_5 или S_1D_5 дает изменение на 33, 39, 18 и 11 соответственно, в то время как выбор S_1D_4 дает изменение только на 18. Другим возможным способом является исследование всех произведений mN и выбор отрицательного числа, соответствующего наибольшему произведению, в нашем примере S_2D_2 .

Выбираем, например, S_1D_4 . Проследив контур, образованный для получения числа «—18» в клетке S_1D_4 , получаем сумму (в символах) $+S_1D_3 - S_2D_3 + S_2D_4$. В обеих клетках со знаком «плюс», т. е. S_1D_3 и S_2D_4 , стоят числа **1**. Следовательно, можно выбрать любое из чисел для удаления из плана. Например, выбираем S_2D_4 . При этом выделенное число **1** записывается в клетку S_1D_4 (см. табл. 11.5а), т. е. в клетку, в которой было число —18 в таблице 11.4 *. Другие выделенные цифры (обозначены жирными точками) расположены в тех же клетках, что и в предыдущем плане (см. табл. 11.5а).

Таблица 11.5а

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	.	.	.	1		9
S_2			.			4
S_3				.	.	8
Итого	3	5	4	6	3	21

Таблица 11.5б

Второй опорный план: $C = 233$

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	3	5	0	1	7	9
S_2	—11	—13	4	18	0	4
S_3	—10	—1	1	5	3	8
Итого	3	5	4	6	3	21

Новый план получается путем заполнения клеток с точками новыми числами в соответствии с поставленными ограничениями (вход — выход). В таблице 11.5б эти числа выделены жирным шрифтом. Полученный план

* Следовательно, получим улучшение плана по сравнению с таблицей 11.4 на $1 \times 18 = 18$ единиц стоимости. Таким образом, следующему плану (табл. 11.5б) будет соответствовать стоимость $251 - 18 = 233$ единицы.

оценивается, как и предыдущий, и оказывается, что вновь среди невыделенных оценок есть отрицательные числа. Процесс улучшения плана повторяется (табл. 11.6, 11.7 и 11.8) до тех пор, пока в таблице 11.8 не получается план, в котором все невыделенные числа неотрицательны. Следовательно, получен оптимальный план, или решение.

Таблица 11.6

Третий опорный план: $C = 181$

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	3	1	4	1	7	9
S_2	2	4	13	31	13	4
S_3	-10	-1	1	5	3	8
Итого	3	5	4	6	3	21

Таблица 11.7

Четвертый опорный план: $C = 151$

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	10	1	4	4	7	9
S_2	12	4	13	31	13	4
S_3	3	-1	1	2	3	8
Итого	3	5	4	6	3	21

Таким образом, в таблице 11.8 приводится оптимальный план, который обеспечивает минимальные транспортные издержки по перегону порожних товарных вагонов. Минимальные суммарные затраты равны 150 долларам по сравнению с 251 долларом в исходном плане (возможном, но не оптимальном решении задачи).

Различные оптимальные планы. В заключение следует отметить, что во многих задачах (в противоположность только что рассмотренной) могут существовать несколько оптимальных планов. Если какая-либо из оценок

Таблица 11.8

Оптимальный план (решение): $C = 150$

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	10	1	4	5	7	9
S_2	11	4	12	30	12	4
S_3	3	1	1	1	3	8
Итого	3	5	4	6	3	31

в оптимальном плане равна нулю, то существует несколько оптимальных планов. Эти различные оптимальные планы можно получить с помощью только что приведенного процесса последовательного приближения. Различие состоит только в том, что нули (если они есть), которые фигурируют в оптимальном плане, рассматриваются так же, как отрицательные числа *.

Таблица 11.9

Матрица удельных затрат

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	-2	-1	-4	-3	0	5
S_2	+1	-3	-5	-2	-1	7
S_3	-1	-4	-3	-2	-1	6
Итого	2	2	5	4	5	18

Если существуют различные оптимальные планы, например $\{P_1\}$, $\{P_2\}$, ..., $\{P_n\}$, где $\{P_n\}$ означает набор X_{ij} , образующих n -й оптимальный план, то **

$$\{P_{n+1}\} = a_1 \{P_1\} + a_2 \{P_2\} + \dots + a_n \{P_n\}$$

* Наличие различных оптимальных планов дает возможность выбора «наилучшего» варианта, позволяет учесть также другие критерии.

** Это равенство означает, что, во-первых, каждый элемент матрицы умножается на соответствующий коэффициент и, во-вторых, путем сложения соответствующих элементов образуется новая матрица, которая содержит то же количество столбцов и строк, что и исходные матрицы.

является также оптимальным планом* при условии, что a_i — неотрицательные постоянные, сумма которых равна 1, т. е.

$$\sum_{i=1}^n a_i = a_1 + a_2 + a_3 + \dots + a_n = 1.$$

Например, в задаче минимизации стоимости, представленной в таблице 11.9, есть два оптимальных плана, приведенных в таблицах 11.10 и 11.11. Таблицу 11.11 можно получить из таблицы 11.10 (и наоборот) с помощью вышеприведенного процесса, считая 0 в клетке S_3D_5 таблицы 11.10 (или в клетке S_3D_4 таблицы 11.11) «отрицательным числом наибольшей величины».

Таблица 11.10
Оптимальный план для задачи с условиями в таблице 11.9

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	4	2	2	2	3	5
S_2	2	1	2	3	2	7
S_3	2	2	5	1	0	6
Итого	2	2	5	4	5	18

Таблица 11.11
Другой оптимальный план для задачи с условиями в таблице 11.9

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1	4	2	2	2	3	5
S_2	2	1	2	4	1	7
S_3	2	2	5	0	1	6
Итого	2	2	5	4	5	18

* Исходные оптимальные планы называются *основными оптимальными планами*, а планы, вновь полученные с помощью коэффициентов a_i , называются *производными оптимальными планами*.

Таблица 11.11а

1/4 таблицы 11.10=

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1		1/2			3/4	5
S_2	1/2			3/4	1/2	7
S_3			5/4	1/4		6
Итого	2	2	5	4	5	18

Таблица 11.11б

3/4 таблицы 11.11а=

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1		3/2			9/4	5
S_2	3/2			3	3/4	7
S_3			15/4		3/4	6
Итого	2	2	5	4	5	18

Образуя «выпуклые линейные комбинации» этих двух основных оптимальных планов, можно получить бесконечное число производных оптимальных планов. Так, если мы возьмем две положительные дроби, сумма которых равна единице, т. е. $1/4$ и $3/4$, то можно получить новый оптимальный план, умножая каждый элемент первого плана на $1/4$ и каждый элемент второго плана на $3/4$ и складывая соответствующие клетки. При этом получим производный оптимальный план (табл. 11.12) с помощью таблиц 11.11а и 11.11б. Аналогично можно получить другие оптимальные планы, используя другие неотрицательные дроби, сумма которых равна 1*.

* Следует отметить, что в общем случае в производные оптимальные планы будут входить дробные числа. Очевидно, что такие планы используются только в том случае, когда допустимы нецелочисленные результаты.

Таблица 11.12

1/4 таблицы 11.10 + 3/4 таблицы 11.11а =

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Итого
S_1		2			3	5
S_2	2			3^{3/4}	5/4	7
S_3			5	1/4	3/4	6
Итого	2	2	5	4	5	18

	D_1					Итого
S_3	3					8
Итого	3					

Замечания относительно метода решения транспортной задачи. Метод решения транспортной задачи был рассмотрен на примере задачи (линейной) минимизации, однако очевидно, что этот метод также применим к задаче (линейной) максимизации. Единственная разница при решении задач максимизации заключается в образовании матрицы «прибылей». Если в задаче минимизации все стоимости входят с отрицательным знаком, то в задаче максимизации все прибыли (или другие величины, которые входят в задачу) входят без знаков. Когда образована матрица исходных данных, то можно решать задачу с помощью приведенного выше метода.

Следует отметить, что существует много разновидностей транспортного метода. Одна из таких разновидностей приведена в приложении 11А. На другие варианты уже были ссылки в тексте. В приведенном методе можно уменьшить число итераций при соответствующем преобразовании матрицы стоимостей; проиллюстрируем этот способ на примере задачи, приведенной в таблицах 11.1 и 11.2.

1. Построим новую матрицу, в которой первая строка и первый столбец соответствуют клетке с минимальной стоимостью. В нашем примере это

S_3D_1 . Поместим количество имеющегося продукта в пункте S_3 (8) и количество потребного продукта в пункте D_1 (3) в соответствующие клетки новой матрицы. Положим значение переменной в S_3D_1 $X_{31} = 3$, т. е. минимальному из двух чисел (8 и 3).

	D_1					Итого
S_3	3					8
Итого	3					

2. При этом удовлетворяются требования для D_1 , но в S_3 остается еще 5 единиц продукта. Поэтому выберем следующую минимальную удельную стоимость среди значений стоимости для S_3 . Получаем -4 в клетке S_3D_5 . Следовательно, помещаем D_5 во второй столбец и записываем количество продукта, потребного в D_5 (3). Сравнивая количество потребного в D_5 продукта (3 единицы) с оставшимся в S_3 продуктом (5 единиц), помещаем в клетке S_3D_5 число 3.

	D_1	D_5				Итого
S_3	3	3				8
Итого	3	3				

3. Поскольку в S_3 все еще осталось 2 единицы продукта, выбираем следующее значение минимальной стоимости среди оставшихся значений S_3 , а именно -7 в S_3D_3 . Записываем D_3 в третий столбец и соответственно количество потребного в D_3 продукта, т. е. 4 единицы. Сравниваем эти 4 единицы с оставшимся в S_3 продуктом ($8-3-3=2$) и помещаем в S_3D_3 число 2. При этом весь продукт в S_3 будет израсходован, а в D_3 требуется еще 2 единицы продукта.

4. Сравниваем стоимости, связанные с D_3 ($C_{13} = -5$ и $C_{23} = -7$), выбираем S_1 и помещаем его во вторую строку, одновременно записываем количество имеющегося в S_1 продукта, т. е. 9. Сравнивая количество

имеющегося продукта в S_1 (т. е. 9) с недостающим в D_3 продуктом (т. е. $2 = 4 - 2$), помещаем 2 единицы в клетку S_1D_3 , так что теперь требования D_3 в продукте удовлетворены.

	D_1	D_6	D_3			Итого
S_3	3	3	2			8
Итого	3	3	4			

	D_1	D_6	D_3			Итого
S_3	3	3	2			8
S_1						9
Итого	3	3	4			

5. Продолжая этот процесс, получим следующую матрицу:

	D_1	D_6	D_3	D_4	D_2	Итого
S_3	3	3	2			8
S_1			2	6	1	9
S_2					4	4
Итого	3	3	4	6	5	21

Этому исходному плану соответствует суммарная стоимость

$$(3)(-1) + 3(-4) + 2(-7) + 2(-5) + 6(-9) + 1(-20) + 4(-10),$$

т. е. без учета знака получаем

$$\text{Суммарная стоимость} = \$153$$

по сравнению с суммарной стоимостью исходного плана, полученного по правилу северо-западного угла, 251 доллар (оптимальный план 150 долларов). Такая перестановка матрицы стоимостей * обычно приводит к лучшему (т. е. с меньшими расходами или с большей прибылью) исходному плану, так что оптимальный план можно получить за меньшее число итераций.

Наконец **, можно отметить, что транспортный метод (1) можно использовать, чтобы оценить, какой проигрыш дает отклонение плана от оптимального, и, кроме того, (2) в нем нет сложных математических выкладок, используется только сложение, вычитание и умножение. Следовательно, при решении задач большего размера можно использовать различные настольные счетные машины, так что можно ожидать более широкого применения транспортного метода к решению различных задач.

ОБЩАЯ ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

В начале этой главы рассматривался определенный класс задач оптимизации и их решение с помощью транспортного метода. Теперь перейдем к рассмотрению общей задачи линейного программирования и общего метода получения оптимального плана, или решения *** этой задачи. Этот метод был сформулирован Данцигом [10] и получил название симплексного метода ****.

Симплексный метод представляет собой метод решения задач оптимизации линейной формы при наличии ограничений в форме линейных неравенств. Как и транспортный метод, симплексный метод рассматривается на конкретном примере. Однако, прежде чем перейти к этому примеру, введем некоторые понятия.

Математические основы. Прежде всего полезно показать разницу между уравнением и неравенством. Уравнение, например $y = 2x$, можно представить геометрически в виде прямой линии с «наклоном» 2 (см. рис. 11.1). Это простая и краткая форма записи следующего положения: «имеем соотношение между двумя переменными x и y такое, что при любой заданной величине x величина y будет в 2 раза больше». Кроме того, любая точка с абсциссой x и ординатой y , удовлетворяющими линейному соотношению $y = 2x$, будет лежать на данной прямой. Например, точка с декартовыми координатами $x = 1$, $y = 2$ [записывается как $(1, 2)$] и точка $(3, 6)$ лежат на данной прямой и, следовательно, являются решениями данного

* Легко заметить, что этот исходный план, соответствующий 153 долларам, можно получить без преобразования матрицы. Надо только начать процесс с клетки с минимальной стоимостью (в нашем примере S_3D_1).

** Более полное изложение транспортного метода, включая рассмотрение вырожденных случаев, можно найти в [18]. Математическое обоснование транспортного метода дано в [29, гл. 23].

*** Непосредственным результатом является оптимальное распределение, а не оптимальный план.

**** Здесь не приводится математическое обоснование симплексного метода. Оно дано в [29, гл. 21].

уравнения *. И обратно, всякая точка, например (1, 3) и (3, 8), которая не лежит на данной прямой линии, не удовлетворяет соответствующему уравнению или функциональной зависимости.

Неравенство записывается с помощью одного из четырех различных знаков в зависимости от содержания. Это следующие знаки: $<$, $>$, $\leq$ или $\geq$ или $\leq$ и $\geq$. Знак $<$ означает, что величина переменной в левой части меньше, чем величина переменной в правой части. Например, $y < 2x$ означает, что для любого заданного значения x значение y будет меньше удвоенной величины x . Это неравенство удовлетворяется в любой точке, лежащей

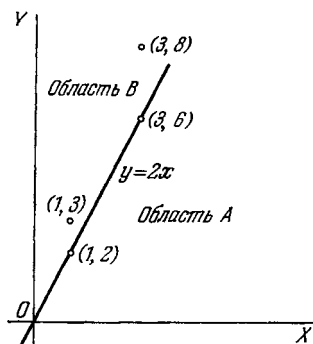


Рис. 11.1. $y = 2x$.

в области A рис. 11.1, т. е. координаты любой точки (x, y) в области A (не включая граничную линию $y = 2x$) таковы, что ордината y меньше удвоенной абсциссы x . Аналогично $y > 2x$ означает, что для любого заданного значения x значение y больше удвоенного x . Геометрически этому неравенству удовлетворяют все точки, лежащие в области B (не включая границы $y = 2x$).

Более общими являются знаки $\leq$ и $\geq$. Знак $\leq$ или эквивалентный знак $\leq$ означает, что значение переменной в левой части меньше или равно значению переменной в правой части. Так, $y \leq 2x$ означает, что для определенного значения x y меньше или равен удвоенному значению x . Другими словами, можно сказать, что значение y не больше удвоенного значения x .

Геометрически это неравенство $y \leq 2x$ включает все точки, принадлежащие области A и границе $y = 2x$. И наконец, $y \geq 2x$ или $y \geq 2x$ означает, что значение y больше или равно удвоенному значению x , и геометрически этому неравенству удовлетворяют все точки области B и на линии $y = 2x$.

Следовательно, неравенства $y \leq 2x$ и $y \geq 2x$ включают в себя также уравнение $y = 2x$, так как всякое решение уравнения автоматически является решением любого (или обоих) неравенства. Однако обратное утверждение неверно, так как решение неравенства $y \geq 2x$ не обязательно будет решением уравнения $y = 2x$. Например, точки (1, 6) и (3, 8) (и бесконечно много других) являются решениями неравенства $y \geq 2x$, но не удовлетворяют уравнению $y = 2x$ (см. рис. 11.1).

В качестве другого примера рассмотрим систему трех неравенств

$$x - 5 \geq 0, \quad y - 2 \geq 0, \quad -6x - 8y + 120 \geq 0.$$

Геометрически решениями этой системы неравенств является (замкнутое) множество точек, лежащих в треугольнике ABC и на его границе (см. рис. 11.2). Существует бесконечное число решений этой системы неравенств. В то же время соответствующая система уравнений

$$x - 5 = 0, \quad y - 2 = 0, \quad -6x - 8y + 120 = 0$$

не имеет решения. (Чтобы существовало решение, все три линии должны пересекаться в одной точке, как на рис. 11.3.)

Таким образом, из предыдущих примеров видно, что уравнение (или система уравнений) является более жестким условием, чем соответствующее

* Когда говорят о точке (x, y) , имеют в виду точку с координатами x и y , где в декартовой системе координат, например, первое число есть абсцисса и второе — ордината.

неравенство (или система неравенств) *. И обратно, неравенство является более свободным условием в том смысле, что решения уравнения, если они есть, являются только подмножеством среди всех возможных решений неравенства. Это очень существенно для многих задач, возникающих в промышленности и торговле. В этих задачах обычно есть ограничения следующего вида: «допуск на отклонение от заданного размера должен быть не больше чем 0,002 дюйма», «произвести по крайней мере не меньше чем...», «доставить эти продукты к определенному сроку», «стоимость не больше

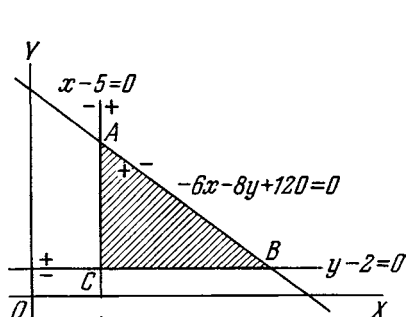


Рис. 11.2. Область, ограниченная тремя прямыми.

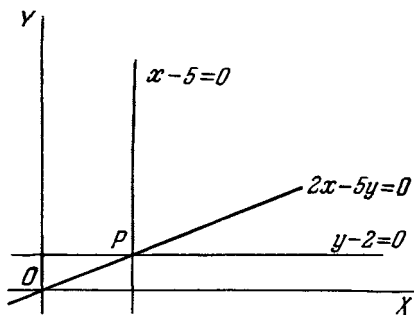


Рис. 11.3. Решение трех линейных уравнений как точка пересечения трех прямых.

чем...», и другие. Эти ограничения носят характер неравенств, т. е. задаются верхний и нижний пределы, а не одна точная граница. Любая неверная словесная или математическая формулировка задачи, которая превращает систему неравенств в систему равенств, ведет к сокращению области возможных решений. Иногда даже возможен ошибочный ответ, что нет никакого решения, как во втором примере, хотя в действительности существует одно или несколько решений. Кроме того, даже если можно определить так называемое оптимальное решение для системы уравнений, может оказаться, что существует много решений действительной системы (неравенств), которые являются более оптимальными для данной функции цели. С другой стороны, в противоположность уравнениям системы неравенств допускают оптимальные решения, которые могут находиться как внутри области, так и на ее границе. Следовательно, чтобы избежать ошибок и неоптимальных результатов, важно правильно определить систему ограничений.

Для обозначения суммирования используется знак Σ . Если есть выражение вида

$$X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7,$$

то его можно прежде всего сократить, используя обозначение в виде многочлена, т. е.

$$X_1 + X_2 + \dots + X_7.$$

Однако более коротким будет обозначение $\sum_{j=1}^7 X_j$, где

$$\sum_{j=1}^7 X_j = X_1 + X_2 + \dots + X_7.$$

* В дальнейшем будет идти речь только о неравенствах вида $\leq$ и $\geq$, но не о неравенствах вида $<$ и $>$.

Этот знак, который является заглавной греческой буквой «сигма», особенно удобен для обозначения бесконечных рядов. Например, бесконечный ряд

$$1 + 1/2 + 1/4 + 1/8 + 1/16 + 1/32 + \dots$$

(который, кстати, имеет предел, равный 2) можно записать более коротко, т. е.

$$\sum_{n=0}^{\infty} 1/2^n.$$

Наконец, будут использованы двойные индексы, например X_{ij} . Использование двойных индексов естественно вытекает из необходимости возможно проще расположить числа (или элементы) в виде прямоугольной таблицы (или матрицы). Например, если рассматривается матрица

Столбец Строка	C_1	C_2	C_3	C_4
R_1	1	2	4	7
R_2	—3	6	1	0
R_3	—2	—1	0	3

то X_{12} означает элемент, расположенный в первой строке и во втором столбце, а именно 2. В общем случае прямоугольная матрица сама будет обозначаться через $\{X_{ij}\}$, а X_{ij} означает отдельный элемент в i -й строке и j -м столбце матрицы.

Теперь вернемся к общей задаче решения совместной системы уравнений или неравенств и выбора среди различных решений наилучшего для данной функции цели. Рассмотрим второй из двух методов — симплексный метод решения общей задачи линейного программирования.

СИМПЛЕКСНЫЙ МЕТОД

В качестве иллюстрации симплексного метода рассмотрим относительно простой пример. Предприниматель хочет получить максимальную прибыль при производстве двух видов продукции R и S . Производство изделий R и S распадается на два этапа, причем все предварительные операции выполняются на участке 1 и все последующие операции выполняются на участке 2А или на участке * 2В. Участки 2А и 2В для любого заданного изделия отличаются по производительности и по прибыли на единицу продукции. Кроме того, допустим, что на участке 2А есть определенное количество сверхурочного времени для производства изделий R и S . Но поскольку использование сверхурочного времени влечет за собой изменение (уменьшение) прибыли на единицу продукции (но не влияет на производительность), примем, что сверхурочное время участка 2А эквивалентно дополнительному участку 2АА.

* Под *участком* понимается группа машин, не обязательно однотипных, но таких, которые можно логически сгруппировать, что необходимо для решения задачи.

В таблице 11.13 приведены продолжительности выполнения различных операций по производству одного изделия типов R и S , общее машинное время, которым располагает каждый участок и прибыль на единицу продукции. Кроме того, для упрощения изложения вводятся обозначения трех возможных комбинаций участков при производстве изделия R , а именно R_1 , R_2 и R_3 , и аналогично для изделия S : S_1 , S_2 и S_3 .

Таблица 11.13

Операция	Участок	Продукт R			Продукт S			Наличное время
		R_1	R_2	R_3	S_1	S_2	S_3	
1	1	0,01	0,01	0,01	0,03	0,03	0,03	850
2 {	2A	0,02			0,05			700
	2AA		0,02			0,05		100
	2B			0,03			0,08	900
Доход		\$0,40	\$0,28	\$0,32	\$0,72	\$0,64	\$0,60	

Задача заключается в том, чтобы определить, какое количество изделий каждого вида следует выпускать, используя различные комбинации участков, с тем чтобы получить максимальную прибыль с учетом ограниченных возможностей участков*.

Запишем теперь задачу в математической форме. Если X_1 , X_2 , X_3 , X_4 , X_5 , X_6 означают соответственно количества изделий R_1 , R_2 , R_3 , S_1 , S_2 , S_3 , то общую прибыль Z можно записать в виде (см. табл. 11.13)

$$Z = 0,4X_1 + 0,28X_2 + 0,32X_3 + 0,72X_4 + 0,64X_5 + 0,60X_6. \quad (1)$$

Кроме того, ограничения записываются в следующем виде:

$$\left. \begin{aligned} 0,01X_1 + 0,01X_2 + 0,01X_3 + 0,03X_4 + 0,03X_5 + 0,03X_6 &\leq 850, \\ 0,02X_1 + 0,05X_4 &\leq 700, \\ 0,02X_2 + 0,05X_5 &\leq 100, \\ 0,03X_3 + 0,08X_6 &\leq 900. \end{aligned} \right\} \quad (2)$$

(Эти ограничения означают, что время, потраченное на производство изделий R и S на каждом участке, не должно превышать машинного времени этого участка.)

Следовательно, задача формулируется следующим образом**. Определить величины $X_j \geq 0$ (где $j = 1, 2, \dots, 6$), которые максимизируют

$$Z = 0,4X_1 + 0,28X_2 + 0,32X_3 + 0,72X_4 + 0,64X_5 + 0,60X_6 \quad (1)$$

* Здесь сделано упрощающее допущение, что можно обеспечить сбыт всей произведенной продукции. Если это не соответствует действительности, то можно наложить дополнительные ограничения в виде максимально допустимого количества продукции каждого вида.

** Ограничения $X_j \geq 0$ ($j = 1, 2, \dots, 6$) означают, что процесс производства носит необратимый характер, поэтому отрицательные значения переменных не имеют смысла.

при наличии ограничений

$$\left. \begin{aligned} 0,01X_1 + 0,01X_2 + 0,01X_3 + 0,03X_4 + 0,03X_5 + 0,03X_6 &\leq 850, \\ 0,02X_1 + 0,05X_4 &\leq 700, \\ 0,02X_2 + 0,05X_5 &\leq 100, \\ 0,03X_3 + 0,08X_6 &\leq 900. \end{aligned} \right\} \quad (2)$$

Чтобы перейти к симплексному методу, следует превратить систему *неравенств* (2) в эквивалентную систему *уравнений* с помощью введения новых неотрицательных переменных X_7, X_8, X_9, X_{10} . Тогда получим:

$$\left. \begin{aligned} 0,01X_1 + 0,01X_2 + 0,01X_3 + 0,03X_4 + 0,03X_5 + 0,03X_6 + X_7 &= 850, \\ 0,02X_1 + 0,05X_4 + X_8 &= 700, \\ 0,02X_2 + 0,05X_5 + X_9 &= 100, \\ 0,03X_3 + 0,08X_6 + X_{10} &= 900. \end{aligned} \right\} \quad (3)$$

Эти новые переменные X_7, X_8, X_9 и X_{10} носят различные названия: «свободные или искусственные» переменные, «псевдопеременные» или «дополнительные переменные». Легко видеть, что в нашей задаче *положительные* значения этих искусственных переменных означают недоиспользованные возможности участков 1, 2А, 2АА и 2В соответственно.

Чтобы записать исходную систему уравнений (1) и (3) в каноническом виде, используем в симплексном методе, а также для получения большей компактности сделаем следующие преобразования. Предположим, что нужно так записать уравнения (3), чтобы соответствующие X_j были записаны в одном и том же столбце. Тогда, считая все пустые места нулями, получим, например, для X_1 столбец коэффициентов: 0,01; 0,02; 0; 0 сверху вниз. Обозначим через P_j столбец коэффициентов при X_j ($j = 1, 2, \dots, 10$) и через P_0 столбец чисел, стоящих в правых частях системы уравнений (3) *.

Кроме того, умножение P_j (или P_0) на действительное число означает, что каждый элемент этого столбца умножается на это число. Так, возвращаясь к коэффициентам при X_1 , получим

$$X_1 P_1 = X_1 \begin{vmatrix} 0,01 \\ 0,02 \\ 0 \\ 0 \end{vmatrix} = \begin{vmatrix} 0,01X_1 \\ 0,02X_1 \\ 0 \\ 0 \end{vmatrix}. \quad (4)$$

Наконец, если P_1 и P_2 являются двумя такими «векторами», то

$$X_1 P_1 + X_2 P_2 = X_2 P_2 + X_1 P_1. \quad (5)$$

Теперь можно сформулировать задачу линейного программирования следующим образом: определить неотрицательные значения переменных X (где $j = 1, 2, \dots, 10$), которым соответствует максимальное значение линейной формы (функционала) **

$$Z = 0,40X_1 + 0,28X_2 + 0,32X_3 + 0,72X_4 + 0,64X_5 + 0,60X_6 + \\ + 0 \cdot X_7 + 0 \cdot X_8 + 0 \cdot X_9 + 0 \cdot X_{10}. \quad (1a)$$

* P_j и P_0 являются *векторами* в четырехмерном пространстве (для нашей задачи); размерность пространства определяется числом ограничений задачи.

** Здесь приняты нулевая прибыль или нулевой убыток, соответствующие каждой искусственной переменной X_7, X_8, X_9 и X_{10} .

при наличии ограничений

$$\sum_{j=1}^{10} X_j P_j = P_0. \quad (3a)$$

Записав задачу в таком виде, мы можем перейти к ее решению с помощью симплексного метода *. Первый шаг состоит в том, чтобы правильно записать векторы-столбцы P_j с помощью уравнений (3), пустые места в которых соответствуют нулевым значениям коэффициентов (см. табл. 11.14).

Таблица 11.14

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_0
0,01	0,01	0,01	0,03	0,03	0,03	1				850
0,02			0,05				1			700
	0,02			0,05				1		100
		0,03			0,08				1	900

Следует отметить, что уравнения (3) можно получить из таблицы 11.14, умножая коэффициенты в столбцах P_j на соответствующие X_j и записывая результаты по строкам. (Вертикальная жирная линия показывает место знака равенства.) Особенно важной является квадратная подматрица, образованная векторами P_7, P_8, P_9, P_{10} , элементами которой являются единицы на главной диагонали и все остальные нули. Эта матрица называется *единичной* или *тождественной* матрицей. Система векторов, образующих *единичную* матрицу, называется *единичным базисом* определенного пространства, в данном случае четырехмерного пространства **.

Чтобы приступить к решению задачи с помощью симплексного метода, следует переставить столбцы, как указано в таблице 11.15а. Кроме того, слева от столбца P_0 вводится столбец «базис», в котором проставляются векторы базиса ***. Вводится также строка C_j , в которой проставляются значения коэффициентов C_j при соответствующих X_i в выражении (1) для Z . Добавляется столбец C_i , аналогичный строке C_j , с тем отличием, что индексы

* В действительности, вследствие существования теоремы *двойственности* можно решать любую из двух задач вместо одной. Это объясняется тем, что с каждой задачей линейного программирования связана двойственная задача, так что если в одной задаче требуется максимизировать линейную функцию, например Z , а в другой минимизировать линейную функцию, например g , то $Z_{\max} = g_{\min}$. Подробнее о проблеме двойственности в линейном программировании см. приложение 11В, а также [4] и [29].

** Базисом являются линейно независимые векторы, с помощью которых можно единственным образом определить любую точку n -мерного пространства (в данном примере $n = 4$); решение (или решения) задачи также выражается через базисные векторы.

*** В данном примере единичный базис образуют дополнительные векторы. В некоторых задачах, в которых в форме равенств или неравенств накладываются слабые ограничения, вводятся так называемые искусственные векторы, которые образуют единичный базис [4, стр. 15]. Следует отметить также, что основные векторы могут быть такими, что их можно ввести в единичный базис.

Таблица 11.15а

$C_j \backslash C_i$							0,40	0,28	0,32	0,72	0,64	0,60
	Базис	P_0	P_7	P_8	P_9	P_{10}	P_1	P_2	P_3	P_4	P_5	P_6
	P_7	850	1				0,01	0,01	0,01	0,03	0,03	0,03
←	P_8	700		1			0,02			0,05		
	P_9	100			1			0,02			0,05	
	P_{10}	900				1			0,03			0,08
	Z_j											
	$Z_j - C_j$						-0,40	-0,28	-0,32	-0,72 *	-0,64	-0,60

i означают строку, а индексы j означают столбец. Выражение для Z можно переписать в виде

$$Z = \sum_{j=1}^{10} C_j X_j. \quad (6)$$

Вводится строка Z_j , где j означает соответствующий столбец. Z_j определяется следующим образом. Считая, что X_{ij} означает элемент в i -й строке и j -м столбце таблицы, определим Z_j (включая Z_0) следующим образом:

$$Z_j = \sum_i C_i X_{ij}. \quad (7)$$

И наконец, вводится еще одна строка $Z_j - C_j$, которая для любого столбца, например для j_0 , определяет собой разность между Z_{j_0} , определенным в предыдущей строке, и соответствующим коэффициентом C_{j_0} .

На этом заканчивается составление таблицы (табл. 11.15а) и первый этап вычислений. Согласно введенной ранее терминологии имеем исходный план (возможное решение) задачи, который выражается через базисные векторы P_7, P_8, P_9 и P_{10} и через значения вектора P_0 , т. е.

$$X_7 = 850, \quad X_8 = 700, \quad X_9 = 100, \quad X_{10} = 900. \quad (8)$$

Таким образом, исходное возможное решение можно сформулировать следующим образом: «Не использовать рабочего времени ни на каком участке», т. е. «ничего не делать»; при этом будет получена чистая прибыль $Z = 0$.

Критерий оптимальности плана. Когда получен план, то возникает вопрос, нельзя ли получить более выгодный план. Возможен один и только один из следующих вариантов.

M1. $\max Z = \infty$ (т. е. максимальное значение Z бесконечно велико) и получено из данного плана.

M2. Максимальное значение Z конечно и получено из данного плана.

M3. Оптимальный план еще не получен, и можно получить план с еще большим значением Z . Симплексный метод дает возможность за конечное число шагов установить, имеет ли место M1 или M2. Кроме того (при условии, что X_{ij} означает элемент i -й строки и j -го столбца), симплексный метод дает возможность на основании данной таблицы определить:

1. Если существует какой-нибудь отрицательный элемент $Z_j - C_j < 0$, то имеет место или M1, или M3.

(а) Если все $X_{ij} \leq 0$ в этом столбце (для которого $Z_j - C_j < 0$), то справедливо M1.

(б) Если некоторые $X_{ij} > 0$, то необходимы дальнейшие вычисления, т. е. справедливо M3.

2. Если все $Z_j - C_j \geq 0$, то получено максимальное значение (M2).

Итеративный процесс получения оптимального плана. В нашем примере (табл. 11.15а) $Z_1 - C_1 < 0$ (отрицательны также разности от $Z_2 - C_2$ до $Z_6 - C_6$), кроме того, некоторые коэффициенты в столбце P_1 положительны. Следовательно, согласно условию 1(б) необходимы дальнейшие вычисления (т. е. выполняется условие M3).

Чтобы получить новые планы, следует выполнять расчеты по симплексному алгоритму. При этом любому новому плану будет соответствовать по крайней мере не меньшее значение функции цели (хотя не обязательно функция цели будет увеличиваться)*.

Процесс вычислений заключается в следующем. Из всех $Z_j - C_j < 0$ выбираем наибольшее отрицательное число (в приведенном примере выбираем $Z_4 - C_4 = -0,72$, которое отмечено звездочкой в таблице 11.15а). Тем самым определяется определенный вектор P_j (а именно P_4), который следует ввести в столбец, обозначенный в таблице 11.15а через «базис». Чтобы определить вектор, который следует исключить из базиса вместо введенного P_j , надо сначала разделить на каждое из положительных значений X_{ij} , соответствующих столбцу P_j , элементы X_{i0} , которые стоят в тех же строках, но в столбце P_0 **. Минимальное число, полученное от деления, определяет вектор, который следует исключить из базиса. Таким образом, в приведенном примере P_4 вводится в базис вместо одного из векторов P_7 , P_8 , P_9 или P_{10} . Вектор P_4 имеет две положительные компоненты X_{ij} , а именно $X_{7,4} = 0,03$ и $X_{8,4} = 0,05$. При делении элементов X_{i0} на соответствующие положительные X_{ij} получим минимальную величину 14 000 (т. е. $700/0,05$). Следовательно, следует исключить вектор P_8 . При этом получим новый базис, состоящий из векторов P_7 , P_4 , P_9 , P_{10} .

Теперь положим, что индекс k означает вектор, который вводится в базис, индекс r означает вектор, который исключается из базиса, X'_{ij} означает элемент новой матрицы и

$$\varphi = \min_i \frac{X_{i0}}{X_{ik}}, \quad X_{ik} > 0 \quad (9)$$

(т. е. φ — минимальное число из всех отношений X_{i0}/X_{ik} для $X_{ik} > 0$). Тогда элементы новой матрицы (X'_{ij}) вычисляются следующим образом.

* Как указывалось выше, симплексный метод гарантирует получение оптимального плана, если он существует, за конечное число шагов. Следовательно, симплексный метод является сходящимся итеративным процессом.

** Поскольку все составляющие P_0 должны быть положительными, эти отношения также будут положительными числами.

Для элементов строки, соответствующей вектору, введенному в базис, получаем

$$X'_{kj} = X_{rj} / X_{rk}. \quad (10)$$

Остальные элементы X'_{ij} новой матрицы вычисляются по формуле

$$X'_{ij} = X_{ij} - \frac{X_{rj}}{X_{rk}} X_{ik}. \quad (11)$$

По формуле (11) определяются также элементы X_{i0} столбца P_0 и элементы последней строки $Z_i - C_j$ (но не предпоследней строки Z_j). Новое значение функции цели вычисляется по формуле *

$$Z'_0 = Z_0 - \varphi (Z_k - C_k), \quad (12)$$

или, поскольку $C_0 = 0$, функцию цели можно задать в виде

$$(Z_0 - C_0)' = (Z_0 - C_0) - \varphi (Z_k - C_k). \quad (13)$$

Например, если имеем исходную таблицу 11.15а, то при переходе к таблице 11.15б наибольшим отрицательным числом $Z_j - C_j$ является

Таблица 11.15б

	P_7	430	1	-0,6		-0,002	0,01	0,01		0,03	0,03
→	0,72	P_4	14 000	20		0,4			1		
←		P_9	100		1		0,02			0,05	
		P_{10}	900			1		0,03			0,08
		Z_j	10 080	14,4		0,288			0,72		
		$Z_j - C_j$	10 080	14,4		-0,112	-0,23	-0,32		-0,64 *	-0,60

$Z_4 - C_4 = -0,72$. Следовательно, $k = 4$. Из уравнения (9) получаем

$$\varphi = \min_i \frac{X_{i0}}{X_{ik}} \quad \text{для всех } X_{ik} > 0,$$

т. е.

$$\varphi = \min \left(\frac{850}{0,03} = 28\,333; \frac{700}{0,05} = 14\,000 \right) = 14\,000.$$

Следовательно P_4 вводится в базис вместо P_8 ; таким образом, в наших уравнениях $k = 4$, $r = 8$. Элементы строки P_4 таблицы 11.15а вычисляются по формуле (см. выражение (10))

$$X'_{4j} = \frac{X_{8j}}{X_{84}} = \left(\frac{X_{8j}}{0,05} \right);$$

* Другими словами, улучшение плана от одной итерации к другой соответствует изменению функции цели на $-\varphi (Z_k / C_k)$.

Следует обратить внимание на сходство в изменении функции цели в симплексном и в транспортном методах.

Таблица 11.15г

	P_7	32,5	1	-0,6	-0,6	$-3/8$	-0,002	-0,002	$-1/800$			
← 0,72	P_4	14 000		20			0,4				1	
	0,64	P_5	2 000		20			0,4				1
→ 0,60		P_6	11 250			$12^{1/2}$			$3/8$			1
	$Z_j - C_j$	18 110		14,4	12,8	$7^{1/2}$	-0,112*	-0,024	-0,095			

Таблица 11.15д

C_j								0,40	0,28	0,32	0,72	0,64	0,60
C_i													
	Базис	P_0	P_7	P_8	P_9	P_{10}	P_{11}	P_2	P_3	P_4	P_5	P_6	
	P_7	102,5	1	$-1/2$	-0,6	$-3/8$		-0,002	$-1/800$	0,005			
→ 0,40	P_1	35 000		50			1			$5/2$			
	0,64	P_5	2 000		20			0,4				1	
← 0,60	P_6	11 250				$100/8$			$3/8$				1
	$Z_j - C_j$	22 030		20	12,8	$7^{1/2}$		-0,024	-0,095*	0,28			

одно из условий M1 или M2. В данном примере решение получено после 6 итераций, т. е. в шестой таблице после первой начальной. В последней таблице 11.15ж дан оптимальный план, или решение *. Этот же оптималь-

* Если бы существовали другие оптимальные планы, то в таблице появилось бы значение $Z_j - C_j = 0$ в столбце, не входящем в базис. В данном примере $Z_j - C_j = 0$ только для $j = 1, 2, 3$ и 7. Следовательно, не существует других оптимальных планов.

Таблица 11.15е

	P_7	140	1	$-1/2$	$-0,6$	$-1/3$		$-0,002$	$0,005$	$1/300$
0,40	P_1	35 000		50			1		$5/2$	
← 0,64	P_5	2 000			20			0,4		1
→ 0,32	P_3	30 000				$100/3$			1	$8/3$
	$Z_j - C_j$	24 880		20	12,8	$32/3$		$-0,024 *$	0,28	$0,25^{1/3}$

Таблица 11.15ж

Оптимальный план (решение задачи)

$C_j \backslash C_i$							0,40	0,28	0,32	0,72	0,64	0,60
	Базис	P_0	P_7	P_8	P_9	P_{10}	P_1	P_2	P_3	P_4	P_5	P_6
	P_7	150	1	$-1/2$	$-1/2$	$-1/3$				$1/200$	$1/200$	$1/300$
0,40	P_1	35 000		50			1			$5/2$		
0,28	P_2	5 000			50			1			$5/2$	
0,32	P_3	30 000				$100/3$						$8/3$
	$Z_j - C_j$	25 000		20	14	$10^{2/3}$				0,28	0,06	$0,25^{1/3}$

ный план представлен в таблице 11.16 (количество единиц каждого изделия) и в таблице 11.17 (время, необходимое для производства изделий) *.

Таким образом, оптимальный план (дающий наибольшую прибыль) заключается в производстве 70 000 единиц изделий R , изделия S полностью

* Геометрическую интерпретацию задач линейного программирования и их решение с помощью симплексного метода см. в приложении 11Б. Сокращенный метод решения см. в приложении 11Г.

Таблица 11.16

Оптимальный план (количество деталей)

Изделия R		Изделия S
R_1 Участки (1—2A)	35 000 штук	S_1 0 штук
R_2 (1—2AA)	5 000 штук	S_2 0 штук
R_3 (1—2B)	30 000 штук	S_3 0 штук
Всего	$R=70\,000$ штук	$S=0$
Общая прибыль	$\$25\,000 + 0 = \$25\,000$	

Таблица 11.17

Оптимальный план (в часах)

Операция	Участок	Изделия R			Изделия S			Использованное время	Наличное время	Неиспользованное время
		R_1	R_2	R_3	S_1	S_2	S_3			
1	1	350	50	300	0	0	0	700	850	150
2	2A	700	100	900	0	0	0	700	700	0
	2AA							100	100	0
	2B							900	900	0

исключены. С помощью выражения (1), а также из таблицы 11.15ж ($Z_0 - C_0$) можно определить общую прибыль:

$$Z = 0,40 \cdot 35\,000 + 0,28 \cdot 5\,000 + 0,32 \cdot 30\,000 + 0,72 \cdot 0 + \\ + 0,64 \cdot 0 + 0,60 \cdot 0 = \$25\,000.$$

ДОПОЛНИТЕЛЬНЫЕ ОГРАНИЧЕНИЯ В ЗАДАЧАХ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Могут возникнуть возражения (или по крайней мере вопросы) относительно решения рассмотренной задачи линейного программирования в связи с тем, что не были учтены минимальные потребности в изделиях S. (Необходимость производства изделий S может быть оговорена в контракте, может определяться требованиями рынка или вытекать из стремления руководства поддерживать репутацию фирмы или в связи с условиями комплексного производства.) Кроме того, могут возникнуть сомнения в ценности оптимального плана (табл. 11.15ж), в частности и симплексного метода в целом, в случае, когда нужно учесть (1) изменения в машинном времени любого из участков, (2) изменения в цене любого из изделий и, следовательно, изменения прибыли от их продажи и (3) изменения производительности, вызванные использованием новых машин, штампов, шаблонов или усовершенствованием производственного процесса. Поскольку эти вопросы важны при решении

любой практической задачи, следует установить, как можно использовать симплексный метод с учетом этих видоизменений задачи.

В общем случае введение этих *дополнительных* ограничений создает фактически новую задачу, причем благодаря использованию симплексного метода нет необходимости решать эту задачу с самого начала. Необходимо знать значения $Z_j - C_j$ оптимального плана исходной задачи, который приведен в таблице 11.15ж. Разности $Z_j - C_j$ представляют собой «стоимости выигрыша» (проигрыша). В нашем примере для $j = 1, 2, \dots, 6$ разности $Z_j - C_j$ представляют собой *минимальные* убытки или потери прибыли, связанные с производством одного изделия, соответствующего X_{ij} . Для $j = 7, 8, 9, 10$ разности $Z_j - C_j$ представляют собой прибыль, получаемую за счет одной дополнительной единицы X_j , а именно одной дополнительной единицы времени на соответствующем участке.

Рассмотрим теперь подробнее четыре различные модификации задачи.

Ограничения на минимальное количество изделий. Предположим теперь, что наложено дополнительное ограничение в виде минимального требования на количества изделий S . Это ограничение может определяться требованиями контракта, потребностями рынка или определяться стремлением предпринимателя иметь определенное количество этих изделий, например, для поддержания репутации фирмы или ввиду необходимости комплексного обеспечения существующего производственного процесса.

Согласно оптимальному плану, приведенному в таблице 11.15ж для исходной задачи, изделия S производить не следует. Если же задано, что необходимо произвести по крайней мере определенное количество изделий S , например S_0 , то ясно, что это отклонение от оптимального плана вызовет уменьшение прибыли. А так как в оптимальном плане (табл. 11.15ж) количество изделий S должно быть меньше S_0 , то ясно, что нужно произвести точно S_0 (и не больше) единиц изделий S . Просматривая разности $Z_j - C_j$ в таблице 11.15ж:

$$Z_4 - C_4 = 0,28, \quad Z_5 - C_5 = 0,06 \quad \text{и} \quad Z_8 - C_8 = 0,253,$$

легко заметить, что наиболее дешевым способом производства изделий S , т. е. способом, который меньше всего сокращает прибыль, является способ, связанный с P_5 , т. е. использующий комбинацию участков 1 — 2АА. Вторым по дешевизне является способ, связанный с P_6 (т. е. 1 — 2В), и третьим — способ, связанный с P_4 (1 — 2А).

Предположим, например, что $S_0 = 1000$, другими словами, предположим, что нужно произвести (по крайней мере) 1000 изделий S . Как только что указывалось, минимальное сокращение прибыли от производства этих 1000 изделий можно получить, используя комбинацию участков 1 — 2АА (P_5).

В таблице 11.18 приведен план, дающий минимальное сокращение прибыли, т. е. оптимальный план для расширенной задачи. Каждый элемент соответствующего столбца оптимального плана (в нашем примере столбец P_5) представляет собой оптимальное изменение, которое следует произвести над базисными векторами (т. е. векторами, которые входят в окончательное решение) для каждой единицы X_j (X_5), которая вводится в оптимальный план.

Важно отметить, что в связи с требованием неотрицательности элементов X_{i0} вектора P_0 ($X_{i0} \geq 0$) оптимальные изменения возможны только до тех пор, пока X_{i0} остаются неотрицательными. Кроме того, «стоимость

выигрыша», заданную в виде $Z_j - C_j$, можно получать только до тех пор, пока изменения значений X_{i0} не дают отрицательных значений X_{i0} .

Таблица 11.18

Оптимальный план таблицы 11.15ж	Новый оптимальный план для $S_0=1000$
$X_7=150$	$X_7=150-1000 \cdot 1/200=145$
$X_1=35\ 000$	$X_1=35\ 000-1000 \cdot 0=35\ 000$
$X_2=5\ 000$	$X_2=5\ 000-1000 \cdot 5/2=2\ 500$
$X_3=30\ 000$	$X_3=30\ 000-1000 \cdot 0=30\ 000$
	$X_5=0+1000=1\ 000$
$Z_0-C_0=\$25\ 000$	$Z_0-C_0=\$25\ 000-\$1\ 000 \times 0,06=\$24\ 940$

Рассмотрим вновь ограничение на минимальное количество изделий S . Легко заметить, что последовательность 1 — 2АА можно использовать для значений S , не превышающих 2000 единиц (так как из $5000 - y(5/2) \geq 0$ следует, что $y \leq 2000$). Предположим тем не менее, что надо произвести 8000 изделий S . Для $S_0 = 8000$ (в действительности для всех S_0 , удовлетворяющих условию $2000 < S_0 < 13\ 250$) соответствующий оптимальный план получается за два шага (см. табл. 11.19).

Таблица 11.19

Оптимальный план таблицы 11.15ж	Оптимальный план для $S_0=8000$	
	Этап I: P_5	Этап II: P_6
$X_7=150$ $X_1=35\ 000$ $X_2=5\ 000$ $X_3=30\ 000$	$X_7=150-2000 \cdot 1/200=140$ $X_1=35\ 000-2000 \cdot 0=35\ 000$ $X_2=5\ 000-2000 \cdot 5/2=0$ $X_3=30\ 000-2000 \cdot 0=30\ 000$ $X_5=0+2000=2000$	$X_7=140-6000 \cdot 1/300=120$ $X_1=35\ 000-6000 \cdot 0=35\ 000$ $X_2=0-6000 \cdot 0=0$ $X_3=30\ 000-6000 \cdot 8/3=14\ 000$ $X_5=2000-6000 \cdot 0=2000$ $X_6=0+6000=6000$
$Z_0-C_0=\$25\ 000$	$Z_0-C_0=25\ 000-2000 \cdot 0,06=\$24\ 880$	$Z_0-C_0=24\ 880-6000 \cdot 0,2533=\$23\ 360$

Аналогично для $13\ 250 < S_0 < 27\ 250$ оптимальный план можно получить за три шага, причем на отдельном шаге имеем

$$\left. \begin{aligned} X_1 &= 35\ 000 - (S_0 - 13\ 250) (5/2), & X_4 &= S_0 - 13\ 250, \\ X_5 &= 2000, & X_6 &= 11\ 250, & X_7 &= 102,5 - (S_0 - 13\ 250) (1/200), \\ Z_0 - C_0 &= \$22\ 030,00 - \$ (S_0 - 13\ 250) (0,28). \end{aligned} \right\} \quad (14)$$

Изменения машинного времени участков. Величины $Z_7 - C_7 = 0$, $Z_8 - C_8 = 20$, $Z_9 - C_9 = 14$ и $Z_{10} - C_{10} = 10^2/3$ в таблице 11.15ж представляют собой выигрыш (в виде дополнительной прибыли), который можно получить за счет использования одной единицы дополнительного машинного времени на участках 1, 2А, 2АА и 2В соответственно. Интересно отметить, что было бы более выгодно использовать дополнительное (сверхурочное) время на участке 2АА, чем на участке 2В. Кроме того, $Z_7 - C_7 = 0$ свидетельствует о том, что поскольку на участке 1 осталось лишних 150 часов, то нет смысла использовать дополнительные часы на этом участке.

Исходя из условия $X_{10} \geq 0$, можно исследовать элементы столбцов, как в предыдущем примере, и определить, что было бы выгодно использовать до 300, 300 и 450 дополнительных часов на участках 2А, 2АА и 2В соответственно. (Эти цифры получены из условия, что лишние (свободные) часы участка 1 будут использованы, но без сверхурочного времени на этом участке.) Полагая, что H_0 означает дополнительные часы, получим предельные значения H_0 для различных участков из следующих выражений:

$$150 + H_0(-1/2) \geq 0,$$

что дает $H_0 \leq 300$, и

$$150 + H_0(-1/3) \geq 0,$$

что дает $H_0 \leq 450$. Кроме того, для любого H_0 , лежащего в заданных пределах, получим следующее выражение для суммарной прибыли:

$$(Z_0 - C_0) = 25\,000 + H_0 \begin{vmatrix} 20 \\ 14 \\ 10^2/3 \end{vmatrix}.$$

В общем случае, если на участках 2А, 2АА и 2В будут использованы дополнительно H_1 , H_2 и H_3 часов, то получим суммарную прибыль в виде

$$(Z_0 - C_0) = 25\,000 + 20H_1 + 14H_2 - 10^2/3 H_3. \quad (15)$$

Изменения производительности. Изменения производительности могут объясняться использованием новых штампов, специальных инструментов, усовершенствованием производственного процесса или даже использованием новых машин. Чаще всего в таких случаях требуется совершенно новое решение. Однако в некоторых случаях все же можно использовать оптимальное решение исходной задачи. Например, предположим, что продолжительность обработки одного изделия R на участках 2А и 2АА уменьшилась (или ее можно уменьшить) с 0,02 часа до 0,0175 часа. Поскольку комбинация 1 — 2А наиболее выгодна при производстве изделия R , то прежде всего используем все наличные машинные часы участка 2А (т. е. 700 часов). Это означает, что потребуется 400 машинных часов на участке 1. Переходя ко второму, наиболее выгодному методу*, используем все 900 машинных часов участка 2В, что влечет за собой использование 300 машинных часов на участке 1. И последнее, используются 100 машинных

* Может возникнуть вопрос относительно выбора последовательности 1—2В, а не последовательности 1—2АА, несмотря на коэффициенты дохода, т. е. 0,32 доллара и 0,28 доллара соответственно. Сомнения вызваны различной трудоемкостью в обработке продукта R на 2В (0,03 часа на единицу продукции) и на 2АА (0,02 часа). Однако в данной постановке задачи решающим является не машинное время участков 2А, 2АА и 2В, а скорее машинное время участка 1; поэтому основным критерием является прибыль на единицу (или на единицу продукции относительно времени обработки на участке 1).

В таблицах 11.21 и 11.22 приводятся аналогичные оптимальные планы соответственно для $t_{2A,R} = 0,01 \cdot \frac{3}{11}$ и $t_{2A,R} = 0,01$.

Наконец, в связи с изменением производительности следует отметить, что при условии существенного улучшения процесса производства изделий S можно получить новый оптимальный план, который обеспечит производство определенного количества изделий S .

Изменение прибыли на изделие. Предположим, что из-за конкуренции или по другим причинам понизились рыночные цены на изделия R , например на 4 цента на изделие, что, естественно, вызывает уменьшение прибыли на 4 цента. Оптимальный план можно при этом получить, не решая задачу заново. Для заданного примера достаточно проделать следующие операции.

1. Переписать из таблицы 11.15ж элементы строк базисных векторов *. (Эта часть таблицы выделена в табл. 11.23а жирной линией.)

Таблица 11.23а

$C_j \backslash C_i$							0,36	0,24	0,28	0,72	0,64	0,60
	Базис	P_0	P_7	P_8	P_9	P_{10}	P_1	P_2	P_3	P_4	P_5	P_6
	P_7	150	1	$-1/2$	$-1/2$	$-1/3$				$1/200$	$1/200$	$1/300$
0,36	P_1	35 000		50			1			$5/2$		
0,24	P_2	5 000			50			1			$5/2$	
0,28	P_3	30 000				$100/3$			1			$8/3$
	Z_j	22 200		18	12	$91/3$	0,36	0,24	0,28	0,90	0,60	$0,74^{2/3}$
	$Z_j - C_j$	22 200		18	12	$91/3$				0,18	$-0,04^*$	$0,14^{2/3}$

2. Проставить новые значения прибылей в первую строку и в первый столбец:

$$C_1 = 0,36; C_2 = 0,24; C_3 = 0,28; C_4 = 0,72; C_5 = 0,64; C_6 = 0,60,$$

и все другие $C_j = 0$.

3. Рассчитать

$$Z_j = \sum_{i=1}^m C_i X_{ij}, \quad j = 0, 1, 2, \dots, n. \quad (7)$$

4. Вычислить $(Z_j - C_j)$.

* Эта часть матрицы называется *основной* и следует из ограничений. Следовательно, при изменении удельной прибыли основная часть матрицы остается неизменной.

(а) Если все $(Z_j - \bar{C}_j) \geq 0$, то старый план остается оптимальным для новых цен.

(б) Если хотя бы одна разность $(Z_j - C_j) < 0$, то продолжать вычисления согласно симплексному алгоритму до получения оптимального плана (см. табл. 11.23а и 11.23б).

Таблица 11.23б

$C_j \backslash C_i$							0,36	0,24	0,28	0,72	0,64	0,60
	Базис	P_0	P_7	P_8	P_9	P_{10}	P_1	P_2	P_3	P_4	P_5	P_6
	P_7	140	1	$-1/2$	$-3/5$	$-1/3$		$-0,002$		$1/200$		$1/300$
0,36	P_1	35 000		50			1			$5/2$		
$\rightarrow$ 0,64	P_5	2 000			20			$2/5$			1	
0,28	P_3	30 000				$100/3$			1			$8/3$
	$Z_j - C_j$	22 280		18	12,8	$91/3$		0,016		0,18		$0,142/3$

В таблице 11.24 приведен новый оптимальный план (с учетом изменения заданных цен). Таким образом, если использовать старый оптимальный план при новой системе цен, то общая прибыль уменьшится на 2800 долларов. В то же время при использовании нового оптимального плана прибыль составляет 22 280 долларов, т. е. минимальное сокращение прибыли составит 2720 долларов. Это достигается заменой 5000 изделий R (1 — 2AA) на 2000 изделий S (1 — 2AA).

Таблица 11.24

Исходный оптимальный план	Новый оптимальный план
$X_7 = 150$ $X_1 = 35\,000$ $X_2 = 5\,000$ $X_3 = 30\,000$ $Z_0 - C_0 = \$ 22\,200 *$	$X_7 = 140$ $X_1 = 35\,000$ $X_5 = 2\,000$ $X_3 = 30\,000$ $Z_0 - C_0 = \$ 22\,280$
* Величина $Z_0 - C_0$ отражает изменения в цене продукта R.	

Следует отметить, что симплексный метод не только дает возможность изучать влияние изменений цен и производительности выпуска продукции. С помощью симплексного метода можно также установить, как изменится оптимальный план при ошибочной оценке этих величин.

ЗАКЛЮЧЕНИЕ

В заключение можно отметить, что транспортный метод решения задач линейного программирования включает простейшие арифметические операции и поэтому желательно применять его для решения задач большой размерности всегда, когда это возможно. Кроме решения задач, подобных задаче оптимального распределения товарных вагонов, транспортный метод можно использовать для размещения новых складов и предприятий, для уменьшения простоев на производстве и для распределения изделий по машинам. Однако, в отличие от симплексного метода, с помощью транспортного метода нельзя решать общие задачи линейного программирования*.

С другой стороны, важно отметить, что с помощью симплексного метода** можно не только решать общие задачи линейного программирования, но и определять оптимальные решения новых задач, которые получены из исходных при введении новых ограничений или при изменении параметров задачи. Это очень существенно в связи с тем, что такие изменения или ограничения можно проанализировать заранее, что дает возможность руководству правильно решить ряд вопросов, связанных с (1) использованием дополнительных рабочих смен, (2) использованием сверхурочного времени на одних участках и основного времени на других, (3) добавлением ряда машин (дополнительное машинное время), (4) использованием новых машин, специальных инструментов или усовершенствованием производственного процесса (уменьшением себестоимости), (5) изменением цен в связи с конкуренцией, (6) затратами на производство продукции, обеспечивающей поддержание репутации фирмы (что связано с потерями прибыли), (7) вопросами стратегии сбыта, (8) оптимальной номенклатурой продукции.

Следует также упомянуть о неаналитических методах. Неаналитические методы решения применяются в тех случаях, когда размерность задач настолько велика, что аналитические методы, подобные симплексному и транспортному методам, практически неприменимы. Эти методы рассматриваются в [1] и [35], кроме того, в главе 13 приведен пример использования неаналитического метода.

В заключение сформулируем еще раз общую задачу линейного программирования:

З а д а ч а (формулировка I). Определить значения переменных $X_1, X_2, X_3, \dots, X_n$, которые максимизируют (минимизируют) функцию

$$Z = X_1C_1 + X_2C_2 + \dots + X_nC_n$$

при ограничениях

$$\begin{aligned} X_j &\geq 0, & j &= 1, 2, \dots, n, & X_1a_{11} + X_2a_{12} + \dots + X_na_{1n} &= b_1, \\ & & & & X_1a_{21} + X_2a_{22} + \dots + X_na_{2n} &= b_2, \\ & & & & \dots & \dots \\ & & & & X_1a_{m1} + X_2a_{m2} + \dots + X_na_{mn} &= b_m, \end{aligned}$$

* Связь между транспортным и симплексным методами с математической точки зрения показана Чарнсом и Купером [3].

** Здесь не упоминается о связи между линейным программированием и теорией игр, которая рассматривается в 18-й главе. О связи между этими дисциплинами см. книгу Мак-Кинси [33].

где a_{ij} , b_i и C_j — заданные постоянные, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, или, считая, что заданы векторы-столбцы

$$P_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad P_0 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad j = 1, 2, \dots, n.$$

Можно сформулировать задачу следующим образом.

Задача (формулировка II). Определить значения переменных $X_1, X_2, \dots, X_n$, которые максимизируют (минимизируют) линейный функционал

$$Z = X_1 C_1 + X_2 C_2 + \dots + X_n C_n \equiv \sum_{j=1}^n X_j C_j,$$

при условии, что

$$X_j \geq 0, \quad j = 1, 2, \dots, n,$$

и

$$X_1 P_1 + X_2 P_2 + \dots + X_n P_n \equiv \sum_{j=1}^n X_j P_j = P_0.$$

Приложение 11А

Второй способ оценки клеток в транспортном методе. В 11-й главе изложен метод (часть общего транспортного метода) «оценки» тех клеток таблицы возможного решения, которые не содержат выделенных чисел, т. е. метод оценки «стоимостей выигрыша», связанных с использованием других планов, отличных от плана, представленного в исходной таблице. В данном приложении описывается второй метод, который не только дает возможность достаточно просто производить такие оценки, но, кроме того, позволяет определить минимальные стоимости, связанные с отклонением от исходной программы*.

Чтобы сравнить этот метод оценки со старым, вернемся к таблицам 11.2 и 11.3, в которых приведены удельные затраты и исходный план транспортной задачи. Сначала необходимо образовать новую таблицу (табл. 11А.1), соответствующую таблице 11.3, но заполненную числами, выражающими удельные затраты, а не количество перевозимого продукта. В таблице 11А.1 эти числа выделены жирным шрифтом.

Добавляем к таблице 11А.1 столбец под названием «Переменные строк» и строку под названием «Переменные столбцов» и вычислим эти величины следующим образом.

1. Сначала выберем произвольно одну из переменных (потенциалов) в одном из столбцов или в одной из строк. Например, выберем значение 0 (нуль) в строке S_1 .

* Представленный здесь метод оценки является видоизменением первоначально предложенного Данцигом метода (см. Купманс [29, гл. ХХI]) и частью метода, описанного Гендерсоном и Шлейфером [25]. Относительно определения стоимостей, связанных с отклонением от оптимального плана, см. [25].

Таблица 11А.1

Удельные затраты и фиктивные затраты, соответствующие первому возможному плану

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5	Переменные столб
S_1	-10	-20	-5	-27	-21	0
S_2	-13	-23	-8	-30	-24	-3
S_3	7	-3	12	-10	-4	17
Переменные столбцов	-10	-20	-5	-27	-21	

2. Затем для каждой клетки в строке S_1 , которая содержит выделенное число, представляющее одно из чисел исходного плана, определим значение переменной (которая может быть положительной, отрицательной или нулевой). Величина этой переменной должна быть такой, чтобы сумма переменных строки и столбца была равна удельным затратам соответствующей клетки таблицы.

Другими словами, если r_i есть значение переменной i -й строки, C_j — значение переменной j -го столбца, а C_{ij} — удельная стоимость клетки, стоящей в i -й строке и в j -м столбце, то все переменные должны удовлетворять уравнению

$$r_i + c_j = C_{ij}. \quad (A.1)$$

Таким образом, приняв $r_1 = 0$, можно определить из уравнения (A.1) значения переменных:

$$C_1 = -10, \quad C_2 = -20, \quad C_3 = -5.$$

3. Затем, зная $C_3 = -5$ и $C_{23} = -8$, определим $r_2 = -3$.

4. Поскольку $r_2 = -3$ и $C_{24} = -30$, то $C_4 = -27$.

5. Из $C_4 = -27$ и $C_{34} = -10$ получаем $r_3 = +17$.

6. Наконец, из $r_3 = +17$ и $C_{35} = -4$ получаем $C_5 = -21$.

Следует заметить, что этот процесс определения значений переменных применим к любому невырожденному плану, т. е. к плану, содержащему точно $m + n - 1$ ненулевых элементов, при условии, что в матрице исходной задачи содержится m строк и n столбцов. (Любой план, в котором содержится меньше $m + n - 1$ ненулевых элементов, называется *вырожденным*. Методы решения в случае вырожденности разработаны Чарнсом и Купером [8], Гендерсоном и Шлейфером [25], Данцигом [10] и другими.)

Вычислив все переменные столбцов и строк таблицы 11А.1, можно заполнить оставшиеся клетки таблицы числами, полученными согласно уравнению (A.1). Эти числа набраны в таблице 11А.1 обычным шрифтом.

Таблица 11А.2

Оценки клеток для первого плана

Пункты назначения Пункты отправления	D_1	D_2	D_3	D_4	D_5
S_1	0	0	0	-18	-11
S_2	-11	-13	0	0	-18
S_3	8	17	19	0	0

После заполнения таблицы 11А.1 можно получить оценки различных клеток следующим образом. Образует новую таблицу (табл. 11А.2), которая состоит из чисел, полученных путем вычитания чисел (удельных затрат) таблицы 11.2 из чисел таблицы 11А.1. Это можно записать символически *: {таблица 11А.2} = {таблица 11А.1} — {таблица 11.2}. Очевидно, что клетки, соответствующие плану задачи, содержат нули. Эти нули выделены жирным шрифтом в таблице 11А.2. Остальные числа в таблице являются оценками соответствующих клеток и дают возможность судить об оптимальности выбранного плана. (Сравнение с таблицей 11.4 показывает идентичность всех чисел.) Получив оценки клеток, продолжаем процесс согласно приведенному выше алгоритму.

Приложение 11Б

Геометрическая интерпретация задачи линейного программирования. В этом приложении приведена геометрическая интерпретация задач линейного программирования на примере двумерной задачи.

Задача. Определить $X, Y \geq 0$, которые минимизируют функцию

$$Z = 2X + 5Y$$

при условиях

$$X \leq 4, \quad Y \leq 3, \quad X + 2Y \leq 8. \quad (\text{Б.1})$$

Этой системе линейных неравенств удовлетворяет выпуклое множество точек, ограниченных многоугольником $OABCD$ (рис. 11Б.1). То есть любая точка (X, Y) внутри или на границе многоугольника удовлетворяет всем уравнениям (Б.1). Следовательно, существует бесконечное множество решений системы (Б.1). Задача линейного программирования заключается в том, чтобы выбрать из этого бесконечного множества точек одну или несколько точек, которые максимизируют функцию $Z = 2X + 5Y$.

* Очевидно, можно упростить порядок вычислений, используя одну матрицу вместо двух. Например, удельные стоимости можно поместить в верхнем правом углу каждой клетки, а значения C_{ij} — в верхнем левом углу. Тогда оценки можно получить из одной и той же таблицы.

Функция $Z = 2X + 5Y$ представляет собой однопараметрическое семейство параллельных прямых линий, т. е. эта функция представляет собой семейство параллельных прямых линий (наклон $-\frac{2}{5}$), причем Z увеличивается по мере удаления линии от начала координат (см. рис. 11Б.2). Таким образом, исходную задачу можно рассматривать как задачу определения такой линии семейства $2X + 5Y = Z$, которая 'наиболее удалена от начала координат, при условии что она проходит по крайней мере через одну точку многоугольника $OABCD$.

На рис. 11Б.3 показано, как проходят различные линии семейства $Z = 2X + 5Y$ относительно многоугольника $OABCD$. Из рисунка видно, что координаты точки B являются решением задачи. Точка B является пересечением линий $Y=3$ и $X+2Y=8$. Следовательно, координаты точки B (2, 3) и значение функции $Z_{\max} = 2 \cdot 2 + 5 \cdot 3 = 19$.

Чтобы показать геометрически, что представляет собой решение задачи с помощью симплексного метода, в таблицах 11Б.1 приведено решение задачи рис. 11Б.3. Из таблиц 11Б.1 можно видеть, что решение переходит из точки ($X \equiv X_1 = 0$, $Y \equiv X_2 = 0$) в точку ($X \equiv X_1 = 0$, $Y \equiv X_2 = 3$), а затем в точку ($X \equiv X_1 = 2$, $Y \equiv X_2 = 3$) или на рис. 11Б.3 из точки O (начало координат) в точку A , а затем в точку B .

В более общем случае, если мы назовем вершины O , A , B , C и D *крайними* точками многоугольника $OABCD$, то оптимальное решение задачи линейного программирования находится в одной из этих точек, причем

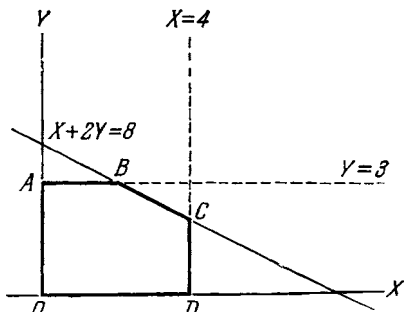


Рис. 11Б.1. Область, соответствующая ограничениям приложения 11Б.

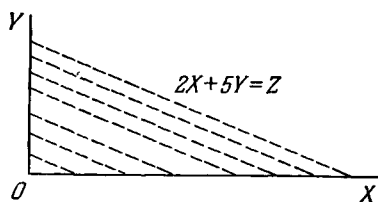


Рис. 11Б.2. Семейство параллельных прямых $Z = 2X + 5Y$.

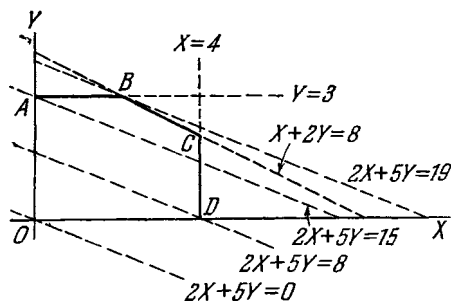


Рис. 11Б.3. Геометрическое решение задачи линейного программирования.

для того, чтобы попасть в эту точку, следует переходить из одной крайней точки в другую *. В нашем примере при определении решения переходим из крайней точки O в крайнюю точку A и наконец в крайнюю точку B .

* Математически многоугольник $OABCD$ представляет собой выпуклое множество точек, т. е. если заданы любые две точки этого многоугольника, то все точки отрезка, соединяющего эти две точки, также принадлежат многоугольнику. Крайней точкой выпуклого множества является любая точка этого множества, которая не лежит на отрезке, соединяющем любые две точки этого множества. Так, крайними точками многоугольника $OABCD$ являются точки O , A , B , C и D .

Таблица 11Б.1а

$C_i \backslash C_j$		0	0	0	0	2	5
	Базис	P_0	P_3	P_4	P_5	P_1	P_2
0	P_3	4	1	0	0	1	0
0	P_4	3	0	1	0	0	1
0	P_5	8	0	0	1	1	2
	Z_j	0	0	0	0	0	0
	$Z_j - C_j$	0	0	0	0	-2	-5

Таблица 11Б.1б

$C_i \backslash C_j$		0	0	0	0	2	5
	Базис	P_0	P_3	P_4	P_5	P_1	P_2
0	P_3	4	1	0	0	1	0
5	P_2	3	0	1	0	0	1
0	P_5	2	0	-2	1	1	0
	$Z_j - C_j$	15	0	5	0	-2	0

Изменим условия задачи.

Задача. Определить $X, Y \geq 0$, которые максимизируют $Z = X + 2Y$ при наличии ограничений

$$X \leq 4, \quad Y \leq 3, \quad X + 2 \leq 8.$$

Нетрудно видеть (рис. 11Б.4), что решение находится в крайней точке В или в крайней точке С. Это объясняется тем, что $X + 2Y = 8$ является

Таблица 11Б.1в

$c_j \backslash c_i$		0	0	0	0	2	5
	Базис	P_0	P_3	P_4	P_5	P_1	P_2
0	P_3	2	1	2	-1	0	0
5	P_2	3	0	1	0	0	1
2	P_1	2	0	-2	1	1	0
	$Z_j - C_j$	19	0	1	2	0	0

одновременно одним из ребер многоугольника $OABCD$ и также одной из прямых семейства линий $Z = X + 2Y$. Следовательно, $B = (2, 3)$ и $C = (4, 2)$ оба являются решениями и дают ответ $Z_{\max} = 8$.

Кроме того, любая выпуклая линейная комбинация точек B и C также будет решением, т. е. множеством всех точек (X^*, Y^*) , заданных в виде

$$X^* = a(2) + (1-a)(4),$$

$$y^* = a(3) + (1-a)(2),$$

где $0 \leq a \leq 1$. Геометрически (X^*, Y^*) представляет собой множество точек, образующих отрезок BC . При $a = 0$ получаем крайнюю точку C , при $a = 1$ — крайнюю точку B и при изменении a от 0 до 1 перемещаемся вдоль отрезка BC из точки C в точку B .

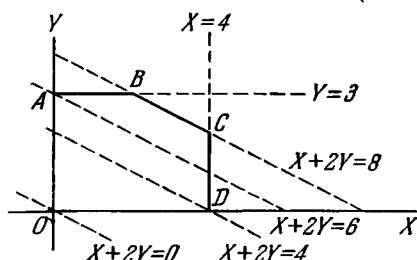


Рис. 11Б.4. Геометрическое решение задачи линейного программирования, имеющей больше одного оптимального решения.

Приложение 11В

Проблема двойственности в линейном программировании. В этом приложении рассматривается проблема двойственности в линейном программировании и, в частности, показывается, как перейти от исходной задачи линейного программирования к двойственной задаче. Кроме того, показывается, как, зная решение исходной задачи линейного программирования, определить решение двойственной задачи. Рассмотрим эти вопросы на примере задачи, приведенной в приложении 11Б.

Задача. Определить $X, Y \geq 0$, которые максимизируют функцию $Z = 2X + 5Y$ при ограничениях

$$X \leq 4, \quad Y \leq 3, \quad X + 2Y \leq 8. \quad (B.1)$$

Эту задачу можно представить в табличном виде (см. табл. 11В.1). Ограничения можно получить, считая тонкую вертикальную линию за знак «+»

Таблица 11В.1

	X	Y	
	1	0	4
	0	1	3
	1	2	8
max	2	5	

Таблица 11В.2

	X	Y	min
W_1	1	0	4
W_2	0	1	3
W_3	1	2	8
max	2	5	

и жирную вертикальную линию за знак « $\leq$ ». Кроме того, функция цели, т. е. $2X + 5Y$, приведена в нижней строке. Чтобы получить двойственную задачу, построим таблицу 11В.2 на основе таблицы 11В.1. Затем, читая каждый столбец сверху вниз, получим двойственную задачу.

Двойственная задача. Минимизировать $g = 4W_1 + 3W_2 + 8W_3$ при следующих ограничениях*:

$$W_1 + W_3 \geq 2, \quad W_2 + 2W_3 \geq 5. \quad (\text{B.2})$$

При решении задачи максимизации с помощью симплексного метода в приложении 11Б получены следующие результаты:

$$\begin{aligned} Z_{\max} &= 19, \\ \left. \begin{aligned} X_1 &= 2, & Z_1 - C_1 &= 0, \\ X_2 &= 3, & Z_2 - C_2 &= 0, \\ X_3 &= 2, & Z_3 - C_3 &= 0, \\ X_4 &= 0, & Z_4 - C_4 &= 1, \\ X_5 &= 0, & Z_5 - C_5 &= 2. \end{aligned} \right\} \quad (\text{B.3}) \end{aligned}$$

Переменные X_3 , X_4 и X_5 являются дополнительными. Теперь перенумеруем $Z_j - C_j$ так, чтобы начать с дополнительных переменных, причем новые значения $Z_j - C_j$ обозначим через $Z'_j - C'_j$. При этом получим:

$$\left. \begin{aligned} Z'_1 - C'_1 &= 0 \quad (\text{соответствует } Z_3 - C_3), \\ Z'_2 - C'_2 &= 1 \quad (\text{соответствует } Z_4 - C_4), \\ Z'_3 - C'_3 &= 2 \quad (\text{соответствует } Z_5 - C_5), \\ Z'_4 - C'_4 &= 0 \quad (\text{соответствует } Z_1 - C_1), \\ Z'_5 - C'_5 &= 0 \quad (\text{соответствует } Z_2 - C_2). \end{aligned} \right\} \quad (\text{B.4})$$

* Неравенства превращаются в равенства путем вычитания неотрицательных дополнительных переменных. Но, поскольку переменные с коэффициентом -1 не могут войти в базис, можно ввести искусственные переменные для получения базиса. Так, неравенство $W_1 + W_3 \geq 2$ сначала превращается в равенство $W_1 + W_3 - W_4 = 2$. Затем вводится искусственная переменная W_6 и получается равенство $W_1 + W_3 - W_4 + W_6 = 2$. Подробное изложение этого вопроса см. в книге Чариса, Купера и Гендерсона [4].

Затем, положив $W_j = Z'_j - C'_j$, получим решение двойственной задачи минимизации.

Если решать задачу минимизации с помощью симплексного метода, то получим те же результаты:

$$\begin{aligned} g_{\min} &= 19, \\ \left. \begin{aligned} W_1 &= 0, & -(g_1 - b_1) &= 2, \\ W_2 &= 1, & -(g_2 - b_2) &= 0, \\ W_3 &= 2, & -(g_3 - b_3) &= 0, \\ W_4 &= 0, & -(g_4 - b_4) &= 2, \\ W_5 &= 0, & -(g_5 - b_5) &= 3, \end{aligned} \right\} \quad (B.5) \end{aligned}$$

где b_j — соответствующие коэффициенты при переменных W_j в линейной функции g .

И обратно, если дано решение задачи минимизации (т. е. если даны уравнения (B.5)), можно определить решение двойственной задачи максимизации, если перенумеровать значения $-(g_j - b_j)$, начиная с дополнительной переменной W_4 . При этом получим решение (B.3).

Можно показать, что для двойственных задач $Z_{\max} = g_{\min}$; другими словами, эти две задачи эквивалентны*. Поэтому при решении задачи линейного программирования мы можем решать любую из задач двойственной пары. Практически считается, что число итераций, потребных для решения задачи линейного программирования, в $(1 + 1,5)$ раза больше числа строк (т. е. ограничений). Поэтому можно сократить вычисления, выбирая одну из двойственных задач, особенно в тех случаях, когда существует значительная разница в числе строк каждой из двух задач.

Приложение 11Г

Сокращенный метод решения задач линейного программирования. Одним из многих преимуществ транспортного и симплексного методов является возможность использования эвристических соображений для облегчения вычислений, связанных с нахождением оптимального решения. В транспортной задаче, в которую входит m строк и n столбцов, эти соображения (или удачная догадка) заключаются в определении возможного решения, т. е. в определении $m + n - 1$ переменных.

Эти переменные вносятся в таблицу в виде выделенных чисел. Затем переходят к оценке оставшихся клеток таблицы, чтобы определить, является ли выбранный план оптимальным.

Чтобы объяснить, как можно облегчить вычисления в симплексном методе, удобнее всего рассмотреть пример. Рассмотрим пример, приведенный в приложении 11Б, и покажем, как при удачной догадке можно построить соответствующую симплексную матрицу и перейти к оптимальному решению (если угаданное решение еще не является оптимальным).

З а д а ч а. Определить $X, Y \geq 0$, которые максимизируют $Z = 2X + 5Y$ при ограничениях

$$X \leq 4, \quad Y \leq 3, \quad X + 2Y \leq 8. \quad (Г.1)$$

* Доказательство см. у Чарнса, Купера и Гендерсона [4, стр. 72] или у Купманса [29, гл. XIX].

Превратив эту систему неравенств в равенства с помощью дополнительных переменных S_3, S_4, S_5 , получим

$$X + S_3 = 4, \quad Y + S_4 = 3, \quad X + 2Y + S_5 = 8. \quad (\Gamma.2)$$

Теперь предположим, что мы «догадываемся» или имеем какие-то основания предполагать, что в оптимальное решение не входит X , т. е. в окончательном решении содержатся переменные Y, S_3 и S_5 . Это означает, что $X = 0$ и $S_4 = 0$. Следовательно, чтобы получить «решение» (план), т. е. элементы базиса, которые проставляются в столбце P_0 симплексной таблицы, требуется подставить значения $X = 0$ и $S_4 = 4$ в уравнения $(\Gamma.2)$. При этом получим

$$S_3 = 4, \quad Y = 3, \quad 2Y + S_5 = 8, \quad (\Gamma.3)$$

откуда имеем

$$Y = 3, \quad S_3 = 4 \quad \text{и} \quad S_5 = 2. \quad (\Gamma.4)$$

Эти значения проставляются в столбце P_0 симплексной таблицы 11Г.1 (P_2 соответствует Y).

Таблица 11Г.1

$C_j \backslash C_i$						2	5
	Базис	P_0	P_3	P_4	P_5	P_1	P_2
0	P_3	4	1	0	0	1	0
5	$P_2(Y)$	3	0	1	0	0	1
0	P_5	2	0	-2	1	1	0
	Z_j	15	0	5	0	0	5
	$Z_j - C_j$	15	0	5	0	-2	0

Теперь нужно заполнить основную часть симплексной таблицы. Так как каждая величина $Z_j - C_j$ соответствует минимальной стоимости отклонения от оптимального плана при введении одной единицы X_j , то можно определить для каждого j соответствующую разность $Z_{1j} - C_j$ и значения X_{1j} в столбце X_j . Например, рассмотрим, каково будет отклонение от данного плана $Y = 3, S_3 = 4$ и $S_5 = 2$, если положить $X = 1$. Определим, как изменятся значения Y, S_3 и S_5 при изменении X на 1. Для этого надо решить следующие уравнения:

$$1 + S_3 = 4, \quad Y = 3, \quad 1 + 2Y + S_5 = 8. \quad (\Gamma.5)$$

Эти уравнения получены из уравнений $(\Gamma.1)$ после подстановки $X = 1$ и $S_4 = 0$. Решая систему $(\Gamma.5)$, получим

$$X = 1, \quad Y = 3, \quad S_3 = 3, \quad S_5 = 1. \quad (\Gamma.6)$$

Сравнивая (Г.4) и (Г.6), определим, насколько изменились значения переменных Y , S_3 и S_5 при изменении X на 1:

$$\Delta Y = 0, \quad \Delta S_3 = 1, \quad \Delta S_5 = 1. \quad (\text{Г.7})$$

Эти значения представлены в столбце P_1 (соответствует переменной X) таблицы 11Г.1. Аналогично для S_4 решаем систему

$$S_3 = 4, \quad Y + 1 = 3, \quad 2Y + S_5 = 8, \quad (\text{Г.8})$$

откуда получаем

$$Y = 2, \quad S_3 = 4, \quad S_5 = 4 \quad (\text{Г.9})$$

и соответственно

$$\Delta Y = 1, \quad \Delta S_3 = 0, \quad \Delta S_5 = -2. \quad (\text{Г.10})$$

Эти значения проставляются в столбце P_4 таблицы 11Г.1.

Векторы P_2 , P_3 и P_5 входят в базис, поэтому заполняем соответствующие столбцы 0 и 1, как показано в таблице 11Г.1.

Остается вычислить разность $Z_j - C_j$, чтобы определить, является ли полученный план оптимальным. Для этого сначала вычисляются значения Z_j по формуле, которая приводилась выше:

$$Z_j = \sum_i C_i X_{ij}, \quad (\text{Г.11})$$

а затем вычитается соответствующее C_j . Для базисных векторов P_2 , P_3 и P_5 значения $Z_2 - C_2$, $Z_3 - C_3$ и $Z_5 - C_5$ все равны 0. Другие значения вычисляются по формуле (Г.11):

$$\begin{aligned} Z_1 - C_1 &= 1(0) + 0(5) + 1(0) - 2 = -2, \\ Z_4 - C_4 &= 0(0) + 1(5) + (-2)(0) - 0 = 5. \end{aligned}$$

Теперь таблица 11Г.1 закончена, и так как полученный план не является оптимальным ($Z_1 - C_1$ — отрицательная величина), то переходим к оптимальному плану согласно симплексному алгоритму (см. приложение 11Б).

Легко заметить, что таблица 11Г.1 совпадает с таблицей 11Б.1б и получена без предварительной таблицы, подобной 11Б.1а. Этот же способ можно применить и к задачам большего размера, благодаря чему при удачном определении переменных, которые войдут в план, можно значительно сократить объем вычислений.

ЛИТЕРАТУРА

1. Arnoff E. Leonard, An Application of Linear Programming, Proceedings of the Conference on Operations Research in Production and Inventory Control, Case Institute of Technology, Cleveland, 1954.
2. Charnes A., Optimality and Degeneracy in Linear Programming, *Econometrica*, 20, 160—170 (1952).
3. Charnes A. and Cooper W. W., The Stepping Stone Method of Explaining Linear Programming Calculations in Transportation Problems, *Mgmt. Sci.*, 1, no. 1, Appendix (Oct. 1954).
4. Charnes A. and Henderson A., An Introduction to Linear Programming, John Wiley & Sons, New York, 1953.
5. Charnes A., Cooper W. W. and Mellon B., Blending Aviation Gasolines—A Study in Programming Interdependent Activities, in ref. 37.
6. Charnes A. and Lemke C. E., A Modified Simplex Method for Control of Round-off Error in Linear Programming, Association for Computing Machinery meeting, Pittsburgh, May 2, 1952.

7. Charnes A. and Lemke C. E., Minimization of Non-Linear Separable Convex Functionals, *Nav. Res. Log. Quart.*, 1, 301—312 (1954).
8. Cooper W. W. and Charnes A., Transportation Scheduling by Linear Programming, *Proceedings of the Conference on Operations Research in Marketing*, Case Institute of Technology, Cleveland, 1953.
9. Cooper W. W. and Farr D., *Linear Programming Models for Scheduling Manufactured Products*, Carnegie Institute of Technology, Pittsburgh, Sept. 1, 1952.
10. Dantzig G. B., Chaps. I, II, XX, XXI, and XXIII of ref. 29.
11. Dantzig G. B., Computational Algorithm of the Revised Simplex Method, RAND Memorandum RM-1266, 1953.
12. Dantzig G. B., Maximization of a Linear Function of Variables Subject to Linear Inequalities, Chap. XXI of ref. 29.
13. Dantzig G. B., Upper Bounds, Secondary Constraints, and Block Triangularity in Linear Programming, RAND Memorandum 1367, 1954.
14. Dantzig G. B. and Orchard-Hays W., Alternate Algorithm for the Revised Simplex Method, RAND Memorandum RM-1268, 1953.
15. Dantzig G. B. and Waters G., Product-Form Tableau for Revised Simplex Method, RAND Memorandum RM-1268-A, 1954.
16. Dantzig G. B., Orden A. and Wolfe P., The Generalized Simplex Method for Minimizing a Linear Form under Linear Inequality Restraints, RAND Memorandum RM-1264, 1954.
17. Dorfman R., Application of Linear Programming to the Theory of the Firm, University of California Press, Berkeley, 1951.
18. Dwyer Paul S., The Solution of the Hitchcock Transportation Problem with a Method of Reduced Matrices, University of Michigan, Ann. Arbor, Dec. 1955 (hctographed).
19. Flood M. M., On the Hitchcock Distribution Problem, *Pac. J. Math.*, 3, 369—386 (1953).
20. The Traveling-Salesman Problem, in J.F. McGloskey and J. M. Copping (eds.), *Operations Research for Management II*, The Johns Hopkins Press, Baltimore, 1956.
21. Ford I. R. and Fulkerson D. R., A Simple Algorithm for Finding Maximal Network Flows and an Application to the Hitchcock Problem, RAND Memorandum P-743, 1955.
22. Fulkerson D. R. and Dantzig G. B., Computation of Maximal Flows in Networks, RAND Memorandum RM-1489, 1955.
23. Gale D., Kuhn H. W. and Tucker A. W., Linear Programming and the Theory of Games, Chap. XIX of ref. 29.
24. Goldstein Leon, The Problem of Contract Awards, in ref. 34.
25. Henderson A. and Schlaifer R., *Mathematical Programming*, Harv. Busin. Rev. (May—June 1954).
26. Hildreth C. and Reiter S., On the Choice of a Crop Rotation Plan, Chap. XI of ref. 29.
27. Hitchcock F. L., The Distribution of a Product from Several Sources to Numerous Localities, *J. Math. Phys.*, 20, 224—230 (1941).
28. Hood W. C. and Koopmans T. C. (eds.), *Studies in Econometric Method*, Cowles Commission Monograph No. 14, John Wiley & Sons, New York, 1953.
29. Koopmans T. C. (ed.), *Activity Analysis of Production and Allocation*, Cowles Commission Monograph No. 13, John Wiley & Sons, New York, 1951.
30. Koopmans T. C., Optimum Utilization of the Transportation System, *Proceedings of the International Statistical Conferences*, Washington, 15 (1947). Cowles Commission Paper, New Series, No. 34.
31. Kuhn H. W., The Hungarian Method for the Assignment Problem, *Nav. Res. Log. Quart.*, 2, 83—98 (1955).
32. Kuhn H. W. and Tucker A. W., Contributions to the Theory of Games, *Annals of Mathematics Study* No. 24, Princeton University Press, Princeton, 1950.
33. McKinsey J. C. C., *Introduction to the Theory of Games*, McGraw-Hill Book Co., New York, 1953.
34. Neumann J. von and Morgenstern O., *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 1947.
35. New Machine Loading Methods, *Fact. Mgmt.*, 112, no. 1, 136-137 (Jan. 1954).
36. Orden A., Survey of Research on Mathematical Solutions of Programming Problems, *Mgmt. Sci.*, 1, 170—172 (1955).
37. Project SCOOP, *Symposium of Linear Inequalities and Programming*, Headquarters, U. S. Air Force, Washington, 1952.

38. Symonds Gifford H., Linear Programming: The Solution of Refinery Problems, Esso Standard Oil Co., New York, 1955.
39. Vidale M. L., A Graphical Solution of the Transportation Problem, J. Opns. Res. Soc. Am., 4, no. 2, 193—203 (Apr. 1956).
40. Voia W. D. F. and Orden A., The Personnel Assignment Problem, Symposium on Linear Inequalities and Programming, Project SCOOP, Headquarters, U.S. Air Force, Washington, 1952.

Глава 12

ЗАДАЧА О НАЗНАЧЕНИИ

ВВЕДЕНИЕ

В 11-й главе рассматривалось решение задач линейного программирования с помощью симплексного и транспортного методов. Однако существует несколько специальных задач линейного программирования, решение которых можно получить с помощью специальных методов, значительно сокращающих объем вычислений по сравнению с симплексным и транспортным методами. В этой главе рассматривается одна из таких задач — задача о назначении, которая встречается во многих областях планирования и распределения.

ЗАДАЧА О НАЗНАЧЕНИИ

Задачу о назначении можно сформулировать следующим образом: имеется n машин и n работ, задана эффективность выполнения каждой работы на каждой машине (таблица, в которой содержатся n^2 чисел, характеризующих эффективность, называется $n \times n$ - или n^2 -матрицей). Задача заключается в том, чтобы назначить на каждую машину одну и только одну работу таким образом, чтобы заданная функция эффективности была оптимальной.

Эта задача часто встречается во многих областях, не связанных с производством. Например, рассмотрим задачу, которая встает перед диспетчером машинного парка, содержащего прицепы и тягачи. Имеется n тягачей, расположенных в разных частях города, кроме того, на $m \leq n$ складах поставщиков расположены n груженых прицепов, которые нужно перевезти на станцию отправления. Задача диспетчера заключается в том, чтобы распределить каждый из n тягачей на соответствующий прицеп таким образом, чтобы определенная функция эффективности (т. е. общее пройденное расстояние или общее время перевозки) была оптимальной.

Следует отметить, что для n^2 -матрицы существует $n!$ возможных вариантов решения задачи. Наиболее тривиальным способом решения был бы перебор всех $n!$ возможных вариантов с оценкой стоимости каждого из них (стоимость каждого варианта оценивается с помощью заданной функции эффективности). В результате такого перебора был бы выбран вариант с наименьшей стоимостью. Легко заметить, что такой способ приводит к огромному числу вычислений даже для небольших значений n . Например, во многих задачах $n = 20$, при этом число возможных вариантов равно

$$n! = 20! = 2\,432\,902\,008\,176\,640\,000.$$

Для быстродействующей вычислительной машины, просчитывающей один вариант за 1 микросекунду и работающей 8 часов в день в течение 365 дней в году, потребовалось бы почти четверть миллиона лет для нахождения

оптимального решения. Этот пример наглядно показывает необходимость специальных приемов для решения задач о назначении.

Математическая модель. Математически задача о назначении формулируется следующим образом.

Д а н о: n^2 -матрица $A_0 = \|a_{ij}^{(0)}\|$ (здесь и в дальнейшем она называется *матрицей оценок*); $a_{ij}^{(0)} \geq 0$ для $i, j = 1, 2, \dots, n$ ($n \geq 3$).

О п р е д е л и т ь: n^2 -матрицу $X = \|x_{ij}\|$ (здесь и в дальнейшем она называется *матрицей назначения* или *перестановочной матрицей*), удовлетворяющую следующим условиям:

$$x_{ij} = x_{ij}^2, \quad i, j = 1, 2, \dots, n \quad (n \geq 3), \quad (1)$$

$$\sum_{i=1}^n x_{ij} = \sum_{j=1}^n x_{ij} = 1, \quad (2)$$

$$T = \sum_{i,j} a_{ij}^{(0)} x_{ij} = \text{minimum}^*. \quad (3)$$

Совместное выполнение условий (1) и (2) эквивалентно следующим условиям:

(а) $x_{ij} = \begin{cases} 1, & \text{если на машине } i \text{ выполняется работа } j, \\ 0 & \text{в противном случае;} \end{cases}$

(б) в каждой строке и столбце матрицы X один элемент равен 1, а остальные равны 0.

Условие (3) совместно с (1) и (2) дает возможность выбрать такой набор элементов из матрицы A_0 , в который входит только по одному элементу из каждой строки и столбца и сумма элементов которого минимальна. Когда все три условия выполняются одновременно, получаем перестановочную матрицу оптимального решения.

Т а б л и ц а 12.1

Матрица A_0 . Матрица оценок

Работа \ Машина	5	6	7	8
1	1	8	4	1
2	5	7	6	5
3	3	5	4	2
4	3	1	6	3

В качестве примера рассмотрим матрицу A_0 4×4 (или 4^2), приведенную в таблице 12.1. В этой таблице даны стоимости $a_{ij}^{(0)}$, соответствующие выполнению каждой из четырех работ (от 5 до 8) на каждой из четырех машин (от 1 до 4). Задача заключается в таком распределении работ по машинам, чтобы общая стоимость была минимальной. (В этой задаче $4! = 24$ возможных решения и соответственно 24 возможных перестановочные матрицы.)

В таблице 12.2 показана перестановочная матрица X оптимального решения. Этой матрице соответствует стоимость

$$T = a_{15}^{(0)} + a_{27}^{(0)} + a_{38}^{(0)} + a_{46}^{(0)} = 1 + 6 + 2 + 1 = 10 \text{ единиц.}$$

Для сравнения в таблице 12.3 приведена перестановочная матрица, соответствующая максимальной стоимости

$$T' = a_{16}^{(0)} + a_{28}^{(0)} + a_{35}^{(0)} + a_{47}^{(0)} = 8 + 5 + 3 + 6 = 22 \text{ единицы.}$$

* Задачу максимизации легко превратить в задачу минимизации, как показано в примере 12.2 этой главы.

Таблица 12.2

Матрица X^* . Оптимальная перестановочная матрица *

Работа \ Машина	5	6	7	8	$\sum_j x_{ij}$
1	1				1
2			1		1
3				1	1
4		1			1
$\sum_i x_{ij}$	1	1	1	1	

* Так же как в главе 11, в этой и последующих матрицах нули опущены.

Таблица 12.3

Матрица X' . Перестановочная матрица, соответствующая максимальной стоимости в задаче о назначении

Работа \ Машина	5	6	7	8	$\sum_j x_{ij}$
1		1			1
2				1	1
3	1				1
4			1		1
$\sum_i x_{ij}$	1	1	1	1	

Оба решения (табл. 12.2 и 12.3) удовлетворяют условиям (1) и (2). Можно отметить, кстати, что существуют задачи о назначении, где условие (1) можно заменить менее строгим условием *

$$x_{ij} \geq 0. \quad (4)$$

Однако в этой главе мы рассматриваем только такие случаи, для которых справедливо условие (1). Тем самым предполагаем, что на одной машине может выполняться только одна работа.

В таком виде (т. е. с уравнением (1)) задача о назначении является наиболее вырожденным случаем задачи распределения Хичкока [12] (которая называется также *транспортной задачей*), что было показано Фладом [5]. Существует несколько алгоритмов ** решения задачи распределения, которые были рассмотрены в 11-й главе. Следует указать, что методы теории графов, разработанные Купмансом и Рейтером [15] для решения невырожденных задач, Флад распространил на решение задачи о назначении, которую он рассматривал как частный вырожденный случай задачи распределения ***.

Можно отметить, что задачу, заданную уравнениями (2), (3) и (4) (т. е. менее строгую задачу о назначении, полученную при замене условия (1) условием (4)), можно решать с помощью симплексного метода линейного программирования. Если же задача задана уравнениями (1), (2) и (3), то специальные методы решения оказываются значительно эффективнее симплексного метода.

* См. [7, в частности, стр. 69—70].

** См. [1, 3, 9, 11, 13 и 15].

*** Связь между транспортной задачей (распределительной) и задачей о назначении с математической точки зрения рассмотрена в [7, стр. 62—63].

Метод решения. В течение последних пяти лет было разработано несколько методов решения задачи о назначении, заданной условиями (1), (2) и (3). И в настоящее время интерес к этой задаче не ослабевает. Значительный вклад в решение этой задачи внесли Двайер [3], Флад [5, 7, 8], Кун [16], Вото и Орден [18]. Кун [16] первым разработал алгоритм для решения задачи на основе следующей теоремы*, сформулированной Эгервари [4] и доказанной венгерским математиком Кёнигом в 1916 году.

Теорема. Если элементы матрицы разделить на два класса на основании свойства R , то минимальное число линий, содержащих все элементы со свойством R , равно максимальному числу таких элементов со свойством R , никакие два из которых не лежат на одной и той же линии.

В этой теореме линия означает строку или столбец матрицы. Эта теорема вместе со следующим важным свойством [17] является основой алгоритма Куна для решения задачи о назначении.

Если задана матрица оценок $A = \|a_{ij}\|$ и построена другая матрица $B = \|b_{ij}\|$, где

$$b_{ij} = a_{ij} - u_i - v_j, \quad (5)$$

а u_i и v_j — произвольные постоянные числа, то решение матрицы A совпадает с решением, полученным на основе матрицы B .

ДВОЙСТВЕННАЯ ЗАДАЧА ДЛЯ ЗАДАЧИ О НАЗНАЧЕНИИ

Двойственная задача** формулируется следующим образом. Определить систему постоянных величин u_i и v_j , для которых справедливы следующие соотношения:

$$a_{ij}^{(0)} \geq u_i + v_j \quad \text{для } x_{ij} = 0, \quad (6)$$

$$a_{ij}^{(0)} = u_i + v_j \quad \text{для } x_{ij} > 0. \quad (7)$$

Осуществим подстановку соотношения (6) в выражение (3). Получим

$$T \geq \sum_{i,j} (u_i + v_j) x_{ij} \geq \sum_i u_i x_{ij} + \sum_j v_j x_{ij},$$

то есть

$$T \geq \sum_i (u_i \sum_j x_{ij}) + \sum_j (v_j \sum_i x_{ij}).$$

Из выражения (2) следует

$$\sum_j x_{ij} = \sum_i x_{ij} = 1;$$

следовательно,

$$T \geq \sum_i u_i + \sum_j v_j.$$

Обозначим правую часть этого выражения через D . Задача минимизации $\sum_{ij} a_{ij}^{(0)} x_{ij}$ идентична следующей двойственной задаче:

Максимизировать

$$D = \sum_i u_i + \sum_j v_j \quad (8)$$

* Оказалось, что эта теорема была впервые доказана Фробениусом [2, 10] в 1912 году.

** Относительно проблемы двойственности в линейном программировании см. главу 11, приложение В.

при условии

$$a_{ij}^{(0)} \geq u_i + v_j.$$

В литературе D называется *суммой граничного множества* *.

Алгоритм решения задачи о назначении. Двайер [3], Флад [7] и Кун [16] использовали двойственную задачу для разработки эффективного метода решения задачи о назначении. В этой главе рассмотрен алгоритм Флада для решения задачи о назначении, поскольку для ручных вычислений он дает значительную экономию во времени по сравнению с другими существующими методами.

Метод Флада заключается в быстром уменьшении величин в матрице оценок A_0 , так что в конце получаем набор n независимых ** нулей по одному в каждой строке и в каждом столбце ***. Этот набор (не обязательно единственный) независимых нулей дает оптимальное решение задачи о назначении.

Рассмотрим этот метод более подробно.

(а) Шаг 1. Просмотреть столбцы матрицы оценок A_0 и обозначить минимальный элемент в каждом столбце **** $v_j^{(0)} (= \min a_{ij})$. Образовать новую матрицу A_1 путем замены $a_{ij}^{(0)}$ через $a_{ij}^{(1)} = a_{ij}^{(0)} - u_i^{(0)} - v_j^{(0)}$ для $i, j = 1, 2, \dots, n$ и $u_i^{(0)} \equiv 0$.

Легко заметить, что здесь мы уменьшаем элементы матрицы A_0 , используя условие (5). Матрица A_1 имеет по крайней мере один нулевой элемент в каждом столбце.

(б) Шаг 2. Определить минимальный набор S_1 линий, включающих все нулевые элементы матрицы A_1 (число таких линий равно n_1). Если $n_1 = n$, то в A_1 существует n нулевых элементов, никакие два из которых не лежат на одной линии. Элементы матрицы A_0 , соответствующие местам этих n элементов, представляют искомое оптимальное решение.

Ц и к л 0

(в) Шаг 2'. Если $n_1 < n$, исследовать строки A_1 и обозначить минимальный элемент в каждой строке через $u_i^{(1)} (= \min a_{ij}^{(1)})$. Образовать новую матрицу A_2 путем замены $a_{ij}^{(1)}$ через

$$a_{ij}^{(2)} = a_{ij}^{(1)} - u_i^{(1)} - v_j^{(1)} \quad \text{для } i, j = 1, \dots, n,$$

где $v_j^{(1)} \equiv 0$.

Матрица A_2 будет содержать по крайней мере один нулевой элемент в каждой линии.

(г) Теперь переходим опять к шагу 2, т. е. обозначим минимальный набор линий через S_2 (число линий n_2). Если $n_2 = n$, то в A_2 существует набор n нулевых элементов, никакие два из которых не лежат на одной линии. Положение этих n элементов матрицы A_2 определяет оптимальное решение.

Ц и к л 1.

(д) Шаг 3. Если $n_2 < n$, то обозначим через h_2 минимальный элемент среди элементов матрицы A_2 , не вошедших ни в одну из линий S_2 . Вычтем h_2

* Экономическая интерпретация величин u_i и v_j приводится в [16, стр. 87].

** Набор нулей называется *независимым*, если никакие два (или больше) нуля из этого набора не лежат на одной линии.

*** Это осуществляется на основе теоремы Кёнига. Отсюда название *венгерский метод* или *метод сокращенной матрицы*.

**** Это означает, что в каждом столбце j мы находим минимальный элемент $a_{ij}^{(0)}$ и называем его $v_j^{(0)}$.

из элементов A_2 , не входящих в S_2 , и добавим h_2 ко всем элементам, лежащим на пересечении линий в S_2 (если такие есть); назовем новую матрицу A_3 .

(е) Проверить, как в шаге 2, не выполняется ли равенство $n_3 = n$. Если $n_3 = n$, то оптимальное решение определяется, как в (г).

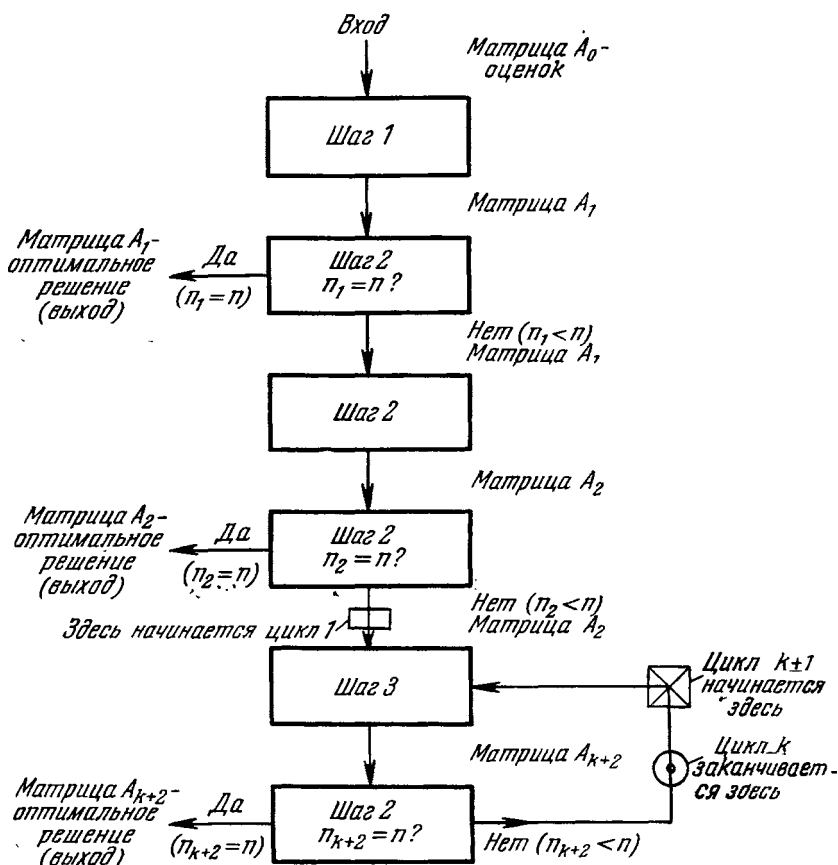


Рис. 12.1. Блок-схема алгоритма Флада решения задачи о назначении

(ж) Если $n_3 < n$, повторять цикл до тех пор, пока в некотором цикле k не получим $n_{k+2} = n$. Тогда оптимальное решение определяется из A_{k+2} , как в (г).

На рис. 12.1 описанный алгоритм решения показан в виде блок-схемы.

На 3-м шаге цикла 1 на входе мы имеем матрицу A_2 и на выходе — матрицу A_3 с характеристиками, приведенными в столбцах 1 и 2 таблицы 12.4. В общем случае на k -м цикле $k = 1, 2, \dots, K$ получим матрицы с характеристиками, приведенными в столбцах 3 и 4 таблицы 12.4.

Следует отметить, что шаг 3 цикла 1 непосредственно следует из условия (5) и эквивалентен следующим двум шагам.

(а) Заменить $a_{ij}^{(2)}$ на $a_{ij}^{(3)'}$, где

$$a_{ij}^{(3)'} = a_{ij}^{(2)} - u_i^{(2)'} - v_j^{(2)'}, \quad (9)$$

$$u_i^{(2)'} = h_2, \quad v_j^{(2)'} \equiv 0 \quad (i, j = 1, 2, \dots, n), \quad (10)$$

т. е. заменить $a_{ij}^{(2)}$ через $(a_{ij}^{(2)} - h_2)$ (по всем i, j).

Таблица 12.4

	(1)	(2)	(3)	(4)
	Вход Матрица A_2	Выход Матрица A_3	Вход Матрица A_{k+1}	Выход Матрица A_{k+2}
Элемент (i, j) матрицы	$a_{ij}^{(2)}$	$a_{ij}^{(3)}$	$a_{ij}^{(k+1)}$	$a_{ij}^{(k+2)}$
Составляющая сум- мы граничного мно- жества D	$\sum_i u_i^{(2)} + \sum_j v_j^{(2)}$	—	$\sum_i u_i^{(k+1)} + \sum_j v_j^{(k+1)}$	—
Минимальный на- бор линий	S_2	S_3	S_{k+1}	S_{k+2}
Число линий в ми- нимальном наборе	n_2	n_3	n_{k+1}	n_{k+2}
Минимальный эле- мент матрицы, не во- шедший ни в одну из линий набора	h_2	—	h_{k+1}	—

(б) Заменить $a_{ij}^{(3)'} на $a_{ij}^{(3)}$, где$

$$a_{ij}^{(3)} = a_{ij}^{(3)'} - u_i^{(2)''} - v_j^{(2)''}. \quad (11)$$

Теперь положим * $u_i^{(2)''} = -h_2$ и $v_j^{(2)''} = -h_2$ для элементов, содержащихся в обеих линиях i и j (записывается как $i, j \in S_2$),

$$\left. \begin{aligned} u_i^{(2)''} = -h_2 \quad \text{и} \quad v_j^{(2)''} = 0 & \quad \text{для} \quad i \in S_2, \quad \text{но} \quad j \notin S_2, \\ u_i^{(2)''} = 0 \quad \text{и} \quad v_j^{(2)''} = -h_2 & \quad \text{для} \quad j \in S_2, \quad \text{но} \quad i \notin S_2, \\ u_i^{(2)''} = 0 \quad \text{и} \quad v_j^{(2)''} = 0 & \quad \text{для} \quad i, j \notin S_2. \end{aligned} \right\} \quad (12)$$

Следовательно,

$$a_{ij}^{(3)} = \begin{cases} a_{ij}^{(3)'} + 2h_2 & \text{для } i, j \in S_2, \\ a_{ij}^{(3)'} + h_2 & \text{или для } i \in S_2, j \notin S \text{ или для } j \in S_2, i \notin S \\ a_{ij}^{(3)'} & \text{для } i, j \notin S_2. \end{cases} \quad (13)$$

Если мы теперь заменим $a_{ij}^{(3)'}$, как указано в шаге (а), то получим

$$a_{ij}^{(3)} = \begin{cases} a_{ij}^{(2)} + h_2, \text{ и } i \in S_2, \text{ и } j \in S_2 \text{ (т. е. для элементов,} \\ \quad \text{находящихся на пересечении } i \text{ и } j), \\ a_{ij}^{(2)} \text{ или для } i \in S_2, j \notin S \text{ или для } j \in S_2, i \notin S, \\ a_{ij}^{(2)} - h_2 \text{ для } i, j \notin S_2. \end{cases} \quad (14)$$

Но поскольку $u_i^{(2)} = u_i^{(2)'} + u_i^{(2)''}$, то

$$u_i^{(2)} = \begin{cases} 0 & \text{для } i \in S_2 \text{ (так как } h_2 - h_2 \equiv 0), \\ h_2 & \text{для } i \notin S_2. \end{cases} \quad (15)$$

* Символ $\in$ означает «принадлежит», символ $\notin$ означает «не принадлежит».

Аналогично, поскольку

$$v_j^{(2)} = v_j^{(2)'} + v_j^{(2)''},$$

то

$$v_j^{(2)} = \begin{cases} -h_2 & \text{для } j \in S_2 \text{ (так как } 0 - h_2 = -h_2), \\ 0 & \text{для } j \notin S_2. \end{cases} \quad (16)$$

Аналогичные замечания относятся к шагу 3 в последующих циклах. Величины u_i и v_j , входящие в условия (6) и (7), можно вычислить для задачи из K циклов по следующим формулам:

$$u_i = u_i^{(0)} + u_i^{(1)} + u_i^{(2)} + \dots + u_i^{(K+1)},$$

то есть

$$u_i = u_i^{(0)} + u_i^{(1)} + \sum_{k=1}^K u_i^{(k+1)}, \quad i = 1, \dots, n. \quad (17)$$

Аналогично

$$v_j = v_j^{(0)} + v_j^{(1)} + v_j^{(2)} + \dots + v_j^{(K+1)},$$

или

$$v_j = v_j^{(0)} + v_j^{(1)} + \sum_{k=1}^K v_j^{(k+1)}, \quad j = 1, 2, \dots, n. \quad (18)$$

Из выражения (8) и теоремы о двойственной задаче можно заключить, что стоимость, связанная с оптимальным решением задачи минимизации (условие (3)), совпадает со стоимостью, связанной с оптимальным решением задачи максимизации (условие (8)), где u_i и v_j вычисляются по формулам (17) и (18).

Уточнение. Из приведенного алгоритма можно сделать вывод, что начать решение задачи можно единственным путем. Поскольку нашей задачей является максимизация D , то желательно первый шаг делать таким, чтобы максимально увеличивать сумму граничного множества $(\sum_i u_i + \sum_j v_j)$. Это осуществляется следующим образом: сравнивается

$$\sum_j \min_i a_{ij}^{(0)} = \sum_j v_j^{(0)}$$

и

$$\sum_i \min_j a_{ij} = \sum_i u_i^{(0)}.$$

Если

$$\sum_j v_j^{(0)} \geq \sum_i u_i^{(0)},$$

то алгоритм начинается с шага 1 и продолжается без изменений, как описано выше. Если же

$$\sum_j v_j^{(0)} < \sum_i u_i^{(0)},$$

то в шаге 1 следует вычитать $u_i^{(0)}$ из каждого элемента i -й строки матрицы A_0 . Тогда в шаге 2' из каждого элемента матрицы A_1 вычитается $v_j^{(1)} = \min_i a_{ij}^{(1)}$.

Остальные шаги проводятся без изменений.

Важной особенностью этого алгоритма является его самокорректировка при следующих ошибках. Если на определенном цикле k мы ошибочно выбрали набор S'_{k+2} , который не является минимальным набором линий матрицы A_{k+2} , т. е. $S'_{k+2} > S_{k+2}$, то можно продолжать вычисления, потому что в дальнейшем ошибка будет автоматически исправлена *.

Пример 12.1. Чтобы проиллюстрировать применение описанного алгоритма к практическим задачам, рассмотрим следующий пример.

Компания по производству горючего обязана распределять горючее из главного хранилища различным промышленным потребителям. Для этой цели компания содержит машинный парк, в котором имеются тягачи и прицепы-цистерны. Поскольку слив горючего из цистерны на складе потребителя обычно занимает около 2 часов, водитель тягача оставляет наполненную цистерну на складе, забирает пустую цистерну из того же или другого склада и перевозит ее на главное хранилище, где она вновь наполняется.

Таблица 12.5

Матрица A_0 . Матрица оценок: стоимость перевозок тягач — цистерна в \$

Цистерна № \ Тягач №	11	12	13	14	15	16	17	18	19	20	$u_i^{(0)} = 0$
1	4,10	4,10	7,90	8,30	2,60	8,70	7,10	9,90	2,20	9,40	0
2	4,60	6,10	4,50	0,90	4,80	6,40	8,80	6,20	0	0,20	0
3	9,70	3,60	5,30	0,90	7,60	2,80	3,50	7,90	4,00	7,00	0
4	2,00	6,50	2,30	5,70	7,40	1,60	2,40	9,70	5,00	1,90	0
5	2,10	1,40	5,00	5,20	1,70	4,70	2,80	6,70	6,30	0,50	0
6	1,70	7,80	3,00	0	1,50	1,00	8,00	6,80	3,00	9,10	0
7	9,10	5,30	0,30	6,40	6,60	3,90	9,80	1,60	2,40	8,10	0
8	4,90	1,70	7,70	5,70	7,90	4,00	9,50	0,80	4,80	2,00	0
9	2,90	8,80	1,80	5,90	6,60	6,20	8,80	0,70	0,30	2,60	0
10	8,40	7,40	0,60	2,60	6,00	7,70	3,70	9,90	0,20	7,90	0
$v_j^{(0)} = \min_i a_{ij}^{(0)}$	1,70	1,40	0,30	0	1,50	1,00	2,40	0,70	0	0,20	

* Метод, в котором нет шага, связанного с проведением линий, описан в [3]

В этой задаче известно, что стоимость (здесь рассматривается только та часть стоимости, которая меняется в зависимости от назначения) перевозки пустых цистерн зависит от (1) пройденного расстояния (учитывается весь путь от склада, на котором тягач оставил наполненную цистерну, до главного хранилища с заходом на склад, где находится пустая цистерна) и (2) скорости перевозки. Зная эти данные, можно составить матрицу оценок (табл. 12.5), в которой приведены стоимости перевозки каждым тягачом каждой цистерны.

Таблица 12.6

Матрица A_1

[illegible]

Задача заключается в определении оптимального распределения, минимизирующего суммарную стоимость перевозки. Очевидным, но чрезвычайно трудным способом решения задачи был бы перебор всех возможных вариантов и выбор оптимального. Однако в данной задаче содержится $10! = 3\,628\,800$ возможных вариантов.

Можно использовать метод Купманса и Рейтера [15] решения транспортной задачи. Решение поставленной задачи этим методом заняло 3 часа ручных вычислений. Ту же задачу автор решил за 20 минут, используя приведенный в этой главе алгоритм.

Проиллюстрируем на данном примере использование этого алгоритма.

Шаг 1. Просматриваем столбцы матрицы A_0 , определяем минимальный элемент в каждом столбце $v_j^{(0)}$ и записываем его в нижней строке под матрицей A_0 . В нашем примере $\sum_j \min_i a_{ij}^{(0)} = \$9,20$ и $\sum_j \min_i a_{ij}^{(9)} = \$6,80$.

Таблица 12.7

Матрица A_2

Работа № Машина №											$u_i^{(2)}$
	11	12	13	14	15	16	17	18	19	20	
1	1,30	1,60	6,50	7,20	0	6,60	3,60	8,10	1,10	8,10	0,30
2	2,90	4,70	4,20	0,90	3,30	5,40	6,40	5,50	0	0	0,30
3	7,10	1,30	4,10	0	5,20	0,90	0,20	6,30	3,10	5,90	0,30
4	0,30	5,10	2,00	5,70	5,90	0,60	0	9,00	5,00	1,70	0,30
5	0,40	0	4,70	5,20	0,20	3,70	0,40	6,00	6,30	0,30	0,30
*** 6	0	6,40	2,70	0	0	0	5,60	6,10	3,00	8,90	0 ***
7	7,40	3,90	0	6,40	5,10	2,90	7,40	0,90	2,40	7,90	0,30
8	3,10	0,20	7,30	5,60	6,30	2,90	7,00	0	4,70	1,70	0,30
9	1,20	7,40	1,50	5,90	5,10	5,20	6,40	0	0,30	2,40	0,30
10	6,50	5,80	0,10	2,40	4,30	6,50	1,10	9,00	0	7,50	0,30
$v_j^{(2)}$	0	-0,30	-0,30	-0,30	-0,30	0	-0,30	-0,30	-0,30	-0,30	

Поскольку

$$\sum_j \min_i a_{ij}^{(0)} > \sum_i \min_j a_{ij}^{(0)},$$

то на 1-м шаге следует вычитать величины $v_j^{(0)}$. Вычитаем $v_j^{(0)}$ из каждого элемента каждого столбца j ($j = 11, \dots, 20$) матрицы A_0 . В результате получаем матрицу A_1 (табл. 12.6), которая содержит по крайней мере один нулевой элемент в каждом столбце.

Шаг 2. Минимальное число линий, проходящих через все нулевые элементы матрицы A_1 , равно 6. Получаем следующий набор линий:

$$S_1 = \{\text{линии } 2, 4, 5, 6, 7, 9\} \text{ и } n_1 = 6 < n \quad (n = 10).$$

Следовательно, оптимальное решение еще не получено.

Шаг 2'. Так как $n_1 < n$, то начинаем цикл 0, определяя минимальный элемент $u_i^{(1)}$ в каждой строке матрицы A_1 и записывая его в крайнем правом

Таблица 12.8

Матрица A_3

Работа №												$u_i^{(3)}$	
Машина №	11	12	13	14	15	16	17	18	19	20			
***	1	1,00	1,60	6,50	7,20	0	6,30	3,60	8,10	1,10	8,10	0	***
	2	2,60	4,70	4,20	0,90	3,30	5,10	6,40	5,50	0	0	0,10	
	3	6,80	1,30	4,10	0	5,20	0,60	0,20	6,30	3,10	5,90	0,10	
***	4	0	5,10	2,00	5,70	5,90	0,30	0	9,00	5,00	1,70	0	***
	5	0,10	0	4,70	5,20	0,20	3,40	0,40	6,00	6,30	0,30	0,10	
***	6	0	6,70	3,00	0,30	0,30	0	5,90	6,40	3,30	9,20	0	***
***	7	7,10	3,90	0	6,40	5,10	2,60	7,40	0,90	2,40	7,90	0	***
	8	2,80	0,20	7,30	5,60	6,30	2,60	7,00	0	4,70	1,70	0,10	
	9	0,90	7,40	1,50	5,90	5,10	4,90	6,40	0	0,30	2,40	0,10	
	10	6,20	5,80	0,10	2,40	4,30	6,20	1,10	9,00	0	7,50	0,10	
	$v_j^{(3)}$	0	-0,10	0	-0,10	0	0	0	-0,10	-0,10	-0,10		

Таблица 12.9

Матрица A_4

Работа № Машина №	11	12	13	14	15	16	17	18	19	20	$u_i^{(4)}$	
*** 1	1,00	1,70	6,50	7,30	0	6,30	3,60	8,20	1,20	8,20	0	***
*** 2	2,50	4,70	4,10	0,90	3,20	5,00	6,30	5,50	0	0	0	***
*** 3	6,70	1,30	4,00	0	5,10	0,50	0,10	6,30	3,10	5,90	0	***
*** 4	0	5,20	2,00	5,80	5,90	0,30	0	9,10	5,10	1,80	0	***
*** 5	0	0	4,60	5,20	0,10	3,30	0,30	6,00	6,30	0,30	0	***
*** 6	0	6,80	3,00	0,40	0,30	0	5,90	6,50	3,40	9,30	0	***
*** 7	7,10	4,00	0	6,50	5,10	2,60	7,40	1,00	2,50	8,00	0	***
8	2,70	0,20	7,20	5,60	6,20	2,50	6,90	0	4,70	1,70	0,20	
9	0,80	7,40	1,40	5,90	5,00	4,80	6,30	0	0,30	2,40	0,20	
*** 10	6,10	5,80	0	2,40	4,20	6,10	1,00	9,00	0	7,50	0	***
$v_j^{(4)}$	0	0	0	0	0	0	0	-0,20	0	0		

столбце таблицы 12.6. Вычитаем $u_i^{(1)}$ из каждого элемента строки i ($i = 1, \dots, 10$), в результате чего получаем матрицу A_2 (табл. 12.7), содержащую по крайней мере один нулевой элемент в каждой линии.

Шаг 2. Минимальное число линий, проходящих через все нулевые элементы матрицы A_2 , равно 9. Следовательно, оптимальное решение еще не получено, и мы переходим к шагу 3 (цикл 1).

Шаг 3. Определяем минимальный элемент h_2 среди элементов матрицы A_2 , не вошедших ни в одну из линий S_2 . В нашем примере таким элементом в таблице 12.7 является $h_2 = a_{4,11}^{(2)} = 0,30$. Добавим 0,30 ко всем элементам, лежащим на пересечении линий в S_2 , и вычтем 0,30 из всех элементов A_2 , не входящих в S_2 . Получим матрицу A_3 (табл. 12.8).

Повторяем шаги 2 и 3, начиная с матрицы A_3 , до тех пор, пока в цикле 3 не получим A_5 (табл. 12.9), в которой $n_5 = 10 = n$. Следовательно, оптимальное решение можно определить из матрицы A_5 . Это оптимальное

Таблица 12.10

Матрица A_5 . Матрица с оптимальным решением

Цистерна № Тягач №	11	12	13	14	15	16	17	18	19	20	
*** 1	1,00	1,70	6,50	7,30	[0]	6,30	3,60	8,40	1,20	8,20	***
*** 2	2,50	4,70	4,10	0,90	3,20	5,00	6,30	5,70	0	[0]	***
*** 3	6,70	1,30	4,00	[0]	5,10	0,50	0,10	6,50	3,10	5,90	***
*** 4	0	5,20	2,00	5,80	5,90	0,30	[0]	9,30	5,10	1,80	***
*** 5	[0]	0	4,60	5,20	0,10	3,30	0,30	6,20	6,30	0,30	***
*** 6	0	6,80	3,00	0,40	0,30	[0]	5,90	6,70	3,40	9,30	***
*** 7	7,10	4,00	[0]	6,50	5,10	2,60	7,40	1,20	2,50	8,00	***
*** 8	2,50	[0]	7,00	5,40	6,00	2,30	6,70	0	4,50	1,50	***
*** 9	0,60	7,20	1,20	5,70	4,80	4,60	6,10	[0]	0,10	2,20	***
*** 10	6,10	5,80	0	2,40	4,20	6,10	1,00	9,20	[0]	7,50	***

*** означает линии из набора S_5 .
 $S_5 = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$; $n_5 = 10 = n$; следовательно, оптимальное решение можно получить из набора нулевых элементов в A_5 .
 Оптимальное решение, которое является единственным для данного примера, выделено квадратными скобками [0].

решение дается набором независимых нулей матрицы A_5 . В таблице 12.10 эти нули выделены квадратными скобками [0] и занимают следующие клетки: 1—15, 2—20, 3—14, 4—17, 5—11, 6—16, 7—13, 8—12, 9—18 и 10—19. Это означает, что, например тягач 1 должен перевезти цистерну 15 и т. д.

Используя выражение (3), можно вычислить суммарные издержки, связанные с полученным решением. Для этого следует обратиться к матрице оценок A_0 (табл. 12.5). Получаем $T^* = 2,60 + 0,20 + 0,90 + 2,40 + 2,10 + 1,00 + 0,30 + 1,70 + 0,70 + 0,20 = \$12,10$.

Хотя для получения решения нет необходимости вычислять величины u_i и v_j , но можно проверить стоимость оптимального решения, используя выражения (15) — (18) и числа u_i и v_j из таблиц 12.5—12.10. Имеем

$$u_1 = u_1^{(0)} + u_1^{(1)} + u_1^{(2)} + u_1^{(3)} + u_1^{(4)} = 0 + 1,10 + 0,30 + 0 + 0 = 1,40,$$

$$u_2 = u_2^{(0)} + u_2^{(1)} + u_2^{(2)} + u_2^{(3)} + u_2^{(4)} = 0 + 0 + 0,30 + 0,10 + 0 = 0,40.$$

Таблица 12.11

Матрица A_0 . Иллюстрация двойственности

Цистерна № Тягач №	11	12	13	14	15	16	17	18	19	20	u_i
1	4,10	4,10	7,90	8,30	[2,60]	8,70	7,10	9,90	2,20	9,40	1,40
2	4,60	6,10	4,50	0,90	4,80	6,40	8,80	6,20	0	[0,20]	0,40
3	9,70	3,60	5,30	[0,90]	7,60	2,80	3,50	7,90	4,00	7,00	1,30
4	2,00	6,50	2,30	5,70	7,40	1,60	[2,40]	9,70	5,00	1,90	0,30
5	[2,10]	1,40	5,00	5,20	1,70	4,70	2,80	6,70	6,30	0,50	0,40
6	1,70	7,80	3,00	0	1,50	[1,00]	8,00	6,80	3,00	9,10	0
7	9,10	5,30	[0,30]	6,40	6,60	3,90	9,80	1,60	2,40	8,10	0,30
8	4,90	[1,70]	7,70	5,70	7,90	4,00	9,50	0,80	4,80	2,00	0,70
9	2,90	8,80	1,80	5,90	6,60	6,20	8,80	[0,70]	0,30	2,60	0,60
10	8,40	7,40	0,60	2,60	6,00	7,70	3,70	9,90	[0,20]	7,90	0,60
v_j	1,70	1,00	0	-0,40	1,20	1,00	2,10	0,10	-0,40	-0,20	

З а м е ч а н и е: «[]» в клетках i, j означают $x_{ij}=1$. Все остальные $x_{ij}=0$.

Аналогично можно вычислить $u_3, \dots, u_{10}$ и $v_1, \dots, v_{10}$. Эти числа вместе с величинами $a_{ij}^{(0)}$ показаны в таблице 12.11. Теперь определим величину

$$D^* = \sum_i u_i + \sum_j v_j = \$12,10,$$

т. е. эти величины совпадают.

Из таблицы 12.11 видно, что во всех случаях, когда $x_{ij}=1$ (a_{ij} , соответствующие этим значениям, выделены квадратными скобками), выполняется равенство $a_{ij} = u_i + v_j$. Во всех других случаях $a_{ij} > u_i + v_j$. Следовательно, условия (6) и (7) также выполняются.

Замечания относительно примера 12.1. Можно заметить, что с точки зрения определения окончательного решения, матрицы от A_0 до A_5 представляют собой одну и ту же задачу. Однако сокращенные матрицы (с умень-

шенными элементами) от A_2 до A_5 не являются единственными, поскольку обычно существует много различных способов выбора линий, покрывающих все нулевые элементы. В зависимости от набора линий после преобразования получаются различные матрицы.

МОДИФИКАЦИИ ЗАДАЧИ О НАЗНАЧЕНИИ

Рассмотрим теперь три различных видоизменения задачи о назначении. Матрица $n \times m$ ($m < n$).

Иногда в задаче об оптимальном назначении матрица не является квадратной. Такую матрицу, однако, легко превратить в квадратную, как показано в следующем примере.

Пример 12.2. Транспортная компания, ведающая перевозками сравнительно небольших грузов (которые можно перевозить в грузовиках), содержит парк грузовиков для внутригородских перевозок и парк грузовиков для междугородных перевозок. Внутригородские грузовики осуществляют местные перевозки и привозят товары на городскую товарную станцию, где их сортируют и передают на соответствующие междугородные грузовики. На товарной станции могут одновременно разместиться шесть внутригородских грузовиков. Помещение каждого грузовика на каждое из шести мест связано с определенной стоимостью (сортировки и передачи грузов).

Допустим, что нужно одновременно разместить на станции четыре внутригородских грузовика (номера от 1 до 4). В таблице 12.12 приведена матрица стоимостей A . Эту матрицу можно превратить в квадратную матрицу A_0 путем введения двух фиктивных грузовиков 5 и 6 (см. табл. 12.13). С назначением этих фиктивных грузовиков не связано никакой стоимости, поэтому соответствующие $a_{ij}^{(0)}$ в этих столбцах равны 0.

Таблица 12.13

Матрица A_0

Таблица 12.12
Матрица A . Матрица оценок: стоимость размещения грузовика на определенном месте

Грузовик \ Место	1	2	3	4
7	3	6	2	6
8	7	1	4	4
9	3	8	5	8
10	6	4	3	7
11	5	2	4	3
12	5	7	6	2

Грузовик \ Место	1	2	3	4	5	6
7	3	6	2	6	0	0
8	7	1	4	4	0	0
9	3	8	5	8	0	0
10	6	4	3	7	0	0
11	5	2	4	3	0	0
12	5	7	6	2	0	0
$v_j^{(0)} = \min_i a_{ij}^{(0)}$	3	1	2	2	0	0

Таблица 12.14

Матрица A_1 . Оптимальное решение представлено элементами [0]

Грузовик Место	1	2	3	4	5	6
7	0	5	[0]	4	0	0
8	4	[0]	2	2	0	0
9	[0]	7	3	6	0	0
10	3	3	1	5	0	[0]
11	2	1	2	1	[0]	0
12	2	6	4	[0]	0	0

Теперь задачу можно решать согласно приведенному алгоритму. Оптимальное решение задачи дано в таблице 12.14 в виде нулей, выделенных квадратными скобками. Это решение можно расшифровать следующим образом:

Поместить грузовик 1 на место 9,

Поместить грузовик 2 на место 8,

Поместить грузовик 3 на место 7,

Поместить грузовик 4 на место 12,

Оставить свободными места 10 и 11.

С таким размещением связана стоимость

$$$(3 + 1 + 2 + 2) = \$8,00.$$

В этом примере

$$v_j = v_j^{(0)} \quad \text{для } j = 1, \dots, 4$$

и

$$u_i = 0 \quad \text{для } i = 7, \dots, 12.$$

Легко видеть (табл. 12.13), что условия двойственности (6) и (7) также выполняются.

ЗАДАЧА МАКСИМИЗАЦИИ

Выше рассматривались примеры, связанные с минимизацией определенной функции. Легко показать, что тот же алгоритм можно применить и к решению задач максимизации.

Одной из таких задач является задача назначения персонала на работы при условии получения максимальных доходов. Метод решения таких задач

будет проиллюстрирован с помощью вымышленного примера, дающего, однако, представление о действительных задачах.

Пример 12.3. Предложим, что некая компания в небольшом торговом городе распределяет женский труд. У нее есть постоянный оплачиваемый штат из четырех женщин (номера от 1 до 4), которые являются специалистами каждая в своей области и, кроме того, разбираются в смежных областях. Вместе с тем у компании есть определенное количество девушек, которые не зачислены в штат, но обучаются различным работам и используются

Таблица 12.15

Матрица C . Матрица доходов:
предполагаемые доходы компании
за день (в долларах)

Работа \ Девушка	5	6	7	8
1	1	8	4	1
2	5	7	6	5
3	3	5	4	2
4	3	1	6	3

временно, когда увеличивается объем работ. Клиенты обычно платят за услуги в соответствии с произведенной работой (т. е. с количеством напечатанных писем, количеством оформленных накладных, заполненных документов и т. д.).

В один из дней компания получает заказ на четыре работы (номера от 5 до 8), для каждой из которых по прошлому опыту известна возможная производительность каждой из оплачиваемых девушек. Поэтому можно составить матрицу предполагаемых дневных доходов компании при назначении каждой из девушек i ($i = 1, \dots, 4$) на каждую из работ j ($j = 5, \dots, 8$) (см. табл. 12.15). Необходимо так распределить работы между девушками, чтобы максимизировать доходы компании за день.

Чтобы применить к решению этой задачи алгоритм Флада, нужно превратить матрицу C в матрицу A_0 с помощью следующего шага 1'.

Шаг 1'. Определяем $\max_{i,j} c_{ij}$ в матрице C . Образует матрицу A , используя соотношение

$$a_{ij}^{(0)} = (\max_{i,j} c_{ij}) - c_{ij}. \quad (19)$$

Матрица A_0 содержит по крайней мере один нулевой элемент.

Теперь задача решается аналогично задаче минимизации, в результате решения выбирается 4 независимых нуля. Читатель легко может проверить, что оптимальным является следующее назначение: 1—6, 2—8, 3—5 и 4—7. При этом предполагаемые доходы компании составят $\$ (8 + 5 + 3 + 6) = \$22,00$.

ДРУГИЕ ОСОБЕННОСТИ ЗАДАЧ О НАЗНАЧЕНИИ

В предыдущих примерах все элементы матрицы стоимости $a_{ij}^{(0)}$ были конечными числами. Однако это условие не является обязательным. Например, если по каким-либо причинам запрещается выполнение определенной работы на определенной машине, то это легко предусмотреть, для чего следует ввести в матрицу стоимостей произвольно большую (бесконечную) стоимость, т. е. $a_{ij}^{(0)} = \infty$. Тем самым нежелательное назначение будет автоматически исключено из оптимального решения. Такой прием используется при решении задачи коммивояжера, которая описывается в 16-й главе.

ЗАКЛЮЧЕНИЕ

В этой главе нами рассмотрен специальный тип задач линейного программирования — задача о назначении. Задача о назначении является математическим «двойником» задачи распределения (так называемой транспортной задачи), для решения которой существует много алгоритмов. Для решения задачи о назначении вручную особенно эффективен венгерский метод решения, разработанный Фладом.

ЛИТЕРАТУРА

1. Charnes A. and Cooper W. W., The Stepping Stone Method of Explaining Linear Programming, Mgmt. Sci., 1, no. 1, 49—69 (Oct. 1955).
2. Dulmage L. and Halperin I., On a Theorem of Frobenius—König and J. von Neumann's Game of Hide and Seek, Trans. Roy. Soc. Can., Third Series, Sec. III, 49, 23—29 (June 1955).
3. Dwyer P. S., The Solution of the Hitchcock Transportation Problem with a Method of Reduced Matrices, Statistical Laboratory, University of Michigan, Dec. 1955 (privately circulated).
4. Egerváry J., Matrixok Kombinatorius Tulajdonságairól, Matematikai és Fizikai Lapok, 38, 16—28 (1931). Translated by H. W. Kuhn as «Combinatorial Properties of Matrices», ONR Logistics Project, Princeton University, Princeton, 1953 (mimeographed).
5. Flood M. M., On the Hitchcock Distribution Problem. Pac. J. Math., 3, no. 2, 369—386 (June 1953).
6. Flood M. M., Operations Research and Logistics, Proceedings of First Ordnance Conference on Operations Research, Report No. 55-1, Office of Ordnance Research, Durham, pp. 3—25, Jan. 1955.
7. Flood M. M., The Traveling-Salesman Problem, J. Opns. Res. Soc. Am., 4, no. 1, 61—75 (Feb. 1958).
8. Flood M. M., The Traveling-Salesman Problem, in F. C. McCloskey and J. M. Copinger (eds.), Operations Research for Management, The Johns Hopkins Press, Baltimore, 1956.
9. Ford L. R., Jr., and Fulkerson D. R., A Simplex Algorithm for Finding Maximal Network Flows and an Application to the Hitchcock Problem, RAND Report RM-1604, RAND Corporation, Santa Monica, Dec. 20, 1955.
10. Frobenius G., Ueber Matrizen mit nicht negativen Elementen, Sitzungsberichte der Berliner Akad., 23, 456—477 (1912).
11. Gleyzal A., An Algorithm for Solving the Transportation Problem, J. Res. Nat. Bur. Stand., 54, no. 4, 213—216 (Apr. 1955).
12. Hitchcock F. L., The Distribution of a Product from Several Sources to Numerous Localities, J. Math. Phys., 20, 224—250 (1941).
13. Houthakker H. S., On the Numerical Solution of the Transportation Problem, J. Opns. Res. Soc. Am., 3, no. 2, 210—214 (May 1955).
14. König D., Theorie der endlichen und unendlichen Graphen, Chelsea Publishing Co., New York, 1950.
15. Koopmans T. C. and Reiter S., A Model of Transportation, in T. C. Koopmans (ed.), Activity Analysis of Production and Allocation, Cowles Commission Monograph No. 13, John Wiley & Sons, New York, 1951.
16. Kuhn H. W., The Hungarian Method for the Assignment Problem, Nav. Res. Log. Quart., 2, nos. 1 and 2, 83—98 (Mar.—June 1955).
17. Neumann J. von, A Certain Zero-Sum Two-Person Game Equivalent to the Optimal Assignment Problem, in H. W. Kuhn and A. W. Tucker (eds.), Contributions to the Theory of Games II, Annals of Mathematics Study No. 28, Princeton University Press, Princeton, pp. 5—12, 1953.
18. Votaw D. F. and Orden A., The Personnel Assignment Problem, in A. Orden and L. Goldstein (eds.), Symposium on Linear Inequalities and Programming, Project SCOOP, Headquarters, U. S. Air Force, Washington, pp. 155—163, 1952.

Глава 13

НЕКОТОРЫЕ ПРИМЕРЫ ЗАДАЧ РАСПРЕДЕЛЕНИЯ

В 11-й главе были рассмотрены два аналитических метода решения задач линейного программирования — транспортный и симплексный. *Транспортный метод* отличается большой простотой, и им можно широко пользоваться, особенно при наличии настольных счетных машин. Однако, как указывалось в 11-й главе, он применим только к ограниченному классу задач линейного программирования. *Симплексный метод*, напротив, является универсальным методом и применим ко всем задачам линейного программирования. Но хотя в этом методе используется простейший математический аппарат на уровне начальной школы, объем вычислений обычно так велик, что приходится прибегать к помощи быстродействующих электронных вычислительных машин или вводить различные упрощения с целью уменьшения размера задачи.

На многих предприятиях, особенно на тех, которые не имеют своих вычислительных машин, использование быстродействующих вычислительных машин оказывается практически невозможным, хотя всегда можно арендовать определенное машинное время в бюро обслуживания. Однако во многих случаях после получения основных данных надо так быстро принимать решения, что уже нет времени пользоваться услугами таких обслуживающих вычислительных центров. Так обстоит дело, например, при решении задач ежедневной загрузки машин и многих других задач, возникающих в повседневной практике производства.

В этой главе приводится неаналитический (так называемый *быстрый* или *грубый*) метод решения таких задач линейного программирования, дающий если не оптимальное, то приближенное к нему решение задачи. Этот метод рассматривается на примере действительной промышленной задачи и на гипотетическом примере. Чтобы не создалось мнения, что все задачи линейного программирования следует решать неаналитическими методами, приводится несколько примеров, в которых используются аналитические методы решения.

НЕАНАЛИТИЧЕСКОЕ РЕШЕНИЕ ЗАДАЧИ РАСПРЕДЕЛЕНИЯ

Задачу, связанную с производством антибиотиков, можно описать следующим образом. Производство антибиотиков состоит из различных этапов, которые заканчиваются полуавтоматической операцией расфасовки, в процессе которой происходит наполнение флаконов лекарствами, и затем операция запечатывания, на которую лекарства поступают по конвейеру из стерильной зоны наполнения в машинный центр. Машины-наполнители являются специализированным оборудованием, которое или приобретается, или берется в аренду. Эти машины отличаются по мощности, по скорости наполнения, по стоимости. При планировании работы участка наполнения бывает известна заранее месячная потребность в лекарствах. Плановик должен разработать производственный план, в котором учитывается не только, *сколько* должно быть расфасовано каждого лекарства и на *какой* машине, но также *когда* должны быть расфасованы эти лекарства. Другими словами, плановик должен дать ответы на следующие вопросы: (1) сколько? (количество различных доз и т. д.), (2) где? (распределение), (3) когда? (календарное планирование).

Главный вопрос «сколько?» учитывает определение экономического размера партии лекарств. Второй вопрос «где?» связан с распределением лекарств по машинам-наполнителям, и «когда?» связано с календарным планированием, или определением времени различных операций и времени поступления материалов на участок расфасовки. Когда составление плана для участка расфасовки закончено, плановик должен, исходя из процесса производства, перейти к планированию задач, т. е. распланировать начальные стадии процесса производства антибиотиков. В данном производстве нет смысла составлять заранее календарный план для машин, запечатывающих флаконы, поскольку крышки одинаковы для всех лекарств, все машины однотипны и работают с неполной загрузкой.

Где же можно использовать линейное программирование? Поскольку количество лекарств, которое нужно произвести, определено, линейное программирование можно использовать для ответа на второй вопрос, т. е. для правильного *распределения* продуктов на различные машины-наполнители. Таким образом, линейное программирование *распределяет*; оно не отвечает на 3-й вопрос, т. е. оно не *планирует* во времени. Другими словами, после того как найдено оптимальное распределение, остается задача календарного планирования.

Допустим, что определенная компания производит 12 различных видов антибиотиков, которые выпускаются или в виде порошков, или в виде масляных или водных смесей. Если принять различные модификации лекарств за разные продукты, то всего получим 16 продуктов. Кроме того, каждое лекарство может быть выпущено в 6 различных дозах, так что всего производится 53 различных комбинации лекарств.

На участке расфасовки работают 9 машин-наполнителей. Из них три машины называются жидкими наполнителями, т. е. они фасуют лекарства, представляющие собой масляные или водные смеси. Остальные шесть машин производят сухую расфасовку. Причем эти машины таковы, что ни одно из названных 16 лекарств не может быть расфасовано на машинах обоих типов. Поэтому возникают две более мелкие задачи.

Рассмотрим задачу распределения лекарств по сухим фасовочным машинам с целью минимизации общей стоимости операции расфасовки. Эти сухие лекарства составляют 10 из 16 и 31 комбинацию из исходных 53.

МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ

В данной задаче примем следующие обозначения:

- x_{ij} — количество j -го продукта, который фасуется на i -й машине;
- c_{ij} — стоимость расфасовки j -го продукта на i -й машине;
- t_{ij} — время, потребное для наполнения одного флакона j -го продукта на i -й машине;
- p_j — потребное количество j -го продукта за планируемый период;
- b_i — производственная мощность i -й машины;
- m — количество сухих фасовочных машин (в нашем случае $m = 6$);
- n — количество различных продуктов*, которые должны быть расфасованы за любой планируемый период, т. е. в нашем примере $n \leq 31$!

* В последующем изложении считается, что лекарства, выпущенные в различных дозах, являются различными продуктами. Определение последовательности расфасовки является отдельной задачей (календарного планирования), которую нельзя решить методами линейного программирования (распределения).

ИСПОЛЬЗУЕМЫЕ КОНСТАНТЫ

Затраты. Производственные затраты учитывают стоимость наладки, ремонта, содержания машины, уборки, настройки в процессе работы, зарплату рабочего, стоимость снабжения машины и отходов производства. Общая стоимость расфасовки всех продуктов выражается в виде

$$C = \sum_{i=1}^m \sum_{j=1}^n x_{ij} c_{ij} = \sum_{i,j} x_{ij} c_{ij}.$$

Другие величины, т. е. потребное время и количество произведенных продуктов, задаются следующими выражениями.

1. Суммарное время, потраченное на i -й машине для получения всех продуктов, $\sum_{j=1}^n t_{ij} x_{ij}$.

2. Общее количество j -го продукта, расфасованного на всех машинах, $\sum_{i=1}^m x_{ij}$.

Таким образом, данную задачу линейного программирования можно сформулировать следующим образом *: при заданных c_{ij} , t_{ij} , p_j , b_i ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$) определить $x_{ij} \geq 0$, которые минимизируют

$$C = \sum_{i,j} c_{ij} x_{ij}$$

при ограничениях

$$\begin{aligned} \sum_{j=1}^n t_{ij} x_{ij} &\leq b_i, & i = 1, 2, \dots, m, \\ \sum_{i=1}^m x_{ij} &\geq p_j, & j = 1, 2, \dots, n. \end{aligned}$$

Все зависимости, связанные с затратами и временем, принимаются линейными, т. е. предполагается, что стоимость расфасовки десяти флаконов в десять раз больше стоимости расфасовки одного флакона и, аналогично, на расфасовку десяти флаконов уходит в десять раз больше времени, чем на наполнение одного флакона. Такое допущение необходимо, чтобы можно было решать задачи с помощью методов линейного программирования **.

В тех случаях, когда перечисленные статьи затрат не зависят от машины, они исключаются из рассмотрения, поскольку они не влияют на общие затраты при различных вариантах распределения. Кроме того, для данной задачи колебания расходов по наладке различных машин для любого продукта так малы, что их также можно не учитывать. (Такое допущение устраняет нелинейность в задаче; однако в данной задаче даже большие затраты по наладке несутся, поскольку все лекарства обычно производятся определенными, фиксированными по размеру партиями, так что всегда можно рассчитать близкие к действительным затраты на наладку в пересчете на один флакон и на одну машину.) Объединяя оставшиеся статьи затрат, получим матрицу производственных затрат на 1000 флаконов (см. табл. 13.1 ***).

* Задачу можно сформулировать также в виде максимизации прибыли, оптимизации запасов и в других формах. Следует, однако, обратить внимание на то, что в задачах, не сходных с данной, минимизация издержек не всегда означает максимизацию прибыли.

** Методы линейного программирования можно применять и к некоторым специальным нелинейным задачам оптимизации. Например, см. Лемке и Чарнс [2].

*** Пустые места в таблицах 13.1 и 13.2 означают, что продукт не может быть расфасован на соответствующей машине.

Таблица 13.1

Производственные затраты на 1000 флаконов

Машина Продукт	1	2	3	4	5	6
1	1 417	2 747	2 373	1 564	3 252	2 060
2	1 425	2 450	2 236	1 509	3 084	2 141
3	1 368	2 402	2 188	1 452	2 994	2 061
4	1 368	2 900	2 686	1 452	2 994	2 061
5	1 355	2 470	2 201	1 460	2 978	2 048
6	2 703			2 806	3 138	2 192
7					6 118	4 227
8				10 040		
9	1 365	2 400	2 185	1 448	2 990	2 056
10					6 109	4 216
11	2 612			2 693	2 990	2 056
12					6 109	4 216
13	1 425	2 450	2 236	1 509	3 085	2 141
14	1 425	2 450	2 236	1 509	3 085	2 141
15					3 355	2 384

Производительность. Производительность машин определяется из непосредственных наблюдений и из заданных нормативов. Соответствующие величины приведены в таблице 13.2 *.

Производственные мощности машин. Общее время работы машины берется с учетом времени перерыва, времени наладки, с учетом возможных поломок и различных нарушений. Из этих соображений получают время

Таблица 13.2

Темп работы фармацевтического участка наполнения (минуты на флакон),
потребное количество каждого продукта и нормативные времена работы
каждой машины

Машина Продукт	1	2	3	4	5	6	Потребное количество продукта
1	0,0125	0,0200	0,0200	0,0125	0,0625	0,0417	55 500
2	0,0125	0,0200	0,0200	0,0125	0,0625	0,0417	22 799
3	0,0125	0,0200	0,0200	0,0125	0,0625	0,0417	35 933
4	0,0125	0,0250	0,0250	0,0125	0,0625	0,0417	53 097
5	0,0125	0,0200	0,0200	0,0125	0,0625	0,0417	514 793
6	0,0250			0,0250	0,0625	0,0417	43 987
7					0,1250	0,0833	77 697
8					0,2000		4 363
9	0,0125	0,0200	0,0200	0,0125	0,0625	0,0417	447 060
10					0,1250	0,0833	11 494
11	0,0250			0,0250	0,0250	0,0417	215 646
12					0,1250	0,0833	12 023
13	0,0125	0,0200	0,0200	0,0125	0,0625	0,0417	25 154
14	0,0125	0,0200	0,0200	0,0125	0,0625	0,0417	44 963
15					0,0625	0,0417	4 046
Время работы машины	7920	7920	7920	7920	7920	7920	

работы каждой машины. В нашем примере принято, что время работы каждой машины равно 7920 минутам в месяц.

Общее количество каждого продукта. Потребное количество каждого продукта определяется планирующими органами исходя из данных отдела сбыта, основанных на прогнозах относительно спроса и на существующих контрактах. Планирующие органы легко могут получить соответствующие данные и установить на их основе потребное количество каждого продукта на определенный месяц. Эти величины приведены в правом столбце таблицы 13.2.

Таким образом, для месячного плана (см. табл. 13.2) может быть сформулирована следующая задача линейного программирования. При заданных c_{ij} , t_{ij} , p_j , b_i ($i = 1, 2, \dots, 6$; $j = 1, 2, \dots, 15$) определить $x_{ij} \geq 0$, которые минимизируют

$$C = \sum_{i=1}^6 \sum_{j=1}^{15} c_{ij} x_{ij}$$

при ограничениях

$$\begin{aligned} \sum_{i=1}^6 x_{ij} &\geq p_j, & j = 1, 2, \dots, 15, \\ \sum_{j=1}^{15} t_{ij} x_{ij} &\leq b_i, & i = 1, 2, \dots, 6. \end{aligned}$$

Таким образом, нужно определить значения 90 переменных, которые минимизируют функцию стоимости, содержащую все 90 переменных, при условии, что эти переменные удовлетворяют 21 ограничивающему неравенству.

Эту задачу, как и любую задачу, представленную в такой форме, можно решить с помощью второго метода линейного программирования, т. е. с помощью симплексного метода. Однако здесь не будет использован этот метод, поскольку только для одного первого шага нужна матрица (таблица), состоящая из 127 столбцов и 23 строк *. Кроме того, воспользовавшись практическим правилом определения числа необходимых итераций, можно подсчитать, что потребуется от 22 до 34 таких таблиц для получения решения.

ПРОЦЕСС ВЫЧИСЛЕНИЯ

Теперь, когда мы знаем объем вычислений для данной задачи, может возникнуть вопрос: как же получить решение этой задачи на практике? И что можно сказать в общем случае о решении задач такого размера?

Прежде всего, что можно сказать о ручных вычислениях с помощью настольных вычислительных машин? Чтобы решить приведенную задачу, одному человеку потребуется 15—22 дня — один рабочий месяц; при этом возможны ошибки в вычислениях. Если не принимать во внимание возможности ошибок, то и тогда такой способ решения имеет сомнительную ценность, за исключением, может быть, задач долгосрочного планирования, в которых месячный срок не имеет значения.

Что можно сказать об автоматических вычислительных устройствах? Как было показано в 11-й главе, в симплексном методе необходимо выбирать наибольшее отрицательное число из данного набора чисел и затем выбирать наименьшее положительное число из соответствующего набора отношений.

* В действительности некоторые из этих столбцов не нужны, так как некоторые значения x_{ij} , очевидно, равны нулю (например, x_{71} , x_{81} , x_{86} и т. д.).

Поэтому, если использовать малые вычислительные машины, т. е. вычислительные машины вплоть до машин мощностью и быстродействием, соответствующими машине IBM CPC (электронная вычислительная машина с вводом программы на перфокартах), то потребуется много ручных вычислений, что значительно увеличит общее время решения, и поэтому такой способ вычислений окажется также неприемлемым, хотя и более быстрым, чем ручные вычисления. Если уж говорить о вычислительных машинах, то приемлемыми оказываются такие машины, как IBM 701, или IBM 704, или Рэмингтон-Рэнд UNIVAC. По сравнению с 15—22 днями работы на настольных вычислительных машинах эти машины решают соответствующую задачу меньше чем за 1 час. Конечно, к этому следует добавить время для подготовки данных для ввода в машину, обработки результатов и т. д., но в общем случае экономия времени будет значительной по сравнению со сроком в 15—22 дня.

Что же можно предпринять, если нет возможности использовать автоматические вычислительные машины или же их использование по каким-либо причинам представляется практически неосуществимым? По-видимому, придется прибегнуть к неаналитическому методу или, иначе, к методам проб и ошибок, которые могут давать удовлетворительные результаты, часто очень близкие к оптимальному решению. И действительно, в приведенной задаче такой метод решения позволил получить очень хорошие результаты.

Что же такое неаналитический метод, или метод проб и ошибок? Чтобы избежать большого количества цифр, в качестве иллюстрации рассмотрим пример, приведенный в статье, написанной М. Г. Мелденом [3]. В этом примере рассматривается задача минимизации машинного времени при производстве шести деталей на трех машинах. В таблице 13.3 приведены основные данные для этой задачи, а именно: потребное число различных деталей, время работы каждой машины и время обработки одной детали на каждой машине. Ниже описывается метод, основанный на эвристических соображениях.

Таблица 13.3

Условия задачи: потребное число деталей, времена обработки каждой детали (часы на одну деталь)

Деталь Машина	1	2	3	4	5	6	Резерв времени машины
1	3	3	2	5	2	1	80
2	4	1	1	2	2	1	30
3	2	2	5	1	1	2	160
Потребное число дета- лей	10	40	60	20	20	30	270

Шаг 1. Для каждой детали выберем машину, на которой ее можно обработать наилучшим образом (т. е. с наименьшей стоимостью, за мини-

мальное время или с максимальной прибылью в зависимости от функции цели). При этом получаем «идеальный» план *, т. е. такой, который не учитывает ограничений на ресурс времени каждой машины **.

Шаг 2. Будем изменять выбранный план с помощью метода проб и ошибок, принимая последующие наилучшие варианты до тех пор, пока не будут выполнены ограничения по ресурсу времени каждой машины или пока не будет получено приемлемое значение функции цели ***.

В приведенном примере (табл. 13.3) при применении первого шага алгоритма получим идеальную программу (см. табл. 13.4). Согласно этой программе машина 2 перегружена, поскольку требуется 100 часов работы, а эта машина может работать всего 30 часов.

Таблица 13.4

«Идеальный план»

Деталь \ Машина	1	2	3	4	5	6	Резерв времени машины	Пот- ребное число дета- лей
1						30	80	30
2		40	60				30	100
3	10			20	20		160	60
Потребное число дета- лей	10	40	60	20	20	30	270	190

В то же время у машин 1 и 3 остается еще 50 и 100 часов неиспользованного времени. Следовательно, необходимо уменьшить или вообще устранить перегрузку машины 2 за счет дополнительной нагрузки машин 1 и 3, учитывая минимально возможное уменьшение функции цели.

Очевидно, что лучше всего изменить нагрузку таким образом, чтобы время обработки детали увеличивалось только на 1 час. Так, деталь 2 следует переместить с машины 2 на машину 3, а деталь 3 — на машину 1. Это приведет к распределению, данному в таблице 13.5, где 25 деталей типа 3 обрабатываются на машине 1 и все 40 деталей типа 2 обрабатываются на машине 3. Теперь машина 1 полностью нагружена, машина 2 перегружена на 5 часов, а у машины 3 еще остается 20 часов неиспользованного времени. После незначительных изменений получаем таблицу 13.6, в которой нагружены все три машины и ни одна не перегружена. Следует отметить, что приведенное в таблице 13.6 решение является действительно оптимальным решением, что можно проверить с помощью симплексного метода.

* Следует подчеркнуть, что в действительности при этом получается «идеальное распределение». План же еще нужно закончить.

** Может встретиться случай, когда нельзя выполнить ограничения по времени работы машин без дополнительных сдвигов или добавочного времени. В этом случае целью может быть минимизация суммарного дополнительного времени работы или равномерное распределение его на определенные машины.

*** Если все ограничения на время работы уже выполнены при этой идеальной программе, то очевидно, что здесь нет никакой задачи.

Таблица 13.5

Промежуточный план

Деталь Машина	1	2	3	4	5	6	Резерв времени машины	Пот- ребное число деталей
1			25			30	80	80
2			35				30	35
3	10	40		20	20		160	140
Потребное число дета- лей	10	40	60	20	20	30	270	255

Таблица 13.6

Оптимальный план

Деталь Машина	1	2	3	4	5	6	Резерв времени машины	Пот- ребное число деталей
1			30			20	80	80
2			30				30	30
3	10	40		20	20	10	160	160
Потребное число дета- лей	10	40	60	20	20	30	270	270

Хотя, очевидно, сравнение времени решения для такого маленького примера непоказательно, следует все же указать, что при решении с помощью симплексного метода необходима матрица 11×34 и требуется 11 шагов, или итераций (т. е. будет 11 таких матриц), так что опытному вычислителю потребуется около 4 часов работы. Это время можно сравнить со временем в несколько минут, которое потребовалось для решения с помощью неаналитического метода (проб и ошибок).

Что касается общей задачи производства антибиотиков, то с помощью неаналитического метода можно получить отличные результаты самое большее за несколько часов, и хотя в общем случае этот метод не дает оптимального решения, тем не менее с его помощью получены значительно лучшие результаты, чем с помощью методов распределения, применявшихся компанией. Другими словами, если этот метод и не является основным методом

решения распределительных задач, поскольку не используются широко электронные вычислительные машины, все же он представляет собой шаг вперед в решении подобных задач, особенно при его дальнейшем усовершенствовании *.

ОПТИМАЛЬНОЕ ИСПОЛЬЗОВАНИЕ ОБОРУДОВАНИЯ ПРИ ПРОИЗВОДСТВЕ ПРОКАТА

Другим примером задачи распределения является задача оптимального использования оборудования при производстве проката. Хотя при решении этой задачи также используются неаналитические методы, но она приводится не в качестве еще одного примера применения неаналитических методов, а чтобы показать специальные приемы, которые применяются для приведения задачи к виду, пригодному для решения.

Рассматривается производство изделий холодной прокатки. Эти изделия можно получить из заготовок, называемых слябами или болванками, которые обрабатываются на станах горячей и холодной прокатки. Кроме того, покупается определенное количество изделий горячей прокатки в виде полосового или листового проката, который затем проходит обработку на станах холодной прокатки. На рис. 13.1 в упрощенной форме показан производственный процесс, который включает операции нагрева, прокатки, резки и отжига. Кроме того, многие изделия должны проходить через некоторые операции несколько раз (например, операции отжига и прокатки обычно повторяются для мелкого проката, так как нельзя получить нужный размер за один раз, поскольку в этом случае характеристики металла не соответствуют заданным). И наконец, каждую операцию можно выполнять на нескольких машинах, поэтому есть возможность выбрать для обработки одну из этих машин.

Таким образом, чтобы определить оптимальный вариант использования оборудования на участке прокатки, необходимо выбрать тип и размер заготовки (т. е. болванки, сляба, листов или полос), а также способ обработки (т. е. последовательность операций и номер машины, на которой производится обработка).

ДАННЫЕ ДЛЯ РЕШЕНИЯ ЗАДАЧИ

Очевидно, что приведенная задача является распределительной задачей, к которой можно применить аппарат линейного программирования. Однако, прежде чем использовать методы линейного программирования, необходимо получить следующие данные: (1) нормативные затраты на производство продуктов, (2) затраты времени на наладку, (3) производительность оборудования, (4) мощности машин и (5) потребное количество каждого продукта. Кроме того, необходимо представить эти данные в удобной форме.

Это станет очевидным, если рассмотреть нормативные затраты на производство продуктов. Эти затраты обычно фиксируются в нескольких карточках на каждую производственную операцию, входящую в процесс обработки данного продукта. Поэтому эти данные нельзя использовать непосредственно **. Например, для определения нормативных затрат на производ-

* Подробнее об использовании этого метода см. [4]. В приведенной задаче, затраты на получение «оптимального» решения превысят дополнительные прибыли, которые можно получить, используя это решение; в этом смысле «быстрое и грубое» решение является оптимальным.

** Следует указать, что при вычислении стоимости каждой операции учитываются потери за счет отходов, накладные расходы и т. д.

ство каждого продукта потребовалось от 60 до 75 минут при условии, что известна последовательность операций и основные характеристики продукта (такие, как отжиг, отделка, толщина, ширина и количество). Таким

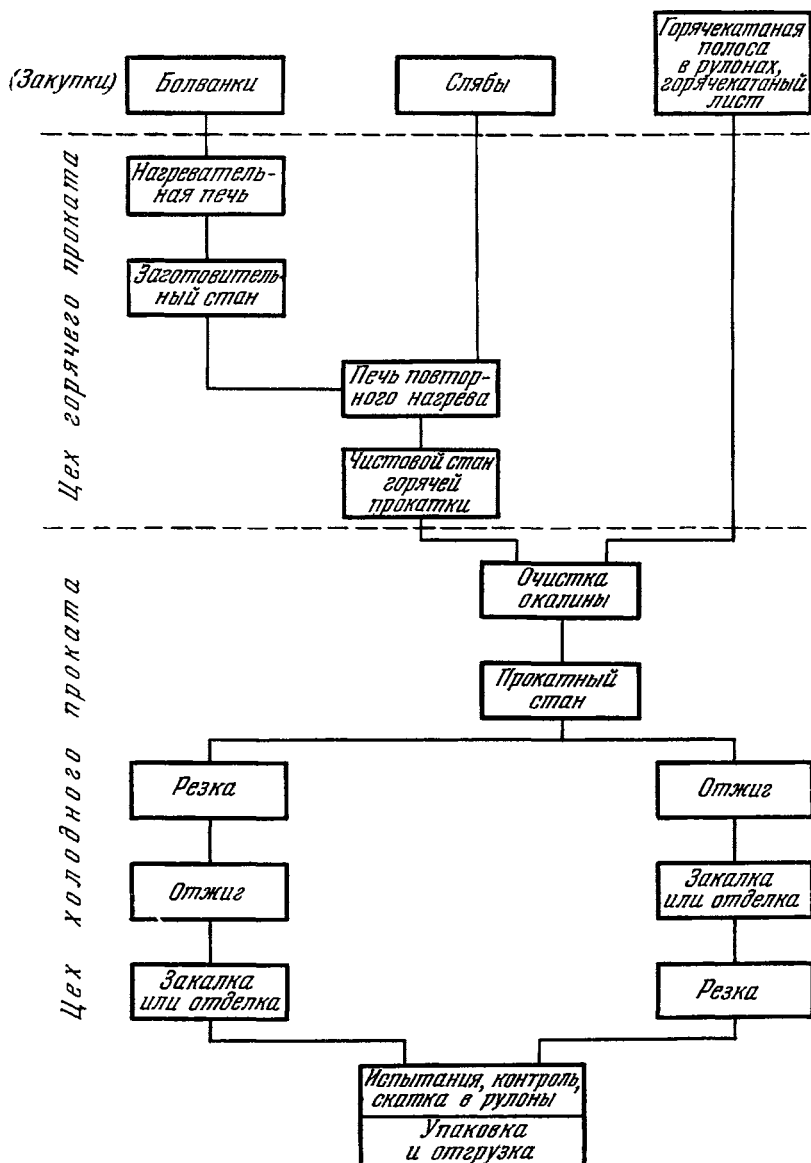


Рис. 13.1. Блок-схема производства холодного проката.

образом, если бы каждую неделю нужно было пропускать 100 заказов при среднем количестве способов обработки на каждый заказ, равном четырем *, то потребовалось бы приблизительно 400—500 человеко-часов в неделю, чтобы определить только нормативные затраты на производство продуктов.

* Хотя может быть много различных вариантов обработки, но на практике достаточно рассмотреть от 2 до 8 вариантов в зависимости от продукта.

Очевидно, что такие затраты времени и сил совершенно исключают практическое использование методов линейного программирования для решения недельной задачи распределения.

Чтобы преодолеть эти затруднения, необходимо разработать способ представления данных по затратам в удобной и доступной форме. Такой формой является графическая, при которой задаются нормативные затраты на продукты (с учетом ширины, толщины и объема заказа) и, кроме того, даются сравнимые затраты при другом способе производства (т. е. при других исходных материалах, другой последовательности операций и другом оборудовании). Ниже рассматривается построение таких графиков.

Построение графиков затрат. Построение графиков затрат проиллюстрируем на примере одного из участков, а именно участка черновой прокатки. Прокат заданной конечной ширины получается путем прокатки большей ширины, кратной заданной, и последующей резки по размеру. Поскольку валки имеют различные характеристики в отношении ширины проката, то выбор кратного числа определяет выбор валков для обработки. Кроме того, осложнения возникают в связи с возможностью выбора одного из видов оборудования для каждой ширины (из-за того, что валки могут обрабатывать несколько размеров). Таким образом, для заданной конечной ширины проката необходимо определить тип и ширину исходных материалов, последовательность операций и выбрать оборудование. Однако, если теоретически существует большое число вариантов обработки, на практике существует всего несколько (приблизительно от двух до восьми).

Можно показать, что при обработке определенного сортамента на заданном стане удельные производственные затраты (без учета стоимости наладки) изменяются обратно пропорционально ширине. Например, если на прокатном стане обрабатывается прокат шириной от 4 до 24 дюймов (т. е. ширина в процессе обработки), то минимальные удельные затраты (т. е. затраты на 1 фунт проката) будут соответствовать ширине 24 дюйма. Затем по мере уменьшения ширины удельные затраты будут увеличиваться по закону обратной пропорциональности до размера 12 дюймов.

Ясно, что на стане, который может прокатывать полосу шириной 24 дюйма, неэкономично прокатывать 12-дюймовый материал. Гораздо выгоднее прокатывать 24-дюймовый материал (т. е. кратное 2 число) и затем разрезать его пополам *. Таким образом, если не учитывать дополнительную стоимость на резку, то можно прокатывать 12-дюймовый материал с теми же удельными затратами, что и 24-дюймовый. Аналогичные рассуждения применимы к прокату шириной 8 дюймов, 6 дюймов и т. д., т. е. при кратных числах 3, 4 и т. д. На рис. 13.2 приведена кривая удельных затрат (включая стоимости наладки), полученная из расчета обработки мерной (кратной) ширины проката. На рис. 13.2 по осям отложены удельные затраты и конечная ширина **.

Кривые затрат, подобные приведенным на рис. 13.2, были рассчитаны для «стандартной» продукции ***, выпускаемой на соответствующем оборудовании.

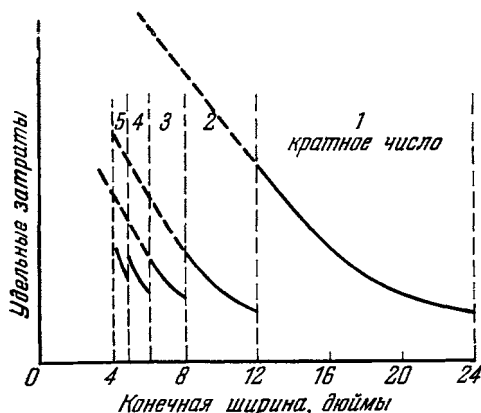
* Для простоты полагаем, что допуск на отходы при резке и зачистке торца равен нулю.

** Из рис. 13.2 можно заметить, что конечную ширину, например в 7 дюймов, можно получить тремя способами на одном стане, т. е. получить прокат шириной 7, 14 или 21 дюймов, и затем разрезать, если потребуется. Сплошные кривые на рис. 13.2 соответствуют наиболее дешевому способу, т. е. прокату 21 дюйм, а пунктирные линии соответствуют более дорогому способу обработки и получены путем экстраполяции кривых для больших размеров.

*** «Стандартная» продукция имеет определенные характеристики, отделку, торцы и калибр. Кроме того, берется также стандартная величина заказа.

довании на различных участках производства. Были построены аналогичные кривые для различных сортментов, и, кроме того, была рассчитана вспомогательная таблица коэффициентов для учета колебаний величины заказа. Затем эти стоимостные кривые были сопоставлены по каждой операции, что позволило наглядно сравнивать удельные затраты на любом участке производства.

Были также получены кривые общих затрат для наиболее часто встречающихся последовательностей операций, т. е. были получены кривые



общих затрат для всего процесса. Эти кривые были построены таким образом, чтобы иметь возможность наглядно сравнивать кривые удельных затрат для различных способов обработки, дающих возможность получить прокат заданной конечной ширины.

Указанные кривые затрат были получены только в итоге многократных сравнений с известными нормативными производственными затратами. После того как они были построены, было произведено сравнение затрат для большого числа производимых продуктов, причем затраты, полученные из этих кривых, лишь незначительно отлича-

Рис. 13.2. Единичные стоимости в функции конечной ширины.

ются от принятых нормативных производственных затрат. Однако если для определения производственных затрат прежними методами требуется 60—70 минут, то кривые затрат позволяют определить соответствующие величины меньше чем за 5 минут.

Замечания. После определения стоимостных данных можно приступить к решению задачи распределения с помощью специально разработанного неаналитического метода.

В заключение нужно сделать два важных замечания. Во-первых, в первоначальной постановке задача сводилась к определению оптимального использования оборудования только на участке горячей прокатки. Однако в результате анализа задачи стало очевидно, что ее нельзя решить без изучения участка холодной прокатки. Рассмотрение оборудования только для участка горячей прокатки не дает оптимального решения задачи. Во-вторых, основные трудности заключаются не столько в решении с помощью матрицы удельных затрат, сколько в определении самой матрицы. При разработанных же способах практического определения удельных затрат сама задача распределения решается достаточно быстро.

ЗАДАЧА ОПТИМАЛЬНОЙ ЗАКУПКИ УГЛЯ

В этом разделе рассматривается еще одна задача распределения, связанная с закупкой угля на шахтах для энергетических станций одной из электрических компаний. Эта задача была быстро решена с помощью симплексного метода после некоторых преобразований.

Известно количество потребной энергии*, которую должны обеспечивать электрические станции данной компании. Эта энергия вырабатывается на нескольких станциях компании с использованием топлива, закупаемого на различных шахтах битуминозного угля. Уголь бывает различного качества и различной цены. Одной из основных характеристик угля, по которой одно месторождение отличается от другого, является содержание Б. Т. Е. (британских тепловых единиц) в одном фунте. Зная это число, можно определить, сколько угля данного сорта необходимо закупить, отгрузить и переработать на станции.

Стоимость тонны угля, доставленной на каждую электростанцию, легко определить. Для этого необходимо учесть стоимость добычи угля на шахте, транспортные издержки и стоимость переработки угля на станции.

Хотя легко определить количество потребной энергии (в киловатт-часах) для каждой станции, но сложнее перевести эти величины в тонны угля из-за различного содержания Б. Т. Е. в различных сортах угля. Был предложен простой, но эффективный выход из этих затруднений: (1) задать количество потребной энергии для каждой станции не в киловатт-часах, а в Б. Т. Е. и (2) выразить стоимость угля как функцию содержания Б. Т. Е., а не тоннажа**.

После таких преобразований задача оптимальной закупки угля легко решается с помощью симплексного метода***. Полученная в результате решения оптимальная таблица была проанализирована, чтобы учесть влияние колебаний различных параметров и тем самым оценить существующий план закупок.

ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНОГО СОСТАВА ШИХТЫ ДЛЯ ДОМЕННОЙ ПЕЧИ

Методы линейного программирования оказались очень эффективны при решении задачи, связанной с определением наилучшей (оптимальной) шихты для производства доменного чугуна. Это в основном задача «диеты», в которой требуется получить конечный продукт (т. е. доменный чугун) определенного качества (т. е. с заданным содержанием марганца, серы, фосфора и т. д.), используя комбинации различных исходных составляющих.

В обычной задаче «диеты» ищется просто смесь различных исходных материалов, которая обеспечивает получение конечного продукта заданного качества. Это относится к задаче составления корма крупного рогатого скота, к задаче смешивания руд для продажи и т. д. Однако в задаче доменного производства существует важная особенность, связанная с превращением (а не смешиванием) исходных материалов в доменный чугун и шлак. Как и в описанной выше задаче о закупке угля, различные руды имеют различные характеристики. Кроме того, для разных руд существуют различные требования на количество кокса и известняка. Поэтому отыскивается общая единица для всех материалов. В данной задаче такой единицей является «количество руды, потребной на тонну горячего металла».

* Уровень потребной энергии определяется с помощью прогнозирования, а также с учетом погоды на планируемый период. Чтобы обезопасить себя от возможных ошибок, используются (1) «аккумуляторы», т. е. машины, которые немедленно могут поглотить любой излишек энергии, и (2) взаимодействие с соседними предприятиями.

** То есть Б.Т.Е. используется как общая единица для измерения тоннажа и киловатт-часов.

*** В задачу входят ограничения на количество потребной энергии на каждой станции, а также максимальный и минимальный объем закупок на каждой шахте.

Чтобы определить «количество руды на тонну горячего металла», нужно знать количество железа и марганца в руде. Например, известно, что железо вместе с марганцем составляет 95% веса горячего металла, т. е. 1900 фунтов на тонну, причем только 75% марганца от его исходного содержания в руде остается в горячем металле. Тогда можно определить теоретическое количество руды, необходимое для получения тонны горячего металла. Если берется руда с содержанием железа 49,78% и марганца 0,77%, то теоретическое количество руды, потребное для получения тонны горячего металла, определяется следующим образом:

$$\frac{1900}{0,4978 + (0,75)(0,0077)} = 3773 \text{ фунта.}$$

К этой величине следует прибавить потери на пыль, уносимую с колошниковым газом. Если принять, что они составляют 25% от материала в 20 меш * в теоретически потребной руде, то получим общее количество руды **:

$$\frac{3773}{1,00 - (0,25)(0,25)} = 4025 \text{ фунтов.}$$

Кроме того, необходимо ввести коэффициент полезного действия. Считая, что соответствующие потери составляют 2%, получаем $(4025 + 80)$, или 4105 фунтов руды, потребной для получения одной тонны горячего металла.

Для каждой руды необходимо также определить соответствующее количество кокса и известняка, необходимых для доменного процесса. Количество кокса (т. е. содержание кокса) определяется из расчетов, основанных на формулах Флинта [1]. В эти расчеты входит определение эффективного углерода в коксе как функции углерода, золы и серы в коксе, влаги в воздухе и т. д.

Для руды, описанной в вышеприведенном примере, содержание кокса составляет 1740 фунтов на тонну горячего металла. Это содержание кокса затем корректируется с учетом (1) изменения объема шлака в руде и (2) более качественной руды (т. е. руды мельче, чем 20 меш). Эта коррекция производится также по формулам Флинта. И наконец, иногда производится добавочная коррекция на содержание кокса для специальных материалов.

Количество известняка (флюса) определяется путем непосредственных расчетов.

Стоимость производства тонны горячего металла определяется как сумма стоимостей руды, теоретически потребной для производства кокса и известняка. В нашем примере руда стоит \$11,93 за тонну (брутто), или всего \$21,86 (т. е. 4105 фунтов, или 1,8326 тонны брутто) для получения тонны (нетто) горячего металла. Если прибавить стоимость кокса и известняка, то общая стоимость получения одной тонны (нетто) горячего металла становится \$36,65. Аналогично определяются стоимости для других руд.

Определив эти стоимости, а также зная металлургические свойства каждой из руд, можно легко получить оптимальную загрузку (т. е. состав смеси руд) доменной печи, чтобы получить доменный чугун заданного качества и минимальной стоимости.

* Меш — единица измерения размера ячейки сита, сквозь которое просеивают руду при сортировке. Размеры стандартизованы и в международной классификации расположены в убывающем порядке. Термин «руда в —20 меш» означает, что гранулы руды прошли сквозь сито в 20 меш, но задержались ситом следующего размера. (Прим. ред.)

** В нашем примере принято, что теоретически потребная руда составляет 25% от руды в —20 меш.

Эта задача была решена симплексным методом на электронной вычислительной машине IBM 701. Кроме того, были решены связанные с ней задачи, в которых оценивалось влияние различных факторов: (1) изменения требований на характеристики доменного чугуна и (2) сравнительных качеств различных руд. Решение этих задач позволяет оценить (1) важность ограничений (например, не заданы ли «слишком высокие» требования на качество доменного чугуна?) и (2) сравнить условия различных конкурирующих поставщиков руды.

На этом примере видно, что, так же как и во многих других задачах линейного программирования, основные трудности заключаются в формулировке задачи и в определении основных данных, в то время как формальное решение получить достаточно легко.

Л И Т Е Р А Т У Р А

1. Flint R. V., Multiple Correlation of Blast Furnace Variables, Blast Furnace, Coke Oven and Raw Materials, Proceedings of American Institute of Mining and Metallurgical Engineers Conference, Pittsburgh, Mar. 31—Apr. 2 1952, pp. 49—73.
2. Lemke C. E. and Charnes A., Extremal Problems in Linear Inequalities, Technical Report No. 36, Carnegie Institute of Technology, Pittsburgh, 1953.
3. Melden Morley G., Operations Research, Fact. Mgmt., 111, no. 10, 113—120 (Oct. 1953).
4. New Machine Loading Methods, Fact. Mgmt. 112, no. 1, 136—137 (Jan. 1954).

Задачи массового обслуживания возникают в том случае, когда некоторые объекты, требующие обслуживания, или само обслуживающее оборудование могут оказаться бездействующими. В зависимости от их структуры задачи массового обслуживания можно разделить на две группы. К первому классу относятся задачи, в которых случайными оказываются поток требований, или длительность времени обслуживания, или и то и другое вместе. При решении этих задач приходится определять либо оптимальное число обслуживающих приборов, либо оптимальную скорость потока (или находить моменты поступления заявок), либо одновременно и то и другое.

Класс моделей, пригодных для решения этих задач, связанных с расчетом оборудования и составлением графиков, и рассматривается в теории массового обслуживания или в теории очередей.

Начало этой теории было положено статьей Эрланга, опубликованной в 1909 г. В первоначальный период своего развития (примерно до 1945 г.) эта теория в качестве приложений в основном рассматривала работу телефонных сетей, однако в последнее время ее удалось привлечь к изучению значительно более широкого круга явлений.

Несмотря на то, что при создании моделей массового обслуживания приходится использовать весьма сложные разделы математики, решение многих задач может быть получено довольно просто, например с помощью метода Монте-Карло. Введение в теорию массового обслуживания, а также упрощенный пример применения метода Монте-Карло и рассматриваются в 14-й главе.

В главе 15 излагается одна «классическая» задача теории массового обслуживания. Эта задача впервые появилась в журнале «*Journal of the Operations Research Society of America*» и была удостоена Ланкастеровской премии университета Джона Гопкинса, как лучшая статья, опубликованная в области исследования операций в 1954 году.

Центральный вопрос второй группы задач теории массового обслуживания состоит в определении последовательности, в которой ряд станций обслуживания выполняет имеющиеся заказы.

Вопросы, связанные с объемом оборудования и моментами поступления заказов, в этих задачах представляют уже меньший интерес. Задачи этой группы получили название задач последовательного обслуживания, и при решении их используются методы комбинаторного анализа. Рассмотрению некоторых несложных задач этого типа посвящена глава 16; здесь также обсуждаются модели последовательного обслуживания и некоторые относящиеся к ним вопросы.

Глава 14

МОДЕЛИ ОЧЕРЕДЕЙ

ВВЕДЕНИЕ

При фиксированных параметрах потока требований и заданном объеме обслуживающего оборудования длина очереди является функцией времени. Поэтому процесс образования очереди иногда называют «стохастическим процессом». Как известно, процесс называется стохастическим, если он состоит из случайных переменных, значения которых меняются во времени.

При построении функции распределения вероятностей для длины очереди необходимо учесть ряд особенностей, присущих процессу образования очереди:

1. Порядок, в соответствии с которым заказы (потребители у прилавка, автомашины на погрузочных пунктах и т. д.) поступают и занимают свое место в очереди.

2. Количество обслуживающих единиц (они называются станциями), исполняющих заказы, и стратегию обслуживания, т. е. ограничения, наложенные на возможности и потребности обслуживания.

3. Последовательность обслуживания — так называемый порядок очереди.

4. Характер обслуживания и его длительность. Это — выход системы (например, обслуженные потребители, загруженные автомашины и т. д.).

При решении задач, связанных с очередями, возможны две ситуации: (1) число заказов слишком велико; в этом случае говорят о большом времени ожидания или о недостаточном объеме обслуживающего оборудования; (2) на станции обслуживания поступает недостаточное число заказов; в этом случае говорят, что слишком велико время простоя оборудования или имеется избыток оборудования. Естественно, что хотелось бы иметь оптимальное соотношение между потерями, вызванными простоем оборудования, и потерями из-за ожидания. Рассмотрим пример, иллюстрирующий это соотношение.

Пусть порядок поступления заказов на единственный пункт обслуживания и очередность их выполнения характеризуются следующими параметрами.

Порядок поступления заказов.

1. В среднем поступает шесть заказов за час, или по одному в течение каждых 10 минут.

2. В пределах 10-минутного интервала момент поступления каждого отдельного заказа случаен.

3. Каждый заказ занимает свое место в очереди в порядке появления и поступает на обслуживание в тот момент, когда пункт освобождается.

Порядок выполнения заказов.

1. В среднем за час выполняется 10 заказов, или по одному в течение каждых 6 минут.

2. На выполнение каждого заказа отводится точно 6 минут; заказы обслуживаются в порядке их поступления.

Допустим, что поступление заказов началось в 8 час 7 мин и продолжалось в следующем порядке: 8 час 25 мин, 8 час 39 мин, 8 час 43 мин и 8 час 56 мин (см. первый столбец табл. 14.1). Далее в этой таблице приведены следующие данные: моменты начала и конца обслуживания; интервалы времени простоя пункта; время ожидания обслуживания; максимальная

длительность очереди в течение каждого десятиминутного интервала в пределах часа.

Т а б л и ц а 14.1

Момент поступления заказа	Момент начала обслуживания	Момент окончания обслуживания	Время простоя пункта	Время ожидания обслуживания	Длина очереди
8,07	8,07	8,13	0	0	0
8,14	8,14	8,20	1	0	0
8,25	8,25	8,31	5	0	0
8,39	8,39	8,45	8	0	0
8,43	8,45	8,51	0	2	1
8,56	8,56	...	5	0	0

Из этой таблицы видно, что в течение часа очередь образовалась только один раз. Заказ поступил в 8 час 43 мин, а его обслуживание началось в 8 час 45 мин. Полное время ожидания составило 2 минуты, а полное время простоя — 19 минут. Можно ожидать, что с увеличением времени обслуживания длина очереди и время ожидания возрастут. В таблице 14.2 приведены данные для той же системы заказов, но для случая, когда время обслуживания составляет 9 минут.

Т а б л и ц а 14.2

Момент поступления заказа	Момент начала обслуживания	Момент окончания обслуживания	Время простоя пункта	Время ожидания обслуживания	Длина очереди
8,07	8,07	8,16	0	0	0
8,14	8,16	8,25	0	2	1
8,25	8,25	8,34	0	0	0
8,39	8,39	8,48	5	0	0
8,43	8,48	8,57	0	5	1
8,56	8,57	...	0	1	1

Из таблицы 14.2 видно, что полное время простоя уменьшилось с 19 до 5 минут, между тем как полное время ожидания обслуживания увеличилось с 2 до 8 минут. В рассмотренном простом примере промежутки времени слишком мал для создания значительных очередей. Однако уже здесь проявились некоторые черты, присущие процессу образования очередей и приводящие к определенному типу задачам. Например, видна необходимость в оперативном управлении средствами обслуживания, позволяющем обеспечить оптимальное соотношение между издержками из-за простоя и издержками, вызванными ожиданием обслуживания.

С ростом капиталовложений в систему обслуживания можно сократить длительность времени ожидания и убытки, обусловленные очередями. При этом желательно добиться минимального размера затрат как производственных, основных и оборотных, так и затрат, связанных с ожиданием обслуживания. Оптимальное соотношение между отдельными группами издержек может быть получено в результате планирования потока заказов, либо путем использования соответствующего объема оборудования, или с помощью

одновременного применения обоих этих методов. В частности, при фиксированном объеме оборудования поток заказов должен быть отрегулирован таким образом, чтобы обеспечить минимальные суммарные издержки, вызванные простоями и ожиданием обслуживания. Если же поток заказов неконтролируем, то минимум всех производственных затрат может быть достигнут путем подбора соответствующего объема оборудования, либо изменением количества обслуживающего персонала, или с помощью регулирования обоих этих элементов. Если же контролю поддаются и моменты поступления заказов и объем средств обслуживания, то пытаются одновременно использовать обе эти возможности.

ПРОБЛЕМА ОБРАЗОВАНИЯ ОЧЕРЕДИ НА ЕДИНСТВЕННОЙ СТАНЦИИ ОБСЛУЖИВАНИЯ

Найдем среднюю длину очереди и вероятность появления очереди заданной длины на единственной станции обслуживания при случайных потоках заказов на входе и выходе. Предположим, что скорость обслуживания не зависит от числа заказов в очереди и что заказы обслуживаются в порядке их поступления.

Введем следующие обозначения:

n — число заказов в очереди к моменту времени t ;

$P_n(t)$ — вероятность образования очереди из n заказов к моменту времени t ;

$\lambda \Delta t$ — вероятность появления в очереди нового заказа в промежутке времени от t до $t + \Delta t$; здесь λ — средняя скорость появления заказов;

$\mu \Delta t$ — вероятность того, что в интервале времени от t до $t + \Delta t$ завершится исполнение заказа, находящегося на обслуживании; здесь μ — средняя скорость обслуживания;

$\bar{n}$ — средняя длина очереди (т. е. число заказов в ней).

Используя основные положения теории вероятности, составим систему дифференциальных уравнений для $P_n(t)$ (и далее для $\bar{n}$).

Вероятность того, что к моменту времени $(t + \Delta t)$ в очереди будет находиться n заказов ($n > 0$), может быть записана в виде следующей суммы вероятностей четырех сложных независимых событий:

1. Произведения вероятностей того, что
 - а) к моменту времени t в очереди имеется n заказов $P_n(t)$,
 - б) в течение интервала времени Δt нет поступления новых заказов $1 - \lambda(\Delta t)$,
 - в) в течение интервала Δt ни на одном из заказов не завершается обслуживание $1 - \mu(\Delta t)$.
2. Произведения вероятностей того, что
 - а) к моменту времени t в очереди имеется $n + 1$ заказов $P_{n+1}(t)$,
 - б) в течение интервала Δt завершается обслуживание одного заказа $\mu(\Delta t)$,
 - в) в течение интервала времени Δt нет поступления новых заказов $1 - \lambda(\Delta t)$.
3. Произведения вероятностей того, что
 - а) к моменту времени t в очереди имеется $n - 1$ заказов $P_{n-1}(t)$,
 - б) в течение интервала Δt поступает один новый заказ $\lambda(\Delta t)$,
 - в) в течение интервала Δt не завершается обслуживание ни одного из заказов $1 - \mu(\Delta t)$.
4. Произведения вероятностей того, что
 - а) к моменту времени t в очереди имеется n заказов $P_n(t)$,

- б) в течение интервала Δt поступает один новый заказ $\lambda(\Delta t)$,
 в) в течение интервала Δt завершается обслуживание
 одного заказа $\mu(\Delta t)$.

Вероятности поступления или обслуживания более одного заказа за промежуток времени Δt предполагаются пренебрежимо малыми.

Преобразуем выражения для этих четырех вероятностей следующим образом:

1. $P_n(t)(1 - \lambda \Delta t)(1 - \mu \Delta t) = P_n(t)[1 - \lambda \Delta t - \mu \Delta t] + o_1(\Delta t)^*$.
2. $P_{n+1}(t)(\mu \Delta t)(1 - \lambda \Delta t) = P_{n+1}(t)\mu \Delta t + o_2(\Delta t)$.
3. $P_{n-1}(t)(\lambda \Delta t)(1 - \mu \Delta t) = P_{n-1}(t)\lambda \Delta t + o_3(\Delta t)$.
4. $P_n(t)(\lambda \Delta t)(\mu \Delta t) = o_4(\Delta t)$.

Складывая эти формулы, находим выражение для вероятности появления очереди из n заказов к моменту времени $(t + \Delta t)$:

$$P_n(t + \Delta t) = P_n(t)[1 - \lambda \Delta t - \mu \Delta t] + P_{n+1}(t)\mu \Delta t + \\ + P_{n-1}(t)\lambda \Delta t + o_1(\Delta t) + o_2(\Delta t) + o_3(\Delta t) + o_4(\Delta t). \quad (1)$$

Это уравнение может быть переписано в следующем виде:

$$\frac{P_n(t + \Delta t) - P_n(t)}{\Delta t} = \lambda P_{n-1}(t) + \mu P_{n+1}(t) - (\lambda + \mu) P_n(t) + \\ + o_1(\Delta t) + o_2(\Delta t) + o_3(\Delta t) + o_4(\Delta t).$$

Полагая $\Delta t \rightarrow 0$, получаем дифференциальное уравнение

$$\frac{dP_n(t)}{dt} = \lambda P_{n-1}(t) + \mu P_{n+1}(t) - (\lambda + \mu) P_n(t) \quad (n > 0). \quad (2)$$

Предположим, что к моменту времени t в длину очереди включается также заказ, находящийся на обслуживании, и что в любой ситуации $n \geq 0$. В этих условиях ранее записанные выражения для вероятностей должны быть видоизменены с тем, чтобы учесть случай $n = 0$. Вероятность того, что к моменту $t + \Delta t$ число заказов в очереди будет равно 0, выражается суммой вероятностей двух независимых событий:

1. Вероятности $P_0(t)(1 - \lambda \Delta t)$ того, что к моменту t очереди нет и в течение интервала Δt новые заказы не поступают.

2. Вероятности $P_1(t)(\mu \Delta t)(1 - \lambda \Delta t)$ того, что к моменту t в очереди имеется один заказ и в течение интервала Δt завершается его обслуживание; в течение того же интервала Δt новые заказы не поступают.

Складывая вероятность этих двух независимых событий, находим вероятность появления очереди нулевой длины к моменту $t + \Delta t$ в виде

$$P_0(t + \Delta t) = P_0(t)(1 - \lambda \Delta t) + P_1(t)\mu \Delta t - \lambda \mu P_1(t)(\Delta t)^2. \quad (3)$$

Из выражения (3) следует, что

$$\frac{P_0(t + \Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t) + \mu P_1(t) - \lambda \mu P_1(t)(\Delta t)$$

и

$$\frac{dP_0}{dt} = -\lambda P_0(t) + \mu P_1(t). \quad (4)$$

* Члены $o_i(\Delta t)$ являются величинами более высокого порядка по сравнению с Δt , и потому ими можно пренебречь.

Дифференциальные уравнения (2) и (4) отражают в неявном виде связь между временем ожидания и временем обслуживания и являются исходным пунктом для решения многих задач теории очередей. В силу сложности $P_n(t)$ решение этих уравнений связано с трудностями. Однако в частном случае, когда можно предположить, что $P_n(t)$ не зависит от времени и равно P_n , задача решается легко. В силу того, что вероятность P_n во времени не меняется,

$$\frac{dP_n}{dt} = 0, \quad n = 0, 1, 2, \dots,$$

и потому для уравнений (2) и (4) имеем

$$0 = \lambda P_{n-1} + \mu P_{n+1} - (\lambda + \mu) P_n \quad (n > 0), \quad (2a)$$

$$0 = -\lambda P_0 + \mu P_1 \quad (n = 0). \quad (4a)$$

Уравнения (2a) и (4a) являются скорее разностными, нежели дифференциальными. Их решение для $P_0, P_1, \dots, P_n, \dots$ может быть найдено методом последовательных подстановок с учетом того, что $\sum_{i=0}^{\infty} P_i = 1$. Этот метод применяется следующим образом. Положим $P_0 = P_0$.

Из уравнения (4a) находим

$$P_1 = (\lambda/\mu) P_0.$$

Полагая в уравнении (2a) $n = 1$ и подставляя значение P_1 , имеем

$$P_2 = (\lambda/\mu)^2 P_0.$$

Полагая в уравнении (2a) $n = 2$ и подставляя выражение для P_2 , находим

$$P_3 = (\lambda/\mu)^3 P_0$$

и вообще

$$P_n = (\lambda/\mu)^n P_0.$$

Суммируя соответствующие члены, получаем

$$\sum_{i=0}^{\infty} P_i = P_0 \sum_{i=0}^{\infty} (\lambda/\mu)^i. \quad (5)$$

Предположим, что $\lambda/\mu < 1$ (т. е. что средняя скорость появления заказов меньше средней скорости обслуживания заказа; это условие необходимо для того, чтобы рост очереди не был чрезмерным). Так как

$$\sum_{i=0}^{\infty} P_i = 1 \quad \text{и} \quad \sum_{n=0}^{\infty} (\lambda/\mu)^n = \frac{1}{1 - (\lambda/\mu)},$$

то

$$\frac{1}{1 - \lambda/\mu} = P_0.$$

Отсюда

$$P_0 = 1 - (\lambda/\mu). \quad (6)$$

Подставляя значение P_0 в предыдущее выражение для P_n , находим вероятность появления очереди длины n . Имеем

$$P_n = (\lambda/\mu)^n (1 - \lambda/\mu) \quad (7)$$

при $\lambda/\mu < 1$. Отношение λ/μ иногда называют *интенсивностью нагрузки*. Физический смысл его — средний объем обслуживания в единицу времени,

измеряемый в условных, по Кендаллу, единицах — *эрлангах*. Эта единица была введена в честь А. К. Эрланга, внесшего большой вклад в теорию массового обслуживания.

Найдем $\bar{n}$ — среднюю длину очереди. По определению, так как $\sum P_n = 1$, то

$$\bar{n} = \sum_{n=0}^{\infty} n P_n. \quad (8)$$

Подставляя в уравнение (8) значение P_n из (7), находим

$$\begin{aligned} \bar{n} &= \sum_{n=0}^{\infty} n (\lambda/\mu)^n (1 - \lambda/\mu) = (1 - \lambda/\mu) \sum_{n=0}^{\infty} n (\lambda/\mu)^n = \\ &= (1 - \lambda/\mu) [\lambda/\mu + 2(\lambda/\mu)^2 + 3(\lambda/\mu)^3 + \dots] = \\ &= (\lambda/\mu) (1 - \lambda/\mu) [1 + 2(\lambda/\mu) + 3(\lambda/\mu)^2 + \dots]. \end{aligned} \quad (9)$$

Это выражение можно вычислить, применяя поочередно операции интегрирования и дифференцирования к ряду, стоящему в (9) в скобках.

Обозначим этот ряд через $S(\lambda/\mu)$ и проинтегрируем его почленно. Выражение

$$\int_0^{\lambda/\mu} S(\lambda/\mu) d(\lambda/\mu) = \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 + \dots + \left(\frac{\lambda}{\mu}\right)^n + \dots$$

представляет собою геометрическую прогрессию с суммой $(\lambda/\mu)/[1 - (\lambda/\mu)]$. Дифференцируя эту сумму по λ/μ , получаем $1/[1 - (\lambda/\mu)]^2$. Следовательно, сумма ряда

$$S(\lambda/\mu) = \frac{1}{[1 - (\lambda/\mu)]^2}. \quad (10)$$

Подставляя это выражение в уравнение (9), находим

$$\bar{n} = (\lambda/\mu) (1 - \lambda/\mu) \frac{1}{[1 - (\lambda/\mu)]^2}. \quad (11)$$

Таким образом, при данных условиях средняя длина $\bar{n}$ очереди

$$\bar{n} = \frac{\lambda/\mu}{1 - \lambda/\mu}, \quad \lambda/\mu < 1. \quad (12)$$

Пример образования очереди на единственной станции обслуживания. Пусть известно, что среднее число λ заказов, поступающих на единственную станцию обслуживания в течение дня, равно 10, а средняя скорость обслуживания $\mu = 20$ заказов в день. Тогда отношение λ/μ равно $10/20 = 1/2$. Подставляя эту величину в уравнения (7) и (12), находим

$$P_n = \left(\frac{1}{2}\right)^n \left(1 - \frac{1}{2}\right) \quad \text{и} \quad \bar{n} = \frac{1/2}{1 - 1/2} = 1.$$

Вероятность появления очереди из 0, 1, 2, ... заказов в любой момент времени имеет вид

n	0	1	2	3	4	...
P_n	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	...

Интересно отметить (см. уравнение (12)), что с увеличением интенсивности нагрузки (λ/μ) средняя длина очереди быстро возрастает и при $\lambda/\mu \rightarrow 1$ величина $\bar{n}$ становится бесконечно большой. (Строго говоря, уравнение для $\bar{n}$, конечно, перестает иметь силу при $\lambda/\mu = 1$.)

Подставляя несколько значений λ/μ в уравнение (12), можно найти, как изменяется средняя длина очереди в зависимости от λ/μ .

Интенсивность нагрузки	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{15}{16}$	$\frac{31}{32}$	...
Средняя длина очереди	1	3	7	15	31	...

Интервалы времени между заказами. Пусть время обслуживания заказа равно T . Предположим, что распределение числа заказов, поступивших в течение этого периода, подчиняется закону Пуассона *. Тогда, если средняя продолжительность интервала времени между заказами равна a , то среднее число λ поступивших заказов $\lambda = T/a$. Пусть, например, T равно одному часу, а λ — шести новым заказам, поступившим в течение часа. Тогда средний интервал времени, разделяющий моменты поступления заказов, составляет 10 минут ($1/6$ часа): $a = T/\lambda = 1/6$.

Найдем плотность распределения вероятностей для промежутков времени t_a между моментами поступления заказов. Пусть $p(t_a)$ — вероятность того, что в течение интервала времени t_a , следующего за моментом появления заказа, новых поступлений нет, а $\lambda \Delta t_a$ (или, что то же, $\Delta t_a T/a$) — вероятность поступления одного нового заказа в течение промежутка времени $(t_a, t_a + \Delta t_a)$. Тогда вероятность сложного события, состоящего в том, что в течение t_a новые заказы не поступают, а на интервале $(t_a, t_a + \Delta t_a)$ поступает точно один заказ, равна произведению

$$p(t_a) \left[\frac{T}{a} \cdot \Delta t_a \right].$$

Без потери общности можно положить $T = 1$; тогда последнее выражение примет вид

$$p(t_a) \cdot \Delta t_a / a.$$

Таким образом,

$$p(t_a) - p(t_a + \Delta t_a) = p(t_a) (\Delta t_a) / a$$

или, так как

$$p(t_a) - p(t_a + \Delta t_a) = -\Delta p(t_a),$$

то

$$-\Delta p(t_a) = \frac{1}{a} p(t_a) \Delta t_a. \quad (13)$$

Физический смысл выражения (13) — вероятность появления точно одного нового заказа в течение интервала Δt_a . Из (13) видно, что плотность распределения вероятностей для t_a имеет вид $\frac{1}{a} p(t_a)$.

* Это означает следующее. Пусть λ — среднее число заказов, поступающих в течение некоторого промежутка времени. Тогда вероятность появления точно n заказов в течение того же промежутка времени

$$f(n) = \frac{\lambda^n e^{-\lambda}}{n!}.$$

Проиллюстрируем соотношение (13) следующим примером. Пусть некоторый интервал времени (рис. 14.1) разделен на четыре равных промежутка, в течение каждого из которых либо поступает один заказ, либо заказы вообще не поступают. Пусть вероятность отсутствия новых заказов

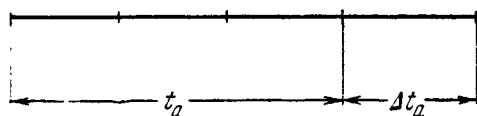


Рис. 14.1. Шкала времени.

в течение любого из этих четырех промежутков времени равна $1/3$. Тогда вероятность отсутствия новых заказов в течение первых трех промежутков $p(t_a) = (1/3)^3 = 1/27$. Вероятность поступления одного нового заказа в течение четвертого (либо любого другого) промежутка $\Delta t_a/a = 1 - 1/3 = 2/3$.

Отсюда вероятность того, что в течение первых трех интервалов новых поступлений не будет, а в течение четвертого появится один заказ, равна

$$p(t_a)(\Delta t_a/a) = 1/27 \cdot 2/3 = 2/81.$$

Этот результат может быть получен и иначе. Действительно,

$$p(t_a + \Delta t_a) = (1/3)^4 = 1/81, \quad p(t_a) = (1/3)^3 = 1/27,$$

откуда

$$-\Delta p(t_a) = p(t_a) - p(t_a - \Delta t_a) = 1/27 - 1/81 = 2/81,$$

что согласуется с уравнением (13). Разделив обе части уравнения (13) на $(-\Delta t_a)$, находим

$$\Delta p(t_a)/\Delta t_a = -p(t_a)/a.$$

Устремляя $\Delta t \rightarrow 0$, получаем

$$dp(t_a)/dt_a = -p(t_a)/a, \quad (14)$$

но, так как $1/a = \lambda$, то

$$dp(t_a)/dt_a = -\lambda p(t_a). \quad (15)$$

Решение этого дифференциального уравнения имеет вид

$$p(t_a) = Ke^{-\lambda t_a}. \quad (16)$$

Постоянная K равна единице. Это вытекает из начальных условий

$$p(0) = 1. \quad (17)$$

Окончательно получаем

$$p(t_a) = e^{-\lambda t_a}. \quad (18)$$

Таким образом, плотность распределения вероятностей для интервала времени между моментами поступления заказов имеет вид*

$$\lambda e^{-\lambda t_a}.$$

* Приведем подробное решение уравнения (15):

$$dp(t_a)/dt_a = -\lambda p(t_a), \quad \text{или} \quad dp(t_a)/p(t_a) = -\lambda dt_a.$$

Интегрируя, находим

$$\ln p(t_a) = -\lambda t_a + C, \quad \text{или} \quad p(t_a) = e^{-\lambda t_a + C}.$$

Эта формула может быть переписана иначе:

$$p(t_a) = Ke^{-\lambda t_a}, \quad \text{где} \quad K = e^C.$$

Так как $p(0) = 1$, то $K = 1$, и окончательно

$$p(t_a) = e^{-\lambda t_a}.$$

Время обслуживания. Ранее уже рассматривалась задача определения средней длины очереди на пункте обслуживания для случая, когда потоки на входе и выходе его носят случайный характер. Кендалл [20] показал, что если входной поток на единственной станции обслуживания подчиняется экспоненциальному закону распределения (см. предыдущий раздел), то средняя длина очереди $\bar{n}$ может быть выражена через среднюю скорость поступления заказов, среднюю скорость обслуживания и дисперсию времени обслуживания. В обозначениях, используемых в настоящей главе, это уравнение имеет вид

$$\bar{n} = \frac{\lambda}{\mu} + \frac{\lambda^2 \sigma_{t_s}^2 + (\lambda/\mu)^2}{2[1 - (\lambda/\mu)]}, \quad (19)$$

где $\sigma_{t_s}^2$ — дисперсия времени обслуживания t_s . Из уравнения (19) видно, что при фиксированных значениях λ и μ средняя длина очереди увеличивается с ростом дисперсии $\sigma_{t_s}^2$. При постоянных λ и μ минимальная средняя длина очереди соответствует $\sigma_{t_s}^2 = 0$, т. е. постоянному времени обслуживания. Иными словами, при заданной средней скорости поступления заказов λ и постоянном времени обслуживания μ

$$\bar{n} = \frac{\lambda}{\mu} + \frac{(\lambda/\mu)^2}{2[1 - \lambda/\mu]}. \quad (20)$$

Легко показать, что если время обслуживания подчиняется экспоненциальному закону с отрицательным показателем и средним значением μ , то дисперсия $\sigma_{t_s}^2 = 1/\mu^2$. В этом случае уравнение (19) принимает вид

$$\bar{n} = \frac{\lambda/\mu}{1 - \lambda/\mu},$$

что согласуется с (12). Из уравнения (19) еще раз видно, что, когда средняя скорость поступления заказов приближается к средней скорости обслуживания ($\lambda \rightarrow \mu$), средняя длина очереди безгранично возрастает. Из этого уравнения также следует, что при заданном распределении времени обслуживания средняя длина очереди может быть сокращена только путем уменьшения величины λ/μ . В общем, можно сказать, что характер поведения очереди целиком определяется отношением λ/μ . Например, при уменьшении λ/μ величина $1 - (\lambda/\mu)$ возрастает, а длина очереди сокращается. Решение проблемы очередей и сводится, по сути говоря, к подбору определенного соотношения между затратами на сокращение длины очереди и издержками из-за недостаточного использования средств обслуживания.

Время ожидания обслуживания. Найдем среднее время ожидания обслуживания на единственной станции обслуживания при случайном потоке заказов на входе ее. Пусть $\bar{t}_w$ — среднее время ожидания обслуживания, $\bar{t}_s$ — среднее время обслуживания. Тогда $\bar{t}_w + \bar{t}_s$ — среднее время, затрачиваемое на обслуживание и его ожидание. При средней скорости поступления заказов λ

$$\bar{n} = \lambda(\bar{t}_w + \bar{t}_s),$$

откуда

$$\bar{t}_w = \frac{\bar{n}}{\lambda} - \bar{t}_s. \quad (21)$$

При случайных потоках на входе и выходе станции обслуживания

$$\bar{n} = \frac{\lambda/\mu}{1 - \lambda/\mu}.$$

Можно показать также, что

$$\bar{t}_s = 1/\mu.$$

Подставляя последние два выражения в (21), находим среднее время ожидания обслуживания на единственной станции обслуживания:

$$\bar{t}_w = \frac{1}{\mu - \lambda} - \frac{1}{\mu}. \quad (22)$$

При случайном потоке заказов на входе единственной станции и заданном законе распределения времени обслуживания с дисперсией σ_s^2 , и средним значением $\bar{t}_s$ среднее время ожидания обслуживания *

$$\bar{t}_w = \frac{1}{\lambda} \left[\frac{\lambda}{\mu} + \frac{\lambda^2 \sigma_s^2 + (\lambda/\mu)^2}{2[1 - (\lambda/\mu)]} \right] - \bar{t}_s. \quad (23)$$

ОБЩАЯ ПРОБЛЕМА ОБРАЗОВАНИЯ ОЧЕРЕДИ НА ЕДИНСТВЕННОЙ СТАНЦИИ ОБСЛУЖИВАНИЯ

В общем случае, когда скорость поступления заказов и скорость их обслуживания зависят от длины очереди, основные уравнения имеют вид

$$\frac{dP_n(t)}{dt} = -(\lambda_n + \mu_n) P_n(t) + \lambda_{n-1} P_{n-1}(t) + \mu_{n+1} P_{n+1}(t), \quad (24)$$

$$\frac{dP_0}{dt} = -\lambda_0 P_0(t) + \mu_1 P_1(t). \quad (25)$$

Процесс, приводящий к этим уравнениям, обычно называется процессом «рождения и смерти». Для иллюстрации его рассмотрим следующий пример.

Некоторый ресторан обслуживает посетителей с 17 до 21 часа, и возможности его таковы, что он может обслужить неограниченное число людей. Пусть посетители прибывают в случайном порядке. Тогда, в соответствии с принятыми обозначениями, вероятность появления посетителя в течение интервала $(t, t + \Delta t)$ равна $\lambda \Delta t$. Пусть вероятность ухода посетителя за тот же промежуток $(t, t + \Delta t)$ равна $\mu \Delta t$. Следовательно, по мере заполнения ресторана скорость обслуживания увеличивается. В соответствии с уравнениями (24) и (25) скорости поступления заказов и их обслуживания

$$\lambda_n = \lambda, \quad \mu_n = \mu.$$

Дифференциальные уравнения для $P_n(t)$ — вероятности нахождения n посетителей в ресторане к моменту времени t — записываются следующим образом:

$$\frac{dP_n}{dt} = -(\lambda + n\mu) P_n(t) + \lambda P_{n-1}(t) + (n+1)\mu P_{n+1}(t), \quad n > 0, \quad (26)$$

$$\frac{dP_0}{dt} = -\lambda P_0(t) + \mu P_1(t). \quad (27)$$

Решение этих уравнений при начальных условиях $P_0(0) = 1$, $P_i(0) = 0$ ($i \geq 1$) имеет вид

$$P_n(t) = \frac{e^{-\lambda/\mu(1-e^{-\mu t})} [(\lambda/\mu)(1-e^{-\mu t})]^n}{n!}. \quad (28)$$

* При $\sigma_s^2 = 1/\mu$ и $\bar{t}_s = 1/\mu$ (см., например, экспоненциальное распределение с отрицательным показателем) уравнение (23) принимает вид

$$\bar{t}_w = \frac{1}{\lambda} \left[\frac{\lambda}{\mu} + \frac{(\lambda/\mu)^2}{1 - (\lambda/\mu)} \right] - \frac{1}{\mu} = \frac{1}{\mu - \lambda} - \frac{1}{\mu},$$

что совпадает с (22).

Отметим, что при $t \rightarrow \infty$

$$\lim_{t \rightarrow \infty} P_n(t) = P_n = \frac{e^{-\lambda/\mu} (\lambda/\mu)^n}{n!}. \quad (29)$$

Таким образом, вероятность того, что в ресторане находятся точно n клиентов спустя достаточно длинный промежуток времени задается распределением Пуассона со средним значением λ/μ . Пусть, например, в начале обслуживания $\lambda/\mu = 9$. В этом случае $e^{-(\lambda/\mu)} \approx 0,0001$ и вероятность наличия необслуженных клиентов $P_0 = 0,0001$. Для более позднего времени можно принять $\lambda/\mu = 0,1$, и здесь $P_0 = e^{-(\lambda/\mu)} = 0,9048$.

Пусть $\bar{n}(t)$ — среднее число посетителей, обслуживаемых в течение любого времени t . Эту величину можно определить и не решая уравнений. Действительно,

$$\bar{n}(t) = \sum n P_n(t), \quad (30)$$

$$\frac{d\bar{n}}{dt} = \sum \frac{n dP_n(t)}{dt} = \lambda - \mu \bar{n}(t). \quad (31)$$

Полагая, что в момент открытия ресторана очереди не было, находим

$$\bar{n}(t) = \frac{\lambda}{\mu} (1 - e^{-\mu t}). \quad (32)$$

ПРОБЛЕМА ОБРАЗОВАНИЯ ОЧЕРЕДЕЙ ПРИ НАЛИЧИИ МНОГИХ СТАНЦИЙ (КАНАЛОВ) ОБСЛУЖИВАНИЯ

Приведенный выше пример стал бы реалистичнее, если бы было установлено предельное число клиентов, которое ресторан может обслужить. Рассмотрим это ограничение на другом примере. Пусть заказы, поступающие в отдел погрузки, носят, как и в предыдущем примере, случайный характер, а вероятность поступления одного заказа за период $(t, t + \Delta t)$ равна $\lambda \Delta t$. Предположим, что эти заказы обрабатывает определенное число S служащих отдела. Рассмотрим случай, когда весь персонал S отдела настолько занят, что вновь поступившие заказы не могут быть обслужены немедленно, а ставятся в очередь. Если длительность обработки одного заказа порождает ситуацию, сходную с обслуживанием в ресторане, то говорят, что и в данном случае время обслуживания распределено экспоненциально. Будем говорить, что система находится в состоянии E_n , если суммарное число обслуживаемых и находящихся в очереди заказов равно n . В этом состоянии очередь существует, если $S < n$, и длина очереди равна $n - S$ заказов.

До тех пор, пока свободен хотя бы один служащий, мы имеем ситуацию, аналогичную разобранный выше. Однако в том случае, когда система пребывает в состоянии E_n а $n > S$, только S заказов находятся в обслуживании, и потому $\mu_n = S\mu$ при $n \geq S$. Эта ситуация описывается следующей системой дифференциальных уравнений:

$$dP_0/dt = -\lambda P_0(t) + \mu_1 P_1(t) \quad (n=0), \quad (33)$$

$$dP_n(t)/dt = -(\lambda + n\mu) P_n(t) + \lambda P_{n-1}(t) + (n+1)\mu P_{n+1}(t) \quad (0 < n < S), \quad (34)$$

$$dP_n(t) = (\lambda + S\mu) P_n(t) + \lambda P_{n-1}(t) + S\mu P_{n+1}(t) \quad (n \geq S). \quad (35)$$

Решение этих уравнений имеет очень сложный вид. Мы рассмотрим лишь предельное распределение вероятностей при $t \rightarrow \infty$. Можно показать, что при всех n существует единственный предел

$$\lim_{t \rightarrow \infty} P_n(t) = P_n.$$

Эти предельные вероятности удовлетворяют следующим уравнениям:

$$\lambda P_0 = \mu P_1 \quad (n=0), \quad (36)$$

$$(\lambda + n\mu) P_n = \lambda P_{n-1} + (n+1)\mu P_{n+1} \quad (1 \leq n < S), \quad (37)$$

$$(\lambda + S\mu) P_n = \lambda P_{n-1} + S\mu P_{n+1} \quad (n \geq S), \quad (38)$$

$$P_n = P_0 \frac{(\lambda/\mu)^n}{n!} \quad (n \leq S), \quad (39)$$

$$P_n = P_0 \frac{(\lambda/\mu)^n}{S! S^{n-S}} \quad (n \geq S). \quad (40)$$

Для того чтобы при всех n вероятности P_n подчинялись единственному закону распределения, необходимо выполнение условия $\lambda/\mu < S$. В противном случае очередь неограниченно возрастает.

Вероятность обязательного пребывания в очереди равна вероятности занятости всех средств обслуживания, или, что то же самое, вероятности нахождения в очереди S или более заказов. Обозначим эту вероятность через W . Тогда

$$W = \sum_{n=S}^{\infty} P_n = \frac{P_0}{S} \left(\frac{\lambda}{\mu}\right)^S \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu S}\right)^n = \left(\frac{\lambda}{\mu}\right)^S \frac{P_0}{S! (1 - \lambda/\mu S)}. \quad (41)$$

Известный интерес представляет вероятность того, что суммарное время обслуживания и ожидания обслуживания превзойдет заданную величину t . Обозначая эту вероятность через $P(>t)$ и пропуская громоздкие промежуточные выкладки, выпишем сразу выражение для нее в виде

$$P(>t) = e^{-\mu t} \left\{ 1 + \frac{W}{S} \left[\frac{1 - e^{-\mu S t [1 - (\lambda/\mu S) - (1/S)]}}{1 - (\lambda/\mu S) - (1/S)} \right] \right\}. \quad (42)$$

Вероятность того, что время между t и $t + \Delta t$ израсходовано на ожидание в очереди и обслуживание, равна $\Delta t \, dP(>t)/dt$. Тогда среднее время, затрачиваемое системой, складывается из средней длительности ожидания в очереди $\bar{t}_w$ и среднего времени обслуживания $1/\mu$:

$$\int_0^{\infty} t \frac{d[P(>t)]}{dt} dt = \bar{t}_w + \frac{1}{\mu}. \quad (43)$$

Разрешая последнее уравнение относительно $\bar{t}_w$, находим среднее время ожидания в очереди при наличии многих каналов обслуживания:

$$\bar{t}_w = \frac{P_0}{\mu S (S!) [1 - (\lambda/\mu S)]^2} \left(\frac{\lambda}{\mu}\right)^S. \quad (44)$$

Множитель P_0 определяется из условия $\sum_{n=0}^{\infty} P_n = 1$:

$$P_0 = \frac{1}{\sum_{n=0}^{S-1} (\lambda/\mu)^n / n! + \{(\lambda/\mu)^S / [S! (1 - \lambda/\mu S)]\}}. \quad (45)$$

Аналитический подход к проблемам массового обслуживания был использован в целом ряде случаев. Среди них анализ многоканальных систем с последовательным обслуживанием [26] и математическое описание роста популяций организмов [20]. Кендалл [20] перечисляет ряд задач, которые решались на базе различных предположений относительно

входного потока (регулярного, случайного или «эрланговского»), распределения времени обслуживания и объема обслуживающих средств, а также приводит библиографию работ по системам массового обслуживания.

В статье [10] Кобгам разработал теорию систем массового обслуживания с приоритетом для следующих двух случаев: (а) один канал обслуживания, длительность времени обслуживания произвольна, входной поток случаен; (б) система обслуживания содержит несколько каналов, время обслуживания распределено по экспоненциальному закону, входной поток случаен.

Анализ этих систем привел к бесконечной итерационной процедуре. Холли в [18] предложил другой метод, позволяющий исключить бесконечную итерацию. Гавер [17] исследовал влияние времени обслуживания на время ожидания в очереди. В работе Баррера [3] рассматривается задача обслуживания заказов, доступных лишь в течение определенного промежутка времени (нетерпеливые клиенты), причем выбор заказов для обслуживания носит случайный характер (индифферентный обслуживающий персонал).

ПРИМЕНЕНИЕ МЕТОДА МОНТЕ-КАРЛО К ЗАДАЧАМ МАССОВОГО ОБСЛУЖИВАНИЯ

Теория массового обслуживания дает методику определения средней длины очереди и среднего времени ожидания для случая, когда скорости поступления заказов и их обслуживания известны. Если издержки, связанные с пребыванием в очереди и обслуживанием, определены, то можно установить и оптимальное отношение между ними.

Рассмотрим, например, нерегулируемый поток заказов на входе системы. Требуемое соотношение между обеими группами затрат достигается в этом случае подбором закона управления средствами обслуживания, причем все необходимые величины могут быть найдены чисто аналитическим путем, без эксперимента. Придавая, к примеру, различные значения скорости обслуживания (μ , σ_s^2), можно исследовать ее влияние на среднее время ожидания $\bar{t}_w$ и связанные с ним затраты. Подобный анализ, однако, может быть произведен и другим путем, а именно с помощью метода Монте-Карло.

Рассмотрим частный пример — доставку на дом товаров, закупленных в магазине. Можно организовать большой автомобильный парк с многочисленным персоналом и обеспечить каждому клиенту доставку его товара в течение одного дня. Это, однако, потребует крупных капитальных затрат и, кроме того, может привести к значительным простоям оборудования и персонала. При слишком же малом размере парка появляется угроза потери части клиентов из-за длительности доставки либо возникает необходимость в сверхурочной работе или аренде дополнительных средств обслуживания. Возникает следующий вопрос: какой размер парка целесообразен и какое количество арендной платы или сверхурочного времени себя оправдывает?

Одним из возможных, но малопрактичных подходов к решению этой задачи является краткосрочное испытание каждого возможного варианта и учет получаемых результатов. Очевидно, этот способ связан с разрушением нормально действующего процесса и требует чрезмерных затрат. Кроме того, очень трудно создать одинаковые условия для всех испытываемых вариантов. В этом и состоят большие преимущества метода Монте-Карло: он не требует разрушения действующего процесса и тем не менее дает

возможность оценить (при заданных условиях) сколь угодно большое число различных решений задач. Рассмотрим для примера задачу об оптимальном размере автопарка.

Вначале определяется средняя скорость поступления тьюков на погрузочный пункт. Она может быть постоянной, но может и варьироваться в широких пределах. В последнем случае надо установить причину этих колебаний. Далее, анализируя записи предыдущих операций, можно определить, как изменялось ежедневное количество тьюков, например, за последний год, и выразить изучаемое колебание в величинах стандартных отклонений для времени поступления заказов. Число тьюков, прибывающих для доставки, может подчиняться любому из многочисленных законов распределения. Примем, однако, что это распределение нормальное, со средним значением 1000 тьюков в день и стандартным отклонением в 100 тьюков. Это означает, что средняя скорость поступления заказов $\lambda = 1000$ тьюков в день.

Для определения нужного объема автопарка требуется знать еще два параметра. Первый из них — ежедневные затраты на эксплуатацию грузового автомобиля, включающие в себя как постоянную, так и переменную части. Примем, что эти затраты составляют 25 долларов в день. Во-вторых, должны быть известны ежедневные расходы, связанные с задержкой грузов. Вообще говоря, уровень этих затрат определить очень трудно. Однако зачастую это можно сделать, и притом с удовлетворительной точностью, путем анализа имеющихся записей по продаже и отправке. Основная цель этого анализа — выявить разницу в прибылях, приносимых клиентами, которые обслуживались в течение одного дня, и теми, кто по различным причинам был вынужден длительно ожидать закупленных грузов. Для простоты рассмотрим задачу, в которой затраты, связанные с задержкой груза, не учитываются, но зато стратегия пункта требует, чтобы все тьюки, готовые к отправке, были доставлены в тот же день. Это может потребовать сверхурочной работы. Будем предполагать, что требуемая длительность сверхурочного времени зависит от скорости обслуживания, а один час сверхурочной работы оценивается в 8 долларов.

Составим таблицу, аналогичную 14.3, и изучим деятельность системы отправки в сочетании с каждым из трех вариантов автопарка. В принципе можно исследовать работу системы в течение любого промежутка времени и при любом размере грузового автопарка.

В данном иллюстративном примере рассматривается период из пяти следующих друг за другом дней. Смысл столбцов (1) и (2) ясен. Столбец (3) требует некоторых пояснений. Прежде всего напомним, что распределение числа отправляемых грузов считается известным и в данном случае принято нормальным. Составляется некоторая выборка из этого распределения. Она должна быть построена таким образом, чтобы учесть характер случайного процесса поступления грузов. Так, если наиболее часто поступает 1000 тьюков, то и выборка должна группироваться вокруг этого числа; в этом случае вероятность появления некоторой случайной величины в выборке совпадает с вероятностью появления этого количества заказов в действительности.

Берем пять первых следующих друг за другом значений из таблицы нормальных случайных чисел (табл. 7.5) и переписываем их в третий столбец таблицы 14.3*. Отрицательные числа относятся к объему заказов меньшему, чем средний уровень, положительные — соответствуют объему выше среднего. Для того чтобы преобразовать эти стандартные единицы

* Конечно, можно начать с любого места таблицы случайных чисел.

Таблица 14.3

Число пакетов, подлежащих доставке					Число пакетов, которое может быть доставлено			Издержки на сверхурочное время	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Число грузовиков в парке	Дни	Табличное значение	Число поступивших пакетов = $1000 + 100 \times (3) *$	Общее число пакетов, подлежащих доставке: $(4) + \text{остаток от предыдущего дня}$	Табличное значение + число пакетов, которое может быть доставлено на одном грузовике	Общее количество пакетов, которое может быть доставлено в течение рабочего дня $100 + 10(6)$	Число пакетов, оставшихся для обработки при отсутствии сверхурочного времени $(5) - (7)$	Число пакетов, подлежащих отправке в сверхурочное время $(4) - (7)$	$(9) \div \frac{(7)}{(1)(8)} \times \$8,0$
10	1	2,455	1246	1246	-0,323	968	278	278	\$184
	2	-0,531	947	1225	-1,940	806	419	141	\$112
	3	-0,634	937	1356	0,697	1070	286	0	
	4	1,279	1128	1414	3,521	1352	62	0	
	5	0,046	1005	1067	0,321	1032	35	0	
Итого									\$296
12	1	2,455	1246	1246	-0,323	1161	85	85	\$56
	2	-0,531	947	1032	-1,940	967	65	0	
	3	-0,634	937	1002	0,687	1284			
	4	1,279	1128	1128	3,521	1623			
	5	0,046	1005	1005	0,321	1239			
Итого									\$56
15	1	2,455	1246	1246	-0,323	1452			
	2	-0,531	947	947	-1,940	1209			
	3	-0,634	937	937	0,697	1605			
	4	1,279	1128	1128	3,521	2028			
	5	0,046	1005	1005	0,321	1548			
Итого									\$0

* Цифры, стоящие в скобках, обозначают номер соответствующего столбца.

в истинное число заказов, надо умножить число этих единиц на 100 (получив, таким образом, отклонение от среднего значения) и прибавить к 1000 (средней величине). Полученные значения приведены в столбце (4) таблицы 14.3. Аналогичным образом составляются столбцы (6) и (7), но только здесь стандартное отклонение принято равным 10. В столбце (8) приведено число неотправленных пакетов. На их доставку и требуется сверхурочное время. В столбце (10) приведена стоимость сверхурочной доставки в предположении, что скорость обслуживания остается неизменной.

Теперь можно найти суммарные еженедельные затраты по каждому автопарку. Так как ежедневные затраты на эксплуатацию одного грузовика равны 25 долларам, то за 5 дней издержки составят 125 долларов. Умножая эту величину на число автомобилей в парке и складывая полученный результат со стоимостью сверхурочной работы, получаем следующие результаты:

(а) 10 грузовиков: $(10 \times \$125) + \$296 = \$1546$,

(б) 12 грузовиков: $(12 \times \$125) + \$56 = \$1556$,

(в) 15 грузовиков: $(15 \times \$125) = \1875 .

Следовательно, в данном случае наиболее экономичным оказался парк из 10 грузовиков.

Конечно, при анализе практических задач число недель должно быть взято гораздо большим. Кроме того, необходимо принять во внимание и другие параметры автопарка. В этом случае можно определить с помощью статистических процедур, описанных в главе 20, насколько разница в суммарных затратах существенна. Если это различие невелико, то для исследования надо использовать выборки большого объема.

С помощью сравнительно небольших дополнительных расчетов можно определить эффективность затрат, связанных с изменением стратегии обслуживания. Например, можно оценить режим доставки пакетов в пределах двух, трех или любого заданного числа дней после покупки. Можно также учесть затраты на аренду грузовиков в дни перегрузки. После таких расчетов предприниматель в состоянии оценить различные стратегии обслуживания.

В ряде случаев прошлые данные о скорости поступления заказов и их обслуживании могут не укладываться в рамки известных аналитических законов распределения (таких, как нормальное или пуассоновское). Тогда они проходят еще предварительную обработку. Вначале собираются отдельно данные о скорости поступления заказов и о скорости их обслуживания, и каждая запись нумеруется. Затем с помощью таблицы случайных чисел составляются выборки из этих значений. Они и используются далее обычным образом. Подобная процедура предполагает статистическую независимость скоростей поступления заказов и их обслуживания. Это предположение, однако, не всегда справедливо, например, скорость работы обслуживающего персонала может зависеть от уровня загрузки. Поэтому целесообразно опытные данные сгруппировать вначале попарно (скорость поступления заказов и соответствующая ей скорость обслуживания) и уже из этой популяции извлекать выборки.

Необходимо иметь в виду, что рассматриваемый здесь пример намеренно упрощен. На практике может возникнуть необходимость в учете еще целого ряда факторов, например сезонных колебаний в требованиях по доставке и разницы в скоростях поступления или обслуживания заказов в отдельные дни недели. Может потребоваться учет поломок грузовиков, погоды, прогулов и даже колебаний в скорости доставки в зависимости от загрузки. Тем не менее метод Монте-Карло дает возможность учесть все эти факторы; необходимо только иметь опытные данные, соответствующие поставленным задачам.

ПРИМЕНЕНИЕ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ В АВИАЦИОННОЙ ПРОМЫШЛЕННОСТИ

В компании «Boeing Aircraft» возникла следующая задача. Необходимо было определить оптимальное число служащих на инструментальных складах одной из фабрик этой компании. Число складов равнялось 60, они были разбросаны по всей территории предприятия, и в них хранился инстру-

мент, необходимый для работы механикам. В обязанности обслуживающего персонала входила выдача этого инструмента и его прием. Данная проблема была отнесена к задачам массового обслуживания и изучена секцией анализа при группе математического обслуживания объединения физических исследований этой компании [7].

Длительность обслуживания механиков была установлена в результате выборочного обследования. Наблюдатель фиксировал момент появления механиков у стола выдачи и с помощью секундомера определял время обслуживания одного человека. Таким способом были получены удовлетворительные данные о распределении времени обслуживания.

Далее находилось распределение для моментов поступления заказов. С этой целью моменты поступления всех заказов фиксировались, а для облегчения сбора данных применялась следующая система. Была собрана небольшая установка, на передней панели которой были установлены два ручных переключателя и две сигнальные лампочки. Внутри помещалась шестивольтовая батарея. При каждом нажатии на кнопку одного из переключателей подавался сигнал на один из каналов двухканального самописца фирмы «Brush». Перо прибора делало отметку на ленте, и одновременно с этим поступал световой сигнал. Работа хронометриста сводилась, таким образом, к нажатию на кнопку переключателя при каждом появлении механика.

Этот же прибор был приспособлен и для фиксации интервала времени, требуемого для обслуживания или совершения другой операции. В этом случае положение переключателя фиксировалось, и прибор оставался включенным до конца операции, что отмечалось на ленте непрерывной линией. С помощью этих записей далее определялись моменты начала и конца операции и строились законы распределения для моментов поступления заказов и интервалов их обслуживания.

Если бы каждый механик появлялся на пункте обслуживания в тот момент, когда заканчивалась операция с его предшественником, то длина очереди, а также время простоя или ожидания обслуживания равнялись бы нулю. Однако в действительности время обслуживания различно, а моменты поступления заказов носят случайный характер. Поэтому в некоторые промежутки времени образуется очередь механиков, а в другие простаивает обслуживающий персонал. Возникает проблема минимизации стоимости времени, затрачиваемого непроизводительно как механиками, так и обслуживающим персоналом. Поскольку управляемой величиной в данном случае является число служащих, то эта проблема сводится к задаче о нахождении численности обслуживающего персонала, минимизирующей общие затраты из-за простоя.

Для одного конкретного склада было установлено, что средний интервал времени между очередными появлениями механиков равен 35 секундам, а среднее время обслуживания одного механика — 50 секунд. В предшествующем изложении индексом λ обозначалась средняя скорость поступления заказов, а μ — средняя скорость обслуживания в течение некоторого фиксированного интервала времени. Если принять, что этот фиксированный промежуток времени равен среднему времени обслуживания, то

$$\lambda = \frac{1/35}{1/50} = \frac{\text{число появлений механиков/секунды}}{\text{среднее время обслуживания/секунды}} =$$

$$= 1,43 \text{ механика, появляющихся в течение среднего интервала времени обслуживания.}$$

Иными словами, на данном складе в течение среднего интервала времени обслуживания (т. е. в течение 50 секунд) в среднем появляется 1,43 механика. Средняя скорость обслуживания в течение этого 50-секундного интервала равна, конечно, единице ($\mu = 1$). Для того чтобы найти среднее время ожидания T_w (необходимо только помнить, что в качестве единицы масштаба времени принято 50 секунд — среднее время обслуживания), надо воспользоваться уравнением (44) для случая многоканального обслуживания, где S означает число служащих на складе:

$$T_w = \frac{P_0}{S\mu(S!) [1 - (\lambda/\mu S)]^2} \left(\frac{\lambda}{\mu} \right)^S.$$

Для нахождения P_0 используется уравнение (45). Подставляя в последнее уравнение значения P_0 , S , μ , λ , можно найти среднее время ожидания обслуживания (в 50-секундных единицах) на данном складе. Подобные расчеты при двух, трех и четырех служащих приведены в таблице 14.4.

Таблица 14.4

λ	μ	S	P_0	T_w	
				В 50-сек. единицах	Секунды
1,43	1	2	0,166	1,040	52,0
1,43	1	3	0,228	0,135	6,8
1,43	1	4	0,237	0,025	1,3

Рассмотрим, как было найдено число $P_0 = 0,166$ (первая строка, четвертый столбец). Подставляя в уравнение (45) значения $\lambda = 1,43$, $\mu = 1$ и $S = 2$, находим

$$P_0 = \frac{1}{\sum_{n=0}^{\infty} (1,43)^n / n! + (1,43)^2 / \{2! [1 - (1,43/2)]\}} =$$

$$= 1 / (1 + 1,43 + 3,58) = 1/6,01 = 0,166.$$

Значение T_w , соответствующее $S = 2$, рассчитывается по формуле

$$T_w = \frac{0,166}{2 \cdot 1 \cdot 2! [1 - (1,43/2)]^2} (1,43)^2 = \frac{0,166}{4 \cdot (0,285)^2} (1,43)^2 =$$

$$= 1,04 \text{ 50-секундной единицы.}$$

Среднее время ожидания в секундах равно $1,04 \times 50 = 52$ секунды. Аналогичным образом вычисляются и все остальные члены этой таблицы.

Так как механики появляются на складе в среднем через 35 секунд, то за один 7,5-часовой рабочий день в среднем поступает

$$\frac{7,5 \cdot 3600}{35} = 770 \text{ заказов.}$$

Поскольку на обслуживание одного заказа в среднем тратится 50 секунд, то для удовлетворения 770 заявок потребуется

$$\frac{770 \cdot 50}{3600} = 10,7 \text{ часа}$$

работы обслуживающего персонала в день.

Однако продолжительность рабочего дня служащих также составляет 7,5 часа. Поэтому, если данный склад будет обслуживать один человек, то это приведет к непрерывному росту очереди у склада. При двух клерках время обслуживания возрастает до 15 часов, но зато появляется время простоя этого персонала, равное $15 - 10,7 = 4,3$ часа. Из таблицы 14.4 видно, что среднее время ожидания обслуживания для каждого механика при двух клерках на складе составляет 52 секунды. Поскольку среднее число заявок в день равно 770, то среднее время простоя механиков в течение рабочего дня равно 770×52 секундам, или примерно 11 часам.

Пусть час работы обслуживающего персонала оценивается в 2 доллара, а час работы механика — в 5 долларов. В этом случае время простоя клерков обойдется в $4,3 \times 2 = 8,6$ доллара, а время простоя механиков — в $11,1 \times 5 = 55,5$ доллара. Общие потери из-за простоя составляют 64,10 доллара.

Аналогично рассчитываются издержки при трех клерках на складе, они равны 31 доллару. Но уже при четырех клерках эти затраты возрастают до 40 долларов и продолжают увеличиваться с ростом числа клерков. Следовательно, при данных предположениях оптимальное число служащих равно трем.

В реальной задаче для удобства вначале были подсчитаны затраты при различных скоростях поступления заказов, разной численности обслуживающего персонала и различных отношениях затрат на время простоя механиков и клерков. По результатам этих расчетов было построено семейство кривых, что позволило быстро решать задачу об оптимальном числе служащих для любого инструментального склада.

ЗАКЛЮЧЕНИЕ

Метод Монте-Карло может быть использован для решения любой задачи массового обслуживания, если только требуемые данные можно собрать. Хотя в некоторых случаях отыскание решения связано с большими трудностями, все же в большинстве задач удастся, сочетая теоретические выводы и методы Монте-Карло, найти решение без особых хлопот*.

В настоящее время в Массачусетском технологическом институте проводится работа по созданию таблиц, облегчающих решение задач массового обслуживания. Имеется также успешный опыт использования цифровых и аналоговых машин для моделирования процессов в очередях и нахождения оценок этих процессов. Значительная работа, проделанная к настоящему времени в этой области, показывает, что интуиция плохо помогает решению задач массового обслуживания. Методы же, описанные в настоящей главе, всегда приводили к значительным улучшениям в производстве.

ЛИТЕРАТУРА

1. Adler R. B. and Fricker S. J., The Flow of Scheduled Air Traffic—I and II, M.I.T. RLE Technical Report No. 198, May 2, 1951, and No. 199, Aug. 13, 1951.
2. Bailey N. T. J., On Queuing Processes with Bulk Service, J. R. Statist. Soc., 16, No. 2, 80—87 (1954).

* Кемп [9] разработал аналитические методы, которые, в сочетании с методом Монте-Карло, позволяют снизить объем требуемых вычислений. Это ему удалось сделать на базе идей (а) теории управления (нахождение верхней и нижней границ для математического ожидания) и (б) теории вероятностей.

3. B a r r e r D. Y., A Waiting Line Problem Characterized by Impatient Customers and Indifferent Clerks, Third Annual Meeting, Operations Research Society of America, New York, June 4, 1955.
4. B e l l G. E., Operational Research into Air Traffic Control, J. R. Aero. Soc., 53, 965—976 (Oct. 1949).
5. B e n s o n F. and C o x D. R., The Productivity of Machines Requiring Attention at Random Intervals, J. R. Statist. Soc., 13, 65—82 (1951).
6. B e r k e l e y G. S., Traffic and Trunking Principles in Automatic Telephony, Ernest Benn, Ltd., London, 1949.
7. B r i g h a m G e o r g e s, On a Congestion Problem in an Aircraft Factory, J. Opns. Res. Soc. Am., 3, no. 4, 412—428 (Nov. 1955).
8. B r i s b y M. D. J. and E d d i s o n R. T., Train Arrivals: Handling Costs, and the Holding and Storage of Raw Materials, J. Iron Steel Inst., 172, pt. 2, 171—183 (Oct. 1952).
9. C a m p G. D., Bounding the Solution of Practical Queuing Problems by Analytic Methods, in J. F. McCloskey and J. M. Copping (eds.), Operations Research for Management II, The Johns Hopkins Press, Baltimore, 1956.
10. C o b b a m A l a n, Priority Assignment in Waiting Line Problems, J. Opns. Res. Soc. Am., 2, no. 1, 70—76 (Feb. 1954).
11. C r o m m e l i n C. D., Delay Probability Formulae When the Holding Times Are Constant, P. O. Elect. Engrs' J., 25, pt. 1, 41—50 (Apr. 1932).
12. E d d i s o n R. T. and O w e n D. G., Discharging Iron Ore, Operat. Res. Quart., 4, no. 3, 39—50 (Sept. 1953).
13. E d i e L. C., Traffic Delays at Toll Booths, J. Opns. Res. Soc. Am., 2, no. 2, 107—138 (May 1954).
14. E v e r e t t J. L., State Probabilities in Congestion Problems Characterized by Constant Holding Times, J. Opns. Res. Soc. Am., 1, no. 5, 279—285 (Nov. 1953).
15. F e l l e r W., An Introduction to Probability Theory and Its Applications, John Wiley & Sons, New York, 1950.
16. F r y T. C., Probability and Its Engineering Uses, D. Van Nostrand Co., New York, 1928.
17. G a v e r D. P., The Influence of Servicing Times in Queuing Processes, J. Opns. Res. Soc. Am., 2, 139—149 (May 1954).
18. H o l l e y J. L., Waiting Lines Subject to Priorities, J. Opns. Res. Soc. Am., 2, no. 3, 341—343 (Aug. 1954).
19. J a c k s o n R. R. P., Queuing Systems with Phase Type Service, Operat. Res. Quart., 5, no. 4, 109—120 (Dec. 1954).
20. K e n d a l l D. G., On the Role of Variable Generation Time in the Development of a Stochastic Birth Process, Biometrika, 35, 316 (Dec. 1948).
21. K e n d a l l D. G., Some Problems in the Theory of Queues, J. R. statist. Soc., (B), 13, no. 2, 151—173 (1951).
22. K e n d a l l D. G., Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain, Ann. math. Statist., 24, no. 3, 338—354 (Sept. 1953).
23. L i n d l e y D. V., The Theory of Queues with a Single Server, Proc. Cambr. phil. Soc., 48, pt. 2, 277—289 (Apr. 1952).
24. M a r s h a l l B. D., Jr., Queuing Theory in J. F. McCloskey and F. N. Trefethen (eds.), Operations Research for Management, The Johns Hopkins University Press, Baltimore, 1954.
25. Marshalling and Queuing, Operat. Res. Quart., 3, no. 1, 1—15 (Mar. 1952).
26. M.I.T. Interim Report No. 2, Fundamental Investigations in Methods of Operations Research, Apr. 1, 1954—Nov. 30, 1954.
27. M.I.T. Summer Short Course in Operations Research, Technology Press, Massachusetts Institute of Technology, Cambridge, 1953.
28. M o l i n a E. C., Application of the Theory of Probabilities to Telephone Trunking Problems, Bell Syst. Tech. J., 6, 461 (1927).
29. M o r s e P. M., Stochastic Properties of Waiting Lines, J. Opns. Res. Soc. Am., 3, no. 3, 255—261 (Aug. 1955).
30. M o r s e P. M., G a r b e r H. N. and E r n s t M. L., A Family of Queuing Problems, J. Opns. Res. Soc. Am., 2, no. 4, 444—445 (Nov. 1954).
31. P o l l a c z e k F., Sur l'application de la théorie des fonctions au calcul de certaines probabilités continuées utilisées dans la théorie des réseaux téléphoniques, Ann. Inst. Poincaré, 10, no. 1, 1 (1946).
32. P o l l a c z e k F., Über eine Aufgabe der Wahrscheinlichkeitstheorie, Math. Z., 32, 64—100 and 729—750 (1930).

Глава 15

ЗАДЕРЖКИ ТРАНСПОРТА НА ПУНКТАХ СБОРА ПОШЛИН *

Основными задачами Управления порта г. Нью-Йорк являются развитие торговли в районе порта и транспортное обслуживание, которое выполняется путем организации работы различных служб.

Выполняются следующие виды работ: (1) координация погрузочно-разгрузочных работ, противопожарная и прочая аварийная служба, обработка багажа и организация стоянки машин в аэропортах; (2) распределение мест у причалов, хранение на складе и транспортировка грузов в морских портах; (3) погрузка автомашин, погрузка и отправление автобусов и распределение железнодорожных и автомобильных грузов по станциям назначения; (4) регулировка движения, предупреждение аварий и сбор платы за проезд через туннели и мосты — этому и посвящена данная глава.

Хотя этот перечень и не полон, он достаточен, чтобы показать обширное поле деятельности для специалистов в области исследования операций.

Методы исследования операций применяются для решения задач транспортного обслуживания Отделом стандартов Управления транспортных работ. Главное управление исполняет роль консультанта по вопросам, с которыми сталкиваются четыре отраслевых управления, отвечающих за материальные и финансовые результаты одного из четырех видов работ, указанных выше. В настоящее время методы исследования операций используются и Управлением контроля, где производится выборочная проверка различных счетов. Отдел стандартов использовал методы исследования операций для изучения работы полиции Управления порта — группы в 1000 человек, которая представляет собой самую большую обособленную категорию работников Управления.

Цель изучения состояла в том, чтобы выяснить, является ли состав полиции различных служб достаточным для выполнения работ и экономичным. Для достижения этой цели было необходимо тщательно проанализировать многочисленные операции, проводимые полицией, и установить стандарты на эти операции. В ряде случаев довольно трудно установить хорошие стандарты, в связи с чем изучение деятельности полиции, на которое сначала отводилось 6 месяцев, на самом деле потребовало 14 месяцев. Дополнительное время было израсходовано главным образом на исследование операций (подобные проведенным в данной главе), которые сначала не были предусмотрены, но которые желательно было бы проделать. Годовая экономия, полученная сразу после окончания исследования, более чем в десять раз превысила его стоимость при потенциальной будущей ежегодной экономии, превышающей стоимость проведенных исследований более чем в 20 раз. Кроме того, по капитальным вложениям экономия достигла почти десятикратной стоимости исследования. Значительную часть экономии, а также улучшение обслуживания клиентов работы обслуживающего персонала можно приписать исследованию операций.

* Настоящая глава представляет собой перепечатку статьи Лесли С. Эдая из «Journal of the Operations Research Society of America», 2, no. 2, 107—138 (May 1954).

СБОР ПЛАТЫ ЗА ПРОЕЗД

Сбор платы за проезд машин является важной частью работы полиции порта, им занято более четверти персонала полиции. На предварительных этапах нашего анализа было замечено, что результаты этой работы не вполне удовлетворительны. Качество обслуживания время от времени заметно изменялось. Иногда оно было значительно лучше необходимого и в работе сборщиков платы возникали перерывы, а иногда было неудовлетворительным и от клиентов поступали жалобы. Например, средняя задержка автомашин изменялась от 2 до 50 секунд.

До нашего анализа пункты сбора платы укомплектовывались почти всегда на основе субъективных мнений и взглядов, а количество сборщиков определялось прежде всего финансовыми ограничениями. Предполагалось, что в предстоящем году потребуется такое же количество сборщиков. Эти требования затем пересматривались руководством в свете ожидаемого годового объема движения и прошлого опыта, который помогал установить какой поток машин может обслужить сборщик. Наличный состав сборщиков, предусмотренный этими финансовыми ограничениями, распределялся полицейскими сержантами и их начальниками по различным дням и сменам. Таков типичный процесс управления.

В каждой смене действительное количество пунктов сбора платы, работающих в данный момент, определялось дежурным сержантом полиции, который стремился к наилучшему использованию наличного состава. Сержанты полиции работают в три смены и занимаются то сбором платы, то регулированием движения, поэтому им трудно ознакомиться с потоком транспорта в какой-либо смене. Основная работа сержантов состоит в примирении часто противоречивых требований потока с потребностями сборщиков иметь личные и обеденные перерывы. Так как полицейские сержанты имеют разный опыт и разные взгляды на организацию работы, результаты были неодинаковы. У некоторых выработались четкие навыки работы, у других — нет. Из бесед со сборщиками платы выяснилось, что их требования о перерывах часто не удовлетворялись. Вообще, если возникал конфликт, преимущество давалось клиенту за счет сборщика; но так как сбор платы — работа напряженная и нервная, то длительные периоды без перерывов весьма нежелательны.

Из дальнейшего обсуждения видно, что ставились следующие задачи: (1) оценить качество обслуживания клиентов и определить зависимость его от интенсивности движения на линии * сбора платы; (2) установить оптимальные нормативы обслуживания и (3) развить более точный метод распределения обслуживающего персонала и денежных ресурсов при том же времени на хорошо спланированные перерывы для сборщиков.

НАБЛЮДЕНИЯ

Первая регистрируемая величина — поток машин, прибывающих на площадку сбора платы. Один наблюдатель подсчитывал количество машин, прибывающих за 30 секунд, и записывал число вместе со временем (см. первый и второй столбец табл. 15.1). Было определено, что интервалы в 30 секунд являются самыми короткими из тех, которые наблюдатель может использовать для записи, не сбиваясь со счета.

* Линия сбора платы представляет собой узкий проезд на части шоссе. (Прим. перев.)

Таблица 15.1

Пример записанных данных

Время (после полу- дня), мин	Поток машин	Число машин в линиях				Число занятых линий
		А	6	10	Всего	
8.58	10	2	2	1	5	3
8.59	6 1	0 0	1 1	0 0	1 1	1 1
9.00	3 4	1 0	0 1	0 1	1 2	1 2
9.01	5	1	1	0	2	2
9.15 *	6 5	0 1	1 2	0 0	1 3	1 2
9.16	6 4	5 2	0 1	1 0	6 3	2 2
9.17	4 2	1 0	0 0	1 0	2 0	2 0
9.18	7	1	1	3	5	3
Итого	205	41	55	38	134	76
Число платежей **						
9.18		2102	79785	97466		
8.58		2034	79698	97416		
Итого		68	87	50	205	
* Пропущено 14 минут. ** Номера чеков кассовых аппаратов.						

Вторая регистрируемая величина — длина очереди в каждой открытой линии сбора платы. Она записывалась вторым наблюдателем также через 30 секунд синхронно с записями входного потока.

Третья величина — число денежных сделок. Она записывалась через полчаса и всякий раз, когда имелось изменение в количестве или типе линий сбора платы. В некоторых случаях количество и тип открытых линий регулировались полицейским сержантом, но иногда они регулировались группой обследования, которая должна была выяснить влияние специфических мер и создать небольшую перегруженность движения. Подсчет сделок дает возможность контролировать входной поток, которому он должен соответствовать, если рассматривать разницу в накоплении в начале и конце периода наблюдения. Важнее, что это также позволяет проводить подсчеты для каждой линии отдельно и для всех линий вместе.

В таблице 15.1 сведены все три типа данных для туннеля Линкольна, где были открыты в одном направлении три левосторонних пункта сбора платы, а интенсивность потока составляла 615 машин в час. Показаны также предварительные шаги в анализе данных, а именно: общая длина очереди для трех линий, количество занятых линий в каждом наблюдении и суммы по каждому столбцу.

РАСЧЕТЫ

Одной из основных искомых величин является средняя задержка. Однако сначала следует подсчитать полное время пребывания машины в линии сбора платы; оно включает как задержку, или время ожидания, так и пребывание на пункте, или время обслуживания. Полное время пребывания всех машин в линиях сбора, по данным выборочных наблюдений, равно 4020 секундам, среднее — 19,6 секунды.

Общее время пребывания машин на пункте за период наблюдения, указанный в таблице 15.1, равно суммарному числу занятых линий, умноженному на интервал наблюдения, т. е. равно 2280 секундам. Среднее время пребывания на пункте составляет 11,1 секунды. Средняя задержка на машину равна разности между средним полным временем на машину и средним временем пребывания на пункте, т. е. 8,5 секунды.

Другой исследуемой величиной является средняя задержка, выраженная в единицах времени пребывания на пункте и названная здесь коэффициентом задержки. Эта величина представляет особенный интерес, так как она используется в теории очередей, а также потому, что она обеспечивает измерение задержки, которое не зависит от флуктуаций времени пребывания машины на пункте. Коэффициент задержки равен средней задержке на машину, деленной на среднее время пребывания на пункте, или 0,77.

Как характеристику качества обслуживания пунктов сбора платы можно использовать процент задержанных машин. Эту величину можно получить делением числа случаев обслуживания на одном пункте двух или более машин на общее число зарегистрированных случаев занятости пунктов. Другой характеристикой является средняя задержка задержанных машин или средняя задержка всех машин, деленная на процент задержанных.

Максимальную задержку можно определить из максимальной длины очереди и среднего времени пребывания на пункте. В нашем примере проверкой данных было установлено, что максимальная отмеченная длина очереди равнялась шести машинам. Шестая машина ждала пять предыду-

щих, и, приняв, что время нахождения каждой машины на пункте равно среднему, определили максимальную задержку — 55,5 секунды.

Наличие свободной линии сбора платы также можно использовать для измерения качества обслуживания. Нетрудно установить, что такой фактор дополняет процент задержанных машин, так как любая машина может направиться как в занятую линию и быть задержанной, так и в свободную линию и не быть задержанной. Это было бы так, если бы водители всегда выбирали незанятую линию. К сожалению, водители часто выбирают занятую линию, даже если всегда имеется свободная, и этим могут задержать все машины. В нашем примере в 31 случае из 40 были свободны одна или больше линий, т. е. процент свободных линий составлял 77,5%. Увеличение процента задержанных машин за счет этого явления составляет 55%.

Наряду с ранее упомянутыми величинами, любая из которых может быть использована для определения качества обслуживания, интересно определить процентную занятость пунктов сбора платы, равную числу отмеченных случаев занятости, деленному на общее число наблюдений. В нашем примере она равна 63,3.

Эти расчеты были сделаны для трех линий сбора платы совместно. Используя расчетные величины, приведенные в нижней части таблицы 15.1, можно определить все параметры для каждой линии в отдельности.

Показав, каким образом из собранных данных можно определить ряд экспериментальных критериев обслуживания, мы в рамках этой главы будем интересоваться только теми, которые действительно использовались как показатель качества обслуживания. Прежде чем рассматривать их, произведем анализ входного потока.

АНАЛИЗ ПОТОКА АВТОМАШИН

Входной поток анализировался составлением плотности распределения числа машин, появляющихся в 30-секундные интервалы при разной интенсивности движения. По интенсивности движения наблюдения были разбиты на группы по 200 машин/час, и в каждой группе подсчитывалось и заносилось в таблицу число случаев появления 0, 1, 2, 3 и т. д. машин.

Эмпирическая частота появления каждой группы подсчитывалась как процент к общему числу интервалов. Построенные гистограммы частот показаны на рис. 15.1. Они имеют хорошее сходство с распределениями, которые можно было бы ожидать в случае чисто случайного потока. При этом можно отметить одну особенность: при более высоких интенсивностях движения наблюдается тенденция к «длинным хвостам» в правой части распределений. В частности, для самой большой интенсивности движения следовало бы отметить расширение распределения до 28 машин.

На рис. 15.2 показаны теоретические распределения для тех же интенсивностей движения. Сходство очевидно (рис. 15.1 и 15.2), хотя у теоретических распределений правый хвост значительно короче. При низких интенсивностях движения теоретические распределения представляют собой распределения Пуассона, при высоких — нормальные распределения.

Сравнение между эмпирическим и двумя теоретическими распределениями еще легче провести на рис. 15.3, где они вычерчены вместе. Эти распределения сняты в туннеле Линкольна при интенсивности движения 655 машин в час. Средняя интенсивность потока 5,46 машины за 30-секундный интервал, и среднее квадратическое отклонение — 2,73 машины за 30-секундный интервал. При вычислении параметров пуассоновского и нормального распределений использовалось выборочное среднее, но в случае

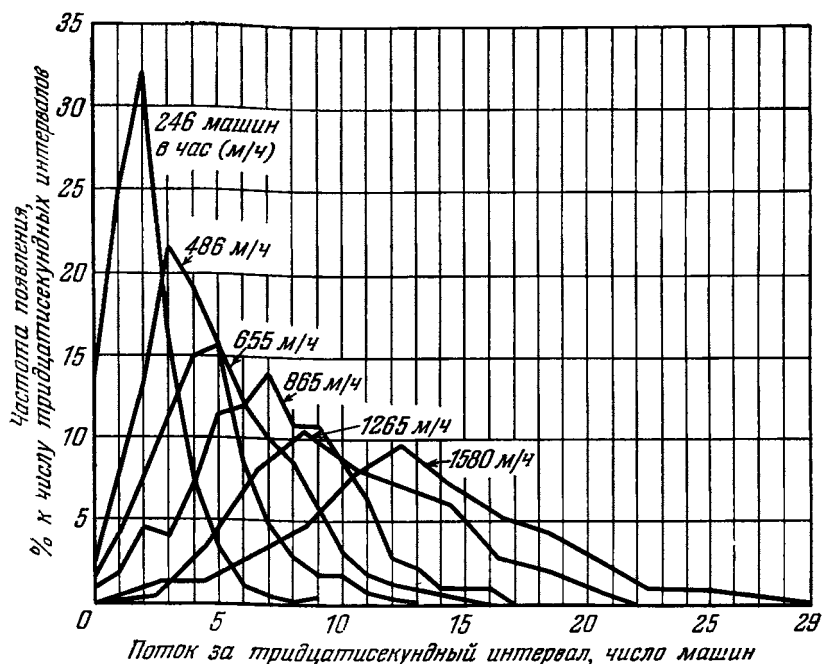


Рис. 15.1. Эмпирическая плотность распределения (полигон частот) потока машин.

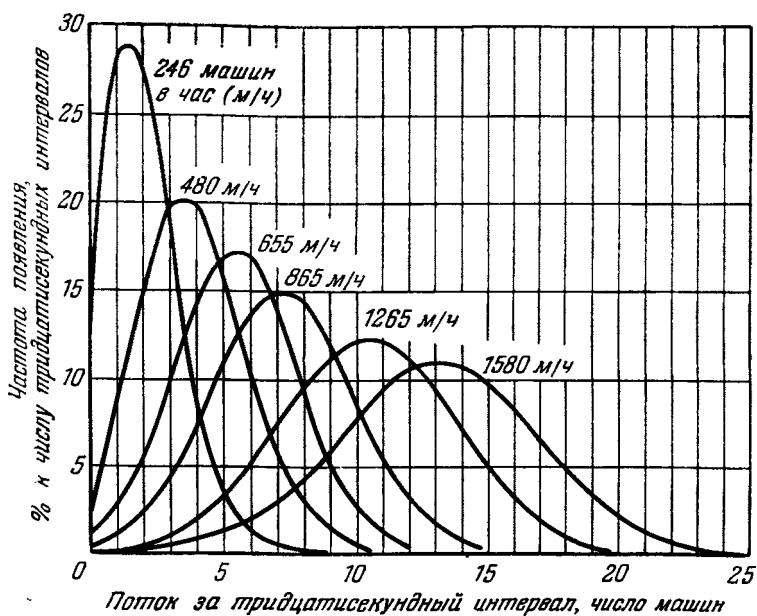


Рис. 15.2. Теоретическая плотность распределения потока машин.

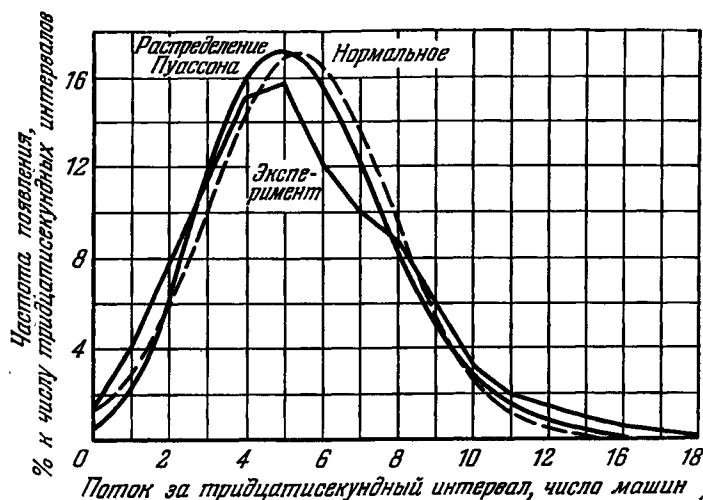


Рис. 15.3. Теоретическая и экспериментальная плотности распределения потока машин интенсивностью 655 машин в час в туннеле Линкольна.

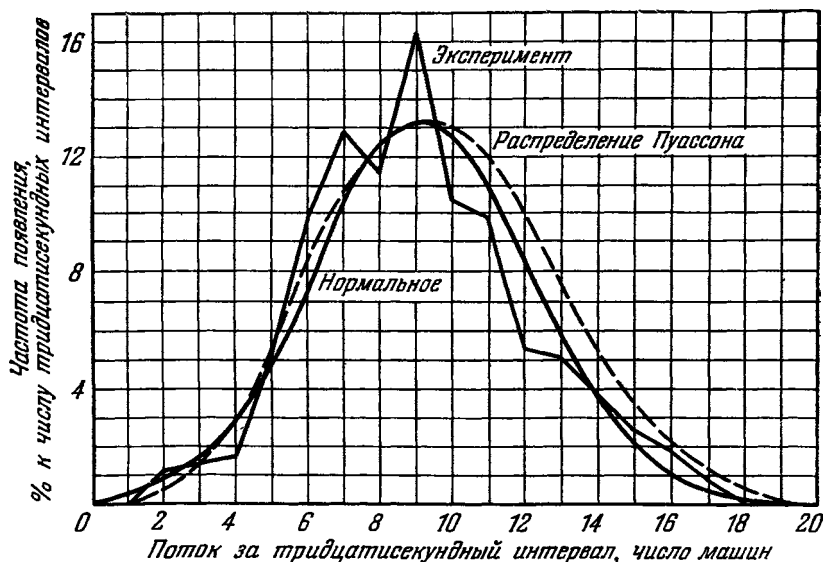


Рис. 15.4. Теоретическая и экспериментальная плотности распределения потока машин интенсивностью 1100 машин в час по мосту Дж. Вашингтона.

нормального распределения в качестве среднего квадратического отклонения принимали теоретическую величину для чисто случайного распределения. В этом примере распределение Пуассона, изображенное сплошной линией, по-видимому, дает лучшее сходство с эмпирическим, чем нормальное.

Таким же образом и с подобными же результатами анализировался поток машин на мосту Дж. Вашингтона. Соответствующие кривые для интенсивности движения 1100 машин в час со средним 9,17 и среднеквадратическим отклонением 3,00 изображены на рис. 15.4. Здесь, по-видимому, нормальная кривая, показанная сплошной линией, дает лучшее согласие, чем кривая распределения Пуассона *.

Таблица 15.2

Согласие распределения потока машин с теоретическими распределениями

Интенсивность потока (машин/час)	Распределение Пуассона	Нормальное распределение	Теоретическое распределение, дающее наилучшее согласие	Интенсивность потока (машин/час)	Распределение Пуассона	Нормальное распределение	Теоретическое распределение, дающее наилучшее согласие
246	0,754	0,235	Пуассона	1100	0,718	0,812	Нормальное
480	0,966	0,743	Пуассона	1265	0,359 *	0,295	Нормальное
655	0,738 *	0,459	Нормальное	1580	0,191	0,575	Нормальное
865	0,842	0,882	Нормальное				

Таблица 15.2 иллюстрирует несомненное согласие показателей для ряда последовательных интенсивностей движения. При высоких интенсивностях движения согласие явно ухудшается, хотя и остается удовлетвори-

* Распределение Пуассона характеризуется выражением

$$P(x) = e^{-m} m^x / x!,$$

где $P(x)$ — вероятность прибытия x машин за какой-либо отрезок времени при средней интенсивности прибытия m машин. Следует отметить, что распределение Пуассона полностью описывается одним параметром m . Нормальное распределение выражается следующим образом:

$$F(x) = \exp \left[-(x-m)^2 / 2s^2 \right] / s \sqrt{2\pi},$$

где $F(x)$ — вероятность прибытия x машин за какой-либо отрезок времени при средней интенсивности прибытия m и среднеквадратическом отклонении s . Для чисто случайного потока $s = \sqrt{npq}$, где p — вероятность появления машины за данный отрезок времени, q — вероятность неоявления машины и n — общее число машин в час.

В том случае, когда n велико и p мало, оба эти распределения являются хорошим приближением биномиального распределения, которое описывается выражением $P(x) = C_n^x p^x q^{n-x}$. При анализе в течение часа число появлений машин за 30-секундные интервалы $p = 1/120$, $q = 119/120$ и n — полная интенсивность потока. Однако нас больше интересуют нормальное и пуассоновское распределения, так как с ними легче иметь дело.

Чтобы узнать, какое из этих двух теоретических распределений дает лучшее согласие, можно использовать критерий χ^2 . Величина χ^2 описывается выражением

$$\chi^2 = \sum_{x=0}^n (f_0 - f_t)^2 / f_t,$$

где f_0 и f_t — эмпирическая и теоретическая частоты. Когда эти величины подсчитаны, надо найти в таблице значений χ^2 , чтобы определить степень вероятности принятой аппроксимации. Абсолютное согласие соответствует вероятности 1,00, но и согласие, соответствующее вероятности, большей 0,05, считается вполне удовлетворительным.

тельным, т. е. лучшим, чем 0,05. Однако это ухудшение представляет некоторый интерес, так как оно соответствует отмеченным ранее удлиненным правым хвостам эмпирических распределений. Как расширенные хвосты, так и меньшее согласие при более высокой интенсивности движения можно объяснить ростом скоплений, при которых действия одной машины служат препятствием для действий другой. Очевидно, что при еще более высоких интенсивностях согласие будет нарушаться, а при скоплениях машин бампер к бамперу функция распределения будет стремиться к постоянной величине. Интенсивность, при которой согласие становится неудовлетворительным, зависит, конечно, от числа линий на шоссе. Столбец, показывающий наиболее близкое теоретическое распределение, основывается на теории, что лучшим приближением к биномиальному распределению при среднем значении меньше 5 является распределение Пуассона, а при среднем больше 5 — нормальное. Это среднее значение соответствует интенсивности движения $120 \times 5 = 600$ машин в час. Два исключения, отмеченные звездочкой в таблице, не являются существенными. Результаты подтверждают уверенность, что действительное распределение при отсутствии перегрузок является биномиальным и, следовательно, чисто случайным распределением.

ЗАВИСИМОСТЬ ЗАНЯТОСТИ ОТ СТЕПЕНИ ЗАДЕРЖКИ

Считая установленным случайный характер потока, мы думали, что могли бы установить зависимость каждого из критериев обслуживания от интенсивности движения, а затем создать теорию задержки, которая достаточно хорошо согласовывалась бы с экспериментальными кривыми, и чтобы по крайней мере некоторые критерии вытекали из теории. К сожалению, этого не произошло; для некоторых коэффициентов задержки вследствие большого разброса точек нельзя было нарисовать удовлетворительные эмпирические кривые. Чтобы прямо по точкам можно было определить кривые корреляции для некоторых критериев обслуживания, потребовался бы очень большой объем информации.

В эту категорию попало наиболее естественное соотношение, которое пытаются установить, — соотношение между интенсивностями движения и средней задержкой в секунду. Причиной является то, что средняя задержка, измеренная в секундах, является функцией не только интенсивности движения, но и времени обслуживания. Вследствие различий в составе потока на разных пунктах сбора платы время обслуживания неодинаково и данные, полученные на одном пункте, нельзя сравнить с данными, полученными на другом. Другая причина состоит в том, что время обслуживания отчасти зависит от сборщиков платы, которые в некоторых случаях знали, что за ними наблюдают и, естественно, старались сделать это время более коротким. Из-за этих причин нельзя было непосредственно вычертить среднюю задержку для каждого пункта.

Чтобы преодолеть эту трудность, мы обратились к кривым зависимости занятости от степени задержки. Эти кривые не зависят от времени обслуживания, следовательно, можно использовать данные разных пунктов. Разброс точек заметно уменьшался, и при еще большем количестве наблюдений, полученных при суммировании данных со всех пунктов, можно вычертить удовлетворительные эмпирические кривые. Далее теория очердедей развивается на основе единиц времени обслуживания, и было бы желательно сравнить эмпирические результаты с теориями Эрланга [2], Молина [1] и теорией Поллачека и Кроммелина [3].

Уравнение Эрланга имеет вид

$$d = \frac{(y^x/x!) [x/(x-y)^2]}{\{1 + y + (y^2/2!) + (y^3/3!) + \dots + [y^{x-1}/(x-1)!] + (y^x/x) [x/(x-y)]\}},$$

где d — средняя задержка в единицах времени обслуживания, x — число каналов движения, y — интенсивность потока в эрлангах.

Эрланг определяется как средняя занятость за время T , деленная на T . Это безразмерная единица, подобная децибелу, который использовался для выражения величины затухания. Например, если каждый из трех каналов занят половину времени T , общая занятость равна $1,5T$ и интенсивность потока составляет $1,5$ эрланга. Число эрлангов также выражает среднее число элементов потока, поступающего одновременно. Уравнение задержки по Эрлангу основывается на предположении об экспоненциальном распределении времени обслуживания, т. е. $P(t) = e^{-t/h}$ — вероятность того, что время обслуживания случайно выбранной заявки больше или равно t , если среднее время обслуживания равно h .

В уравнении Молина вводится поправочный множитель к уравнению Эрланга, который позволяет применять его при постоянном времени обслуживания. Поправочный множитель дается выражением $[x/(x+1)] [1 - (y/x)^{x+1}] / [1 - (y/x)^x]$. Уравнение Поллачека — Кроммелина, основанное на предположении постоянства времени обслуживания [3], имеет вид

$$d = \sum_{w=1}^{\infty} e^{-wy} \left(\sum_{u=wx}^{\infty} \frac{(wy)^u}{u!} - \frac{x}{y} \sum_{u=wx+1}^{\infty} \frac{(wy)^u}{u!} \right).$$

На рис. 15.5 проведено сравнение между величинами, полученными

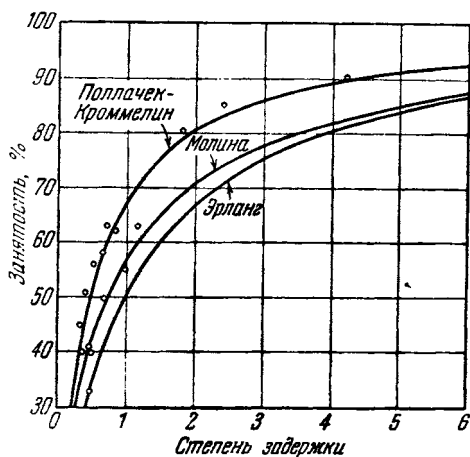


Рис. 15.5. Теоретические кривые занятость — задержка и экспериментальные точки для одного пункта сбора платы.

из этих уравнений, и экспериментальными точками для одного пункта сбора платы. Как и предполагалось, уравнение Поллачека — Кроммелина дает величины, близкие к экспериментальным, в то время как уравнения Молина и Эрланга дают значительно большие задержки. Это свидетельствует о постоянстве времени обслуживания на пунктах и подтверждает, что уравнение Поллачека — Кроммелина точнее описывает среднюю задержку при более высоких нагрузках, чем уравнение Молина, хотя при низких нагрузках между ними нет заметного различия. Определение распределения времени обслуживания посредством хронометрирования также показало, что оно ближе к постоянному, чем к экспоненциальному.

Для четырех пунктов сбора платы (см. рис. 15.6) экспериментальные задержки оказались значительно больше теоретических, причем уравнение Эрланга ближе к экспериментальным результатам, чем уравнение для постоянного времени обслуживания. Причина была отмечена ранее: машины выстраиваются перед одной линией, в то время как другая свободна. Так как было показано, что поток является случайным, а уравнение Поллачека — Кроммелина точно описывает задержку для одного пункта, именно

плохое распределение потока между пунктами является фактически единственной причиной значительного увеличения задержек. Наши попытки учесть это обстоятельство в уравнении не были успешными, поэтому пришлось перейти к экспериментальным значениям.

Следует отметить, что одна эмпирическая кривая построена для четырех левосторонних пунктов, а другая — для трех левосторонних и одного правостороннего. Левосторонними называются пункты сбора платы, расположенные со стороны водителя машины, проходящей линию, а правосторонними — расположенные на противоположной стороне. Обе кривые приведены для того, чтобы показать более низкую эффективность правосторонних пунктов.

Еще более ясно это иллюстрируется в таблице 15.3, которая показывает процентное возрастание степени задержки при равной занятости и уменьшение занятости для равных задержек, когда левосторонний пункт заменяется правосторонним. Прежде всего, следует отметить, что это приводит к увеличению задержки приблизительно на 50%. Задержка увеличивается для всего потока, а не только для машин, имеющих дело с правосторонними пунктами.

Далее, следует отметить, что при умеренных уровнях задержки занятость правостороннего пункта составляет меньше половины занятости левостороннего пункта. По мере увеличения скопления машин на площадке занятость правостороннего пункта растет. Таким образом, особенность правостороннего пункта

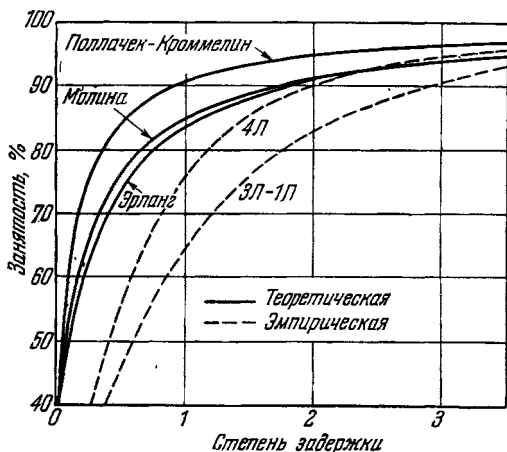


Рис. 15.6. Теоретические и действительные кривые занятости — задержка для четырех пунктов сбора платы.

Таблица 15.3

Сравнение четырех левосторонних пунктов с тремя левосторонними (Л) и одним правосторонним (П)

Процентная занятость для 4 левосторонних пунктов (4Л)	Соответствующая величина для 3Л-1П	Степень задержки для 4Л	Соответствующая степень задержки для 3Л-1П	Увеличение степени задержки для 3Л-1П (в %)	Занятость П по отношению к занятости Л (в %)
1	2	3	4	5	6
50	40	0,40	0,60	50	20
60	50	0,60	0,85	41	33
70	60	0,85	1,25	47	43
80	70	1,25	1,80	44	50
90	83	2,00	3,00	50	60

состоит в том, что он работает при перегрузке. Поэтому Управление порта перестраивает основные площадки таким образом, чтобы все пункты сбора платы были левосторонними.

ЗАВИСИМОСТЬ ВРЕМЕНИ ОБСЛУЖИВАНИЯ ОТ ПОТОКА

Чтобы преобразовать безразмерные степени занятости и степени задержки в практические единицы (*машин/час* и *секунды*), необходимо определить величину времени обслуживания. В некоторых задачах теории задержки транспортные перегрузки и число используемых каналов не влияют на время обслуживания, например в случае потока телефонных вызовов. Но в случае операций по сбору платы время обслуживания является функцией интенсивности потока и числа используемых пунктов. Из рис. 15.7 для пункта, расположенного у моста Дж. Вашингтона, можно видеть, что время обслуживания заметно больше при низкой интенсивности потока линии, чем при высокой. Когда поток на линию приближается к нулю,

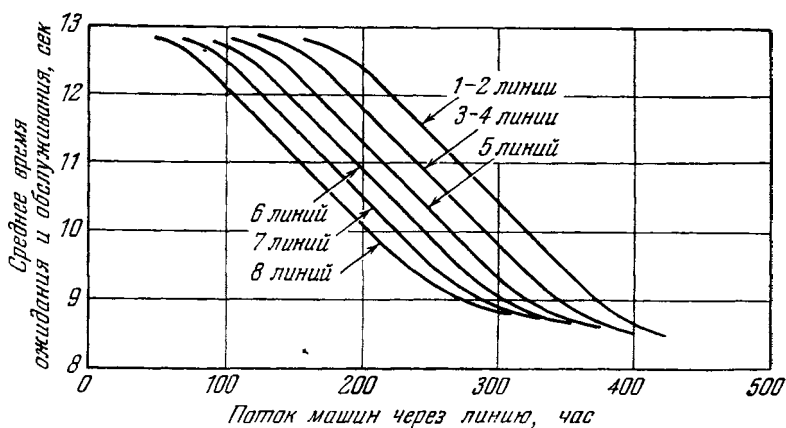


Рис. 15.7. Среднее время задержки и обслуживания на мосту Дж. Вашингтона.

время обслуживания приближается к максимальному значению, равному приблизительно 13 секунд, и когда поток увеличивается до максимального, который может пропустить линия, время обслуживания достигает минимальной величины, равной 8 секундам. Из рисунка можно также увидеть, что при большем числе линий сбора платы время обслуживания спадает быстрее. Однако когда оно начинает спадать, то спадает одинаково для всех кривых, т. е. пропорционально увеличению потока на линию, причем наклон равен приблизительно 1 секунде на 50 машин на линию в час.

Нетрудно найти объяснение этому явлению. Прежде всего, при увеличении потока на линию время обслуживания уменьшается потому, что и сборщики и клиенты стремятся ускорить действия под давлением очереди. Видимо, это является абсолютно общим явлением для очередей с людьми, которые сознают значение пробок. Нами было замечено, что при слабом потоке происходило значительно больше разговоров между сборщиком и клиентом, чем при сильном потоке.

Следующая причина заключается в том, что когда перед пунктом очередь, клиенты имеют возможность приготовить свою плату, тогда как при свободной линии клиент может подъехать прямо к пункту прежде, чем достанет плату, и тогда вынужден искать ее.

По-видимому, более быстрый спад времени обслуживания при большем числе линий объясняется неравномерным распределением потока между

свободными линиями. Клиенты, в частности, предпочитают линии с левосторонними пунктами и расположенные около центра площадки. Большая интенсивность потока в этих популярных линиях приводит к уменьшению времени обслуживания даже в тех случаях, когда средний по всем линиям поток еще мал. Так как эти популярные линии имеют дело с наибольшим потоком, они оказывают относительно большее влияние на средний по всем линиям поток, чем менее популярные.

При построении кривых зависимости времени обслуживания от интенсивности потока на линию было обнаружено, что при высоких интенсивностях мало экспериментальных значений времени обслуживания и трудно четко определить положение кривых. Дело в том, что для получения точек при сильной нагрузке на линию приходится искусственно создавать большие перегрузки. Это может вызвать недовольство, на которое трудно возразить. Поэтому мы искали другие методы.

Основной использованный метод состоял в хронометрировании времени операций по сбору платы и подсчете времени задержки машин для различных типов операций. Время обслуживания состоит из двух частей. Первая — это время, затраченное сборщиком на то, чтобы получить плату от клиента и, если необходимо, дать сдачу или квитанцию. Вторая — время, затраченное машиной на то, чтобы стать в положение для уплаты. За время сбора или операции принимается интервал между моментом, когда колеса машины, занимающей линию, останавливаются, и моментом, когда они снова начинают вращаться и машина освобождает линию. За время задержки машины берется интервал между моментом начала движения одной и моментом остановки следующей в положении для уплаты.

Используя такое деление времени обслуживания, относительно легко было произвести хронометрирование минимального времени операции, отмечая, когда в линиях, имеющих длинную очередь, колеса машин останавливаются и снова трогаются. Подобным же образом легко было бы замерить и время машины, но это не сочли необходимым, так как легко получить информацию об ускорении и торможении машин и определить время задержки машин из имеющихся кривых.

Наблюдения времени операций, проведенные на нескольких сотнях машин, и определение времени задержки машин из кривых торможения—ускорения установили зависимость минимального времени обслуживания от типа машин и типа пункта сбора платы. Это показано в таблице 15.4, по данным

Таблица 15.4

Распределение среднего минимального времени обслуживания для разных машин

Тип машин	Время подачи машины на место сбора (сек)	Время обслуживания (сбора пошлины) (сек)		Полное время обслуживания машины (сек)	
		левосторонний пункт	правосторонний пункт	левосторонний пункт	правосторонний пункт
Легковая	5,0	3	4	8	9
Автобус	6,5	3	4	9,5	10,5
Грузовая	6,0	5	6,5	11,0	12,5
Грузовик с прицепом	7,5	6,5	8,0	14,0	15,5

которой можно подсчитать минимальное время обслуживания для потока, состоящего из различного процентного состава легковых машин, автобусов, грузовиков и тракторов. Например, поток в туннеле Линкольна в часы пик состоял из 64% легковых, 15% автобусов, 14% грузовиков и 7% тракторов. Было подсчитано, что минимальное время обслуживания для левосторонних и правосторонних пунктов равно:

левосторонние пункты

$$\text{Полное время обслуживания} = 0,64 \times 8 + 0,15 \times 9,5 + 0,14 \times 11 + 0,07 \times 14 = 5,1 + 1,4 + 1,5 + 1,0 = 9 \text{ сек.}$$

$$\text{Максимальная пропускная способность пункта} = 3600/9 = 400 \text{ машин/час.}$$

правосторонние пункты

$$\text{Полное время обслуживания} = 0,64 \times 9 + 0,15 \times 10,5 + 0,14 \times 12,5 + 0,07 \times 15,5 = 5,8 + 1,6 + 1,8 + 1,1 = 10,3 \text{ сек.}$$

$$\text{Максимальная пропускная способность пункта} = 3600/10,3 = 350 \text{ машин/час.}$$

Для контроля был использован другой метод. На мосту Дж. Вашингтона за 18 периодов пик, в течение которых происходили сильные перегрузки из-за чрезвычайно перегруженных пунктов сбора платы, средний поток на линию составлял 403 машины. Предполагая в эти часы 95%-ную загруженность, определили минимальное время обслуживания как $0,95 \times (3600/403) = 8,5$. Это точно совпадает с временем задержки машины и времени сбора платы для равного числа лево- и правосторонних пунктов и потока, состоящего исключительно из легковых машин, который в действительности был на мосту Дж. Вашингтона в указанных случаях. Данный метод применялся только на мосту, потому что у входа в туннель в часы пик перегрузки, вызванные самим туннелем, препятствуют движению машин от пунктов, когда сделка уже закончена, тем самым искусственно удлинняя время обслуживания. Если пункты сбора платы располагаются у выхода из туннеля, туннель сдерживает поток, препятствуя насыщению пунктов.

ПОСТРОЕНИЕ КРИВЫХ СРЕДНЕЙ ЗАДЕРЖКИ

При условии, что, кроме отношения процента загруженности к коэффициенту задержки, найдена также зависимость времени обслуживания от интенсивности движения на линии, является целесообразным (с точки зрения практики) получить зависимость найденной ранее средней задержки от интенсивности движения. В таблице 15.5 приведен пример расчета точек кривой такой зависимости для четырех левосторонних пунктов сбора платы. При расчете использовались значения двух указанных выше соотношений. Таблица 15.5 относится только к мосту Дж. Вашингтона, для него же приведены значения времени обслуживания в третьем столбце.

При расчете первоначально задавались интенсивностью движения по каждой линии (1-й столбец). Пусть, например, объем движения составляет 300 машин в час на одной линии. В следующем столбце указан общий объем движения на четырех линиях (1200 машин в час). В 3-м столбце указано время обслуживания у пункта сбора платы 9,8 сек, найденное из зависимости времени обслуживания от интенсивности движения. В 4-м столбце приведено время обслуживания на линию, полученное умножением времени

Таблица 15.5

**Зависимость задержки от интенсивности движения
для четырех левосторонних пунктов**

Интенсивность движения (машин/час) по одной линии	Общее количество машин/час	Полное время обслуживания (сек)	Полное время обслуживания на одном пункте	Процент занятости	Коэффициент задержки	Задержка (сек)
100	400	12,9	1290	36,0	0,20	2,6
150	600	12,7	1910	53,0	0,45	5,7
200	800	11,8	2360	65,5	0,73	8,6
250	1000	10,8	2700	75,0	1,02	11,0
300	1200	9,8	2940	81,6	1,31	12,8
350	1400	8,9	3120	86,7	1,66	14,8
375	1500	8,7	3260	90,6	2,00	17,4
385	1540	8,6	3310	91,9	2,36	20,3
400	1600	8,5	3400	94,4	3,40	28,9

обслуживания машины на интенсивность движения по одной линии. Деля последнее значение на число секунд в часе (3600), получим 81,6% загрузки (5-й столбец). Используя зависимость процента загрузки от коэффициента

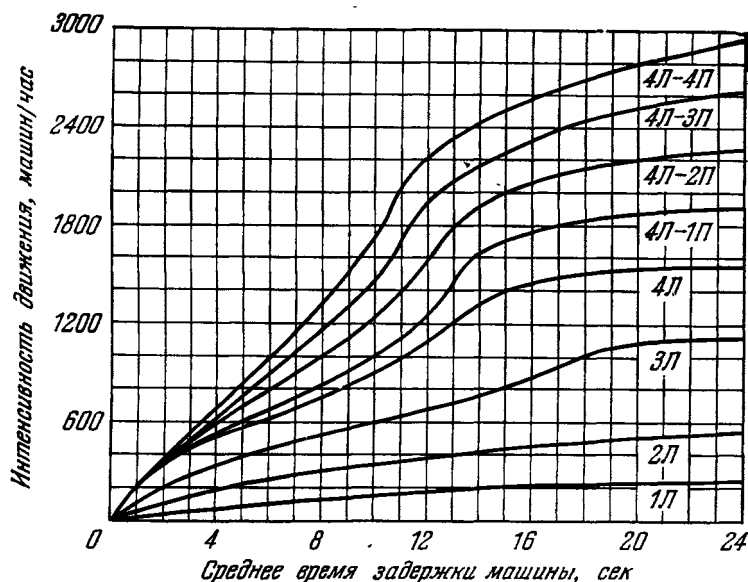


Рис. 15.8. Средняя задержка при различной интенсивности движения через мост Дж. Вашингтона.

задержки для четырех левосторонних пунктов, получим в 6-м столбце коэффициент задержки, равный 1,31. Умножая этот коэффициент на время обслуживания, указанное в 3-м столбце, получим среднее время задержки, равное 12,8 сек.

При помощи данного метода были рассчитаны приведенные на рис. 15.8 кривые для различных комбинаций пунктов у моста Дж. Вашингтона. Из

анализа этих кривых следует, что при заданном времени задержки пропускная способность разных пунктов сбора платы значительно меняется в зависимости от их комбинаций. До проведения этих исследований администрация, составлявшая расписание работы сборщиков, считала, что в любом случае все пункты одинаковы. Рис. 15.8 позволяет также выявить причины перегрузки правосторонних пунктов сбора платы. Кривые для комбинаций четырех левосторонних пунктов с одним, двумя, тремя и четырьмя правосторонними пунктами сливаются с прямой, построенной для четырех левосторонних пунктов при интенсивности движения около 400 машин в час. Ниже этой интенсивности правосторонние пункты не загружены.

Таким образом, получено решение проблемы задержек, но было бы желательным оценить его точность. Чтобы убедиться, что полученные кривые действительно отражают исходные наблюдения, было проведено сравнение значений, найденных по графикам на рис. 15.8, с результатами непосредственных вычислений средних задержек по имеющимся данным. Было найдено, что при периоде наблюдения около 20 мин средняя ошибка составляет 2,64 сек. Считая, что непосредственные вычисления на отрезке времени 20 мин дают нам среднее значение, находим, что вероятная ошибка при периоде 1 час будет меньше в $1/\sqrt{3}$ раза и составит 1,53 сек. Средняя величина задержки за время наблюдений составила 11 сек. Следовательно, по полученным кривым можно предсказывать среднее время задержки с вероятной ошибкой около 15%. К счастью, это вполне приемлемо для наших целей. В противном случае было бы необходимо использовать какой-либо иной критерий для предсказания уровня обслуживания и числа пунктов сбора платы, требующихся для различной интенсивности движения транспорта.

АНАЛИЗ ТРАНСПОРТНЫХ ОЧЕРЕДЕЙ

Очень часто при анализе очередей знания средней задержки недостаточно. Исследователю необходима также информация о максимально возможных задержках. Если в срочной поездке водитель задерживается на несколько минут в длинном хвосте машин у пункта сбора платы, он, вероятно, мало утешится сообщением, что при регулярном проезде по этой дороге среднее время задержки будет весьма небольшим. Данное обстоятельство приводит к необходимости исследования длины очереди.

Простой способ анализа этой величины состоит в построении графика зависимости наибольшей очереди от интенсивности движения для каждой из комбинаций используемых линий. При анализе полученных графиков возникает проблема большого рассеивания данных, причем разброс их больше, чем в рассмотренных ранее соотношениях. Из ограниченного количества имеющихся сведений можно сделать лишь приблизительные выводы о том, какова возможная максимальная очередь и как часто она возникает при известных интенсивностях движения и комбинации пунктов сбора платы. Для нахождения этих характеристик необходимо использовать методы математической статистики.

При обработке данных таблицы 15.1 с целью статистического анализа очередей было решено рассматривать не общее количество машин у всех открытых пунктов, а только число машин в наиболее длинной очереди. Действительно, наиболее логично рассмотреть данную ситуацию с точки зрения водителя, ожидающего максимальное время. Это время не характеризуется общим числом ожидающих машин, так как распределение машин между пунктами неравномерное.

Первый этап анализа аналогичен анализу потока автомашин. Однако разница заключается в значительно меньшем количестве машин в выборке. При анализе потока автомашин данные группировались в партии по интенсивности движения. При такой группировке можно получить выборки, состоящие из нескольких сотен интервалов. Частотное распределение, построенное на основе этих выборок, будет довольно плавным. При исследовании потока автомашин данные должны, однако, приниматься во внимание при анализе очередей. На практике работа пунктов изменяется два или большее число раз в час. Причина состоит в изменении интенсивности

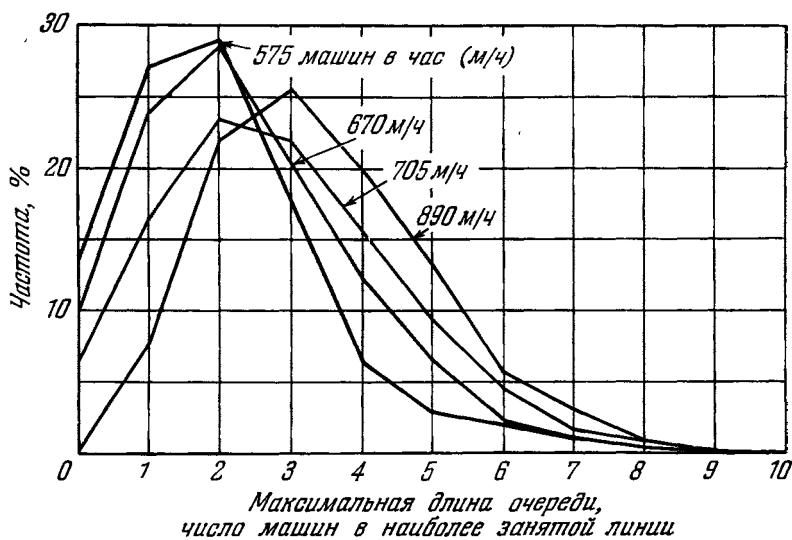


Рис. 15.9. Эмпирические плотности распределения длины очереди для трех левосторонних пунктов.

движения и в смене сборщиков, что приводит иногда к изменению типа пункта сбора платы. Следовательно, при анализе очередей целесообразно использовать интервал 20 мин, как и при определении среднего времени задержки. Это дает выборку с 40 интервалами (2 в минуту). Для сглаживания частотного распределения, построенного по малому числу замеров, можно использовать найденное по трем точкам скользящее среднее взвешенное значение.

На рис. 15.9 приведены частотные распределения максимального числа автомашин в очереди для трех левосторонних пунктов. Эти кривые получены после сглаживания путем усреднения и преобразования масштаба по вертикали (на рис. 15.9 указаны проценты от общего числа наблюдений). Первые две кривые (для интенсивности движения 575 и 670 машин в час) относятся к туннелю Линкольна, а другие две (интенсивность движения 705 и 890 машин в час) — к мосту Дж. Вашингтона. Следует указать, что, как и следовало ожидать, полученные распределения напоминают распределение для потока автомобилей. Действительно, время задержки автомобилей в основном постоянно, и, следовательно вариации количества автомобилей в очереди обусловлены вариациями в потоке автомашин.

На рис. 15.10 показаны распределения Пуассона, которые очень близки к соответствующим эмпирическим распределениям на рис. 15.9. При построении распределений Пуассона использовались средние значения,

найденные из эмпирических распределений. Напротив, при анализе потока автомашин не было необходимости в экспериментальном определении среднего значения, так как оно находилось непосредственно по интенсивности движения и интервалу наблюдения. Нет сомнения в существовании определенной зависимости между интенсивностью движения и средней длиной очереди для заданной комбинации пунктов. Однако мы не можем определить ни теоретической, ни эмпирической формулы для предсказания по интенсивности движения среднего числа машин в очереди.

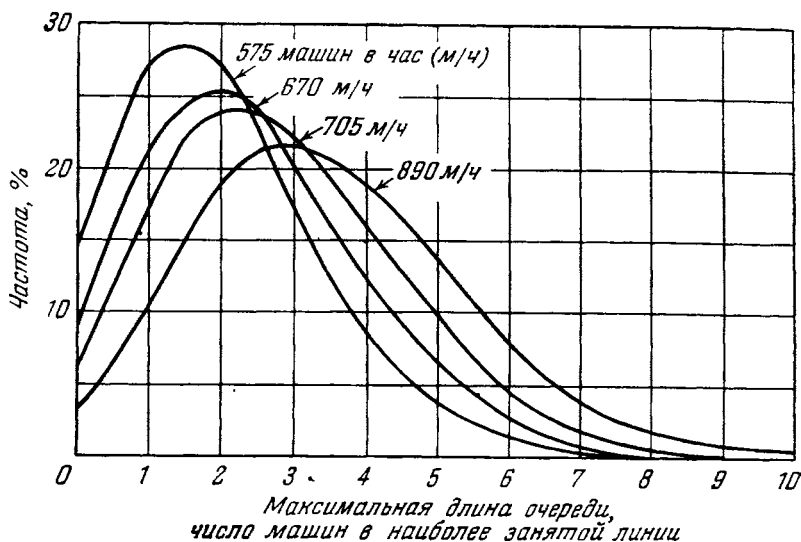


Рис. 15.10. Теоретические плотности распределения длины очереди для трех левосторонних пунктов.

Совпадение эмпирического и пуассоновского распределений иллюстрируется рис. 15.11, где приведены оба распределения. Графики относятся к движению через туннель Линкольна (615 машин в час; три левосторонних пункта). Между прочим, эти значения близки к данным выборки, (см. табл. 15.1). Среднее значение числа машин в очереди равно 2,16, а среднее квадратическое отклонение составляет 1,52. Среднее квадратическое отклонение от среднего значения, которое будет использоваться далее, составляет 0,15. Для нормального распределения эта мера составляет только 0,01.

Рис. 15.12 показывает, что результаты, полученные для моста Дж. Вашингтона, сравнимы с результатами для туннеля Линкольна. При немного большей интенсивности движения (705 машин/час) среднее значение количества машин в очереди составляет 2,79, среднее квадратическое отклонение равно 1,67, а среднее квадратическое отклонение от среднего значения — 0,31. Вероятность принятой аппроксимации, полученная из χ^2 (которая используется как приближенная характеристика близости распределений) равна 0,55 для распределения Пуассона и не превышает 0,01 для нормального распределения.

Графики распределений числа машин в очереди для туннеля и моста весьма близки, так же как близки соответствующие графики для потока автомашин. Совпадение так велико, что мы не могли различить, какой из них относится к туннелю, а какой — к мосту. Это было удивительным, так как оба случая заметно отличаются по времени задержки и составу транс-

порта. Мы потратили много усилий, пытаюсь найти разницу в распределении числа машин в очереди, но безуспешно. Было решено что всегда, кроме условий, близких к насыщению, большее время задержки (при той же интенсивности движения) сказывается скорее на времени ожидания, а не на числе машин в очереди. Например, среднее значение числа машин в очереди при интенсивности движения 615 машин в час и трех левосторонних пунктах равно 2,16 и время задержки составляет 11,1 сек. Это соответствует времени ожидания в очереди $2,16 \times 11,1 = 24$ сек. Если время задержки увеличится на 20%, т. е. 13,3 сек, то время ожидания тоже возрастет до 28,8 сек, т. е. на 20%. Между тем среднее число

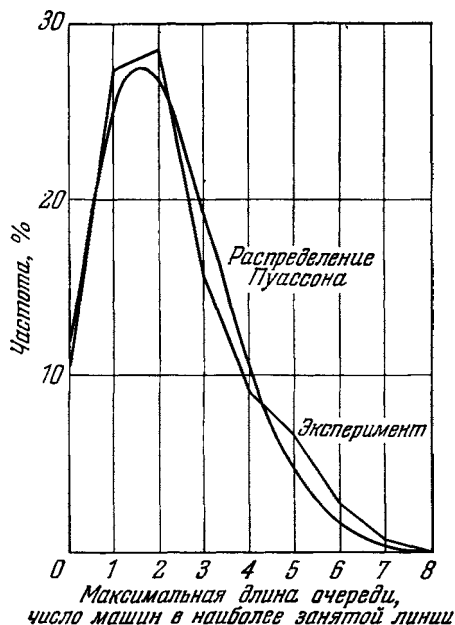


Рис. 15.11. Эмпирическая и теоретическая плотности распределения длины очереди при интенсивности движения 615 машин/час через три левосторонних пункта у туннеля Линкольна.

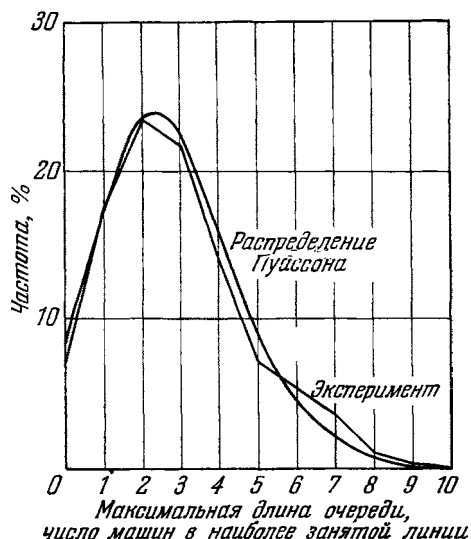


Рис. 15.12. Эмпирическая и теоретическая плотности распределения максимальной длины очереди при интенсивности движения 705 машин/час через три левосторонних пункта у моста Дж. Вашингтона.

машин в очереди останется прежним — 2,16. Вероятно, подобные явления должны происходить и при малых отклонениях во времени задержки.

Во всех рассмотренных случаях эмпирическое распределение ближе к распределению Пуассона, чем к нормальному распределению. Это позволяет считать, что истинный характер распределения числа автомашин в очереди хорошо отражается распределением Пуассона. Но данное утверждение справедливо лишь при определенных условиях. Таблица 15.6 пока-

Таблица 15. 6

Согласие распределения длины очереди с пуассоновским
для трех левосторонних пунктов сбора пошлины

Интенсивность движения	575	615	625	670	705	750	867	890
Степень согласия	0,93	0,64	0,55	0,85	0,55	0,05	0,01	0,32

зывает, что расхождение увеличивается при возрастании интенсивности движения. При интенсивности движения 575 машин в час имеет место довольно хорошее совпадение — 0,93. При большей интенсивности согласие значительно падает, а при интенсивности 800 машин в час согласие кривых совершенно неудовлетворительно. Цифра 800 машин в час получена для трех левосторонних пунктов, но и для других комбинаций пунктов наблюдалось аналогичное ухудшение совпадения, когда интенсивность движения составляла 60—75% от насыщения. Кажется, что причина ухудшения состоит в увеличении числа машин, проходящих в промежуток между интервалами наблюдения, при приближении к насыщению. Интенсивность потока, при которой наблюдается значительное расхождение между истинным и пуассоновским распределением, называется «точкой Пуассона».

ЗАВИСИМОСТЬ СРЕДНЕЙ ДЛИНЫ ОЧЕРЕДИ ОТ ИНТЕНСИВНОСТИ ДВИЖЕНИЯ

После того, как установлено, до какой интенсивности движения эмпирическое распределение совпадает с распределением Пуассона, необходимо найти зависимость средней длины очереди от интенсивности движения.

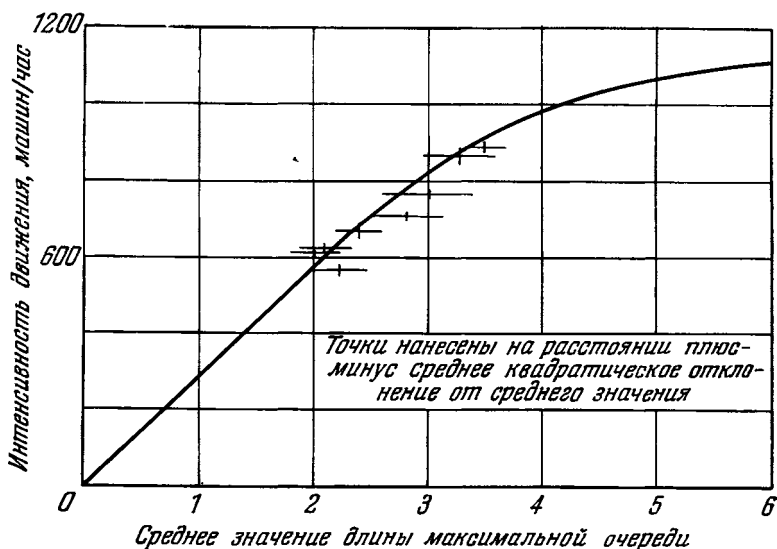


Рис. 15.13. Среднее значение длины максимальной очереди.

Известно, что распределение Пуассона определяется только одним параметром — средним значением. Единственный удовлетворительный метод нахождения средних значений состоял в построении экспериментальной зависимости, показанной на рис. 15.13. Для облегчения построения кривой были нанесены точки на расстоянии плюс — минус среднее квадратическое отклонение от среднего значения. В данном случае, как и во многих других, большинство точек было более или менее сгруппировано в диапазоне интенсивности движения, преобразованной обычно для соответствующей комбинации пунктов сбора платы. Получение экспериментальных данных при высокой интенсивности движения могло бы потребовать искусственной перегрузки пунктов, что доставило бы некоторым водителям немало неприятностей. К счастью, в этих данных нет необходимости, так как ясно, что

кривая асимптотически приближается к полной загруженности для соответствующей комбинации пунктов.

Известно, что полная загруженность для четырех левосторонних пунктов составляет около 400 машин в час у туннеля Линкольна и 450 машин в час у моста Дж. Вашингтона. Из-за безопасности опыты проводились при меньшей интенсивности движения. На графике (рис. 15.13) кривая,

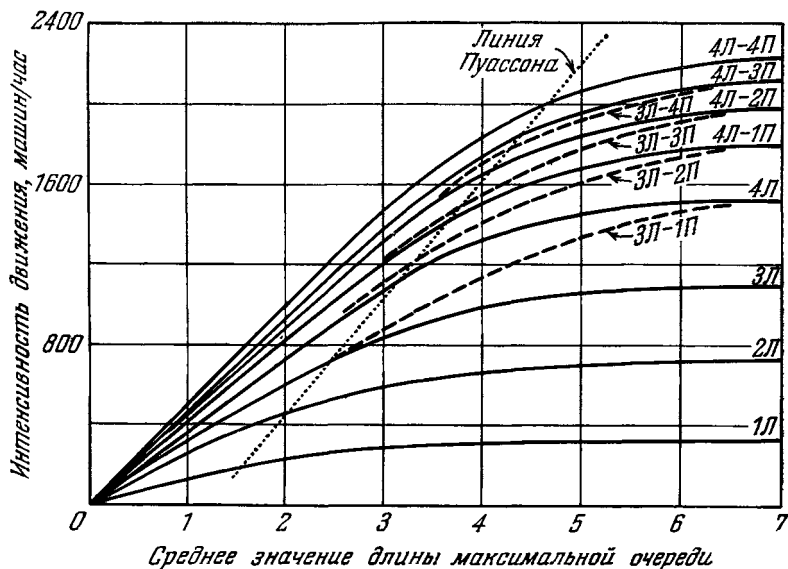


Рис. 15.14. Среднее значение длины максимальной очереди.

построенная для трех левосторонних пунктов, приближается к интенсивности 1200 машин в час.

На рис. 15.14 построено семейство подобных характеристик для различных комбинаций пунктов. На этих характеристиках были нанесены точки Пуассона. Оказалось, что точки Пуассона располагаются очень близко к прямой линии, изображенной на рис. 15.14 пунктиром и названной линией Пуассона.

МАКСИМАЛЬНО ВОЗМОЖНАЯ ДЛИНА ОЧЕРЕДИ

Знание средних значений длины наибольшей очереди, распределенной по закону Пуассона, позволяет исследовать граничные значения длины очереди. Возникает вопрос: в каких граничных значениях мы заинтересованы? Этот же вопрос можно сформулировать в терминах телефонной связи: какой надо принять вероятность потерь?

Ответ на этот вопрос зависит в некоторой степени от следующих рассуждений. Если вероятность потерь выбрана слишком большой, например 0,1, то граничное значение будет превышаться очень часто и будет плохой мерой максимальной задержки, от которой страдают клиенты. С другой стороны, если вероятность потерь слишком мала, например 0,001, то граничное значение встречается очень редко и мы обманываемся уже в другом направлении.

Рассмотрим этот вопрос с несколько иной точки зрения. Пусть интересующий нас отрезок времени — 1 час — состоит из 120 тридцатисекундных интервалов. Пусть максимальное значение длины очереди выбрано таким, что оно может быть достигнуто или превышено с вероятностью 0,1. Тогда ожидаемое число интервалов, на которых это произойдет, составляет $0,1 \times 120 = 12$. Если вероятность потерь взята равной 0,001, то ожидаемое число аналогичных интервалов равно 0,12, т. е. можно ожидать, что максимальная очередь будет встречаться раз в 8 часов. Следовательно, кажется логичным выбрать вероятность потерь, равную 0,01, так как при этом можно ожидать достижения граничного значения лишь на одном тридцатисекундном интервале в течение часа. Можно сказать, что вероятность потерь, равная 0,01, определяет максимальную длину возможной очереди.

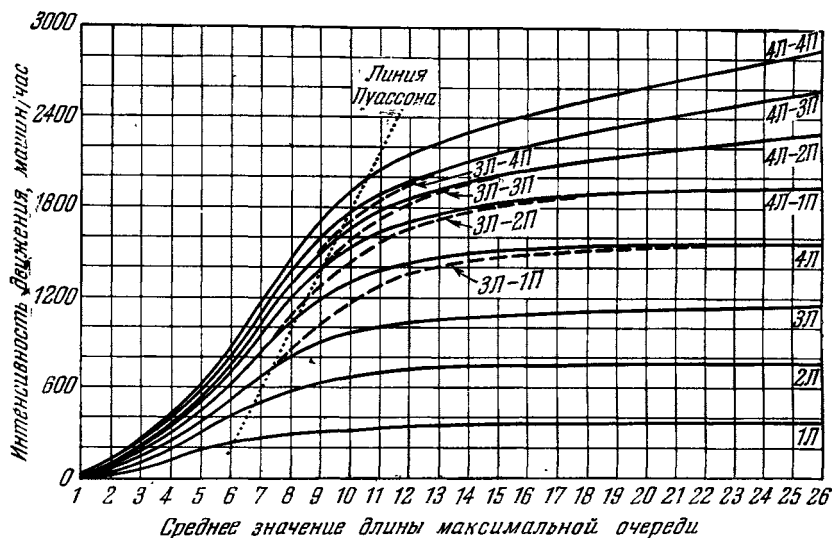


Рис. 15.15. Вероятная максимальная длина очереди.

На рис. 15.15 приведена зависимость максимальной длины возможной очереди от интенсивности движения при вероятности потерь, равной 0,01. Эти кривые были построены по имеющимся данным и далее продолжены путем экстраполяции до интенсивности движения, асимптотически приближающейся к значению, равному числу пунктов, помноженному на 400.

Так же, как при построении кривых средней задержки, сочли целесообразным проверить надежность полученных результатов путем сравнения их с данными непосредственных наблюдений. Было взято 53 периода наблюдений на отрезке времени 20 часов для 1—8 пунктов у туннеля Линкольна и моста Дж. Вашингтона. За это время на 26 интервалах из 2379 действительная максимальная очередь была равна (или превышала) значению максимально возможной очереди на рис. 15.15. Следовательно, вероятность достижения значений, указанных на кривых (рис. 15.15), составляла $26/2379 = 0,0109$, т. е. отличалась лишь на 9% от значения, соответствующего вероятности потерь, равной 0,01. Кроме этих 53 периодов (каждый из которых длился 20 мин) было 10 периодов, на которых длина очереди на одном или нескольких 30-секундных интервалах превышала предсказанное значение. Превышение составляло в среднем 1,4 машины, а максимум превышения был равен 4 машинам для 41 интервала наблюдения, действительный максимум был меньше предсказанного в среднем на 1,5 машины.

ПРОБЛЕМА ОПТИМИЗАЦИИ

После исследования проблемы очередей, нахождения средней задержки и максимальной длины очереди необходимо установить оптимальный уровень обслуживания, или стандарт обслуживания. Возможный способ решения этой задачи состоит в том, чтобы определить верхний уровень средней задержки (например, 20 сек) и при превышении данного уровня открывать дополнительный пункт сбора платы. Такое компромиссное решение вряд ли можно обосновать, и его нецелесообразно рекомендовать в задачах исследования операций. Более того, наша цель состояла не в том, чтобы найти верхний предел задержки и контролировать этот предел лучше, чем раньше. Главная неудовлетворенность прежними методами работы пунктов сбора платы заключалась в том, что средняя задержка менялась (от 2 до 50 сек при нормальной, по существу, загрузке).

Чтобы уменьшить диапазон изменений среднего времени задержки (при том же среднем значении) и оптимизировать обслуживание, было предложено считать стандартным среднее, а не минимальное время задержки. Когда интенсивность движения возрастает, средняя задержка превышает стандартный уровень на величину, равную уменьшению времени задержки (по сравнению с тем же уровнем) при работе дополнительных пунктов сбора платы. Задача состоит в уменьшении произвола при определении среднего значения задержки.

Первый метод решения этой задачи мог бы состоять в определении приращения времени задержки при увеличении интенсивности движения. Например, допустим, что увеличение интенсивности движения на 10 машин в час соответствует дополнительной задержке в 1 сек. Такую линеаризацию вряд ли можно логически обосновать. Подобный, но несколько лучший метод решения состоит в том, чтобы считать одинаковым время, затрачиваемое клиентом, и время, затрачиваемое сборщиком платы. Таким образом, дополнительный пункт надо открыть тогда, когда производство интенсивности движения на уменьшение задержки, обусловленное открытием этого пункта, составит 3600 сек. Хотя этот метод имеет определенный смысл, он не использовался.

Третий метод решения состоит в нахождении точки уменьшения дохода. Этот метод менее спорный и основан на понятиях, широко распространенных среди администраторов. В данном случае стоимость характеризуется задержкой, а доход — интенсивностью движения. Можно определить значения, при которых начнет уменьшаться отношение дохода к стоимости. Эти значения определяют точку минимальной кривизны на соответствующих кривых. Выше этой точки увеличение интенсивности движения, соответствующее увеличению задержки, становится все меньше и меньше, приближаясь к нулю при стремлении задержки к бесконечности.

Определенные таким образом точки уменьшения дохода могут быть найдены экспериментально. Значения их для моста Дж. Вашингтона лежат в диапазоне 10,5—16 сек со средним взвешенным значением 12 сек. Для туннелей Линкольна и Голланда взвешенные средние составляли соответственно 10 и 11 сек. Так как было желательным обеспечить одинаковый уровень обслуживания для двух туннелей и моста, для них было принято в качестве стандарта среднее значение, равное 11 сек.

Теперь можно установить стандартную пропускную способность для различных комбинаций пунктов. При этом производится выравнивание отклонений вверх и вниз от стандартной задержки путем предусмотрения дополнительных пунктов сбора платы. Результаты вычислений для моста

Дж. Вашингтона приведены в таблице 15.7. Следует заметить, что длина очереди при занятости пунктов, соответствующей стандартной средней задержке, возрастает с числом пунктов, изменяясь от 6 (при одном пункте) до 12 (при восьми пунктах). К счастью, этот результат вполне приемлем,

Таблица 15.7

Пропускная способность пунктов у моста Дж. Вашингтона

Пункты		Пропускная способность (машин/час)	Задержка (сек)	Число машин в очереди	Пункты		Пропускная способность (машин/час)	Задержка (сек)	Число машин в очереди
лево-сторонние	право-сторонние				лево-сторонние	право-сторонние			
1		225	0—16,9	6	4	1	1250	10,5—12,0	9
2		450	5,1—15,5	7	4	2	1525	10,0—11,8	10
3		750	6,5—14,0	8	4	3	1850	10,2—11,5	11
4		1050	8,0—11,5	8	4	4	2150	10,5—11,5	12

так как опыты показали, что клиенты спокойней относятся к более длинным очередям при большей интенсивности движения. Очевидно, они интуитивно чувствуют, что очередь из 12 машин при 8 открытых пунктах качественно отличается от очереди из 12 машин при одном открытом пункте.

ПОЧАСОВОЙ ГРАФИК ИНТЕНСИВНОСТИ ДВИЖЕНИЯ

При известной интенсивности движения мы можем теперь определить два критерия обслуживания — среднюю задержку и максимальную очередь, если установлен стандарт для одного из них. Возникает вопрос, насколько

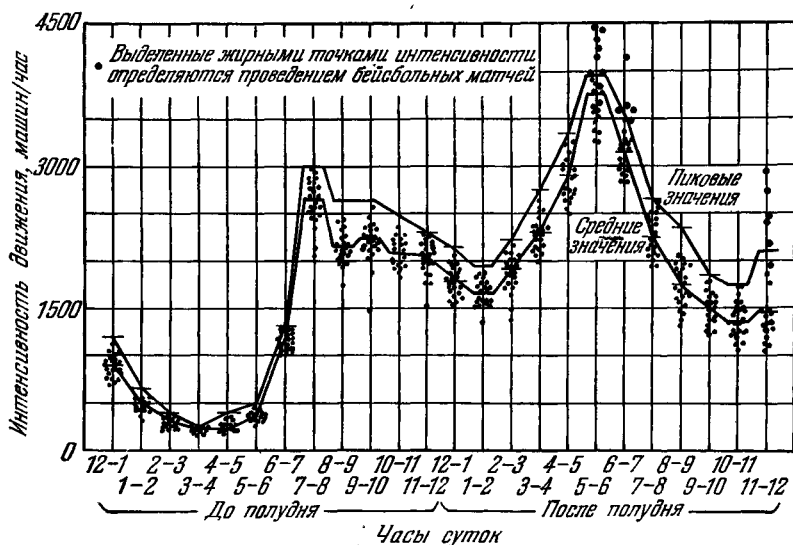


Рис. 15.16. Почасовая интенсивность движения через мост Дж. Вашингтона.

точно можно предсказать интенсивность движения. Для ответа на этот вопрос необходимо изучить почасовой график интенсивности движения

в один из дней, а также дисперсию интенсивности движения от дня ко дню. При построении графика по оси абсцисс откладывается время суток, а по оси ординат — интенсивность движения. Было установлено, что графики для дней в середине недели почти идентичны и могут быть объединены. На рис. 15.16 приведены объединенные графики для вторника, среды и четверга (мост Дж. Вашингтона, лето 1952 г.). Графики для других дней недели значительно отличались и требовали отдельного рассмотрения.

На рис. 15.16 построены две кривые. Одна из них проходит через средние точки; построение такой кривой является наиболее простым способом оценки ожидаемых значений интенсивности движения. Другая кривая проходит через пиковые значения. Она позволяет найти простую оценку ожидаемых максимальных значений интенсивности движения.

Изучение графиков показало, что разность между максимальными и средними значениями меняется от 10% до 60% (мост Дж. Вашингтона). Для туннелей разброс был меньше: 10—30%. Эти вариации определяют, насколько точно можно заранее составить расписание работы пунктов, с тем чтобы обеспечить оптимальное обслуживание.

Теперь необходимо рассмотреть последнюю часть общей задачи — составление графика работы пунктов и сборщиков платы.

СОСТАВЛЕНИЕ ГРАФИКА

При составлении расписания в первую очередь определяют число требуемых пунктов, исходя из пропускной способности различных комбинаций пунктов, найденной для оптимальной средней задержки и средней интенсивности движения. Из-за быстрых изменений интенсивности движения в дневное время нужно определять числа пунктов через каждые полчаса. После проведения этих расчетов для каждого получаса определялась максимальная возможная очередь при пиковом значении интенсивности движения (далее рассматривались те случаи, где были указаны максимальные очереди, превышающие на несколько машин точки Пуассона). Несомненно, что в этом случае наша способность к удовлетворительным предсказаниям довольно быстро падает. При приближении интенсивности потока машин к пропускной способности пунктов небольшое увеличение интенсивности может вызвать значительное увеличение очереди. Хотя для исследования ситуации можно было использовать точные методы, здесь это не обязательно. Следующее решение определяет степень нашего риска: предлагается допустить выше точек Пуассона увеличение очереди на 3 машины. Следовательно, если разность между средней и максимальной интенсивностью движения достаточно велика для того, чтобы очередь превысила этот уровень, надо предусмотреть дополнительные пункты.

В итоге получим расписание работы пунктов в течение дня, из которого можно определить общее количество часов работы пунктов. Остается определить количество сборщиков платы, необходимое для выполнения намеченного расписания работы пунктов и обеспечивающее для каждого из сборщиков перерывы для личных надобностей и для принятия пищи. При этом должны соблюдаться следующие правила:

1. Периоды работы между перерывами или от перерыва до конца работы не должны быть меньше 1 часа и больше 3 часов.
2. Перерывы на еду должны быть в средние четыре часа рабочего дня сборщика.
3. Начало работы не раньше 6 часов утра и конец работы не позже 0 часов 30 минут.

Составление расписания для людей требует подготовки для каждого дня графиков типа графиков Ганта, показывающих периоды работы и отдыха для каждого сборщика платы. Время начала работы и перерыва для сборщиков должно быть подобрано так, чтобы обеспечить количество сборщиков, необходимое для оптимального обслуживания в течение любого получаса дня. Эта проблема решается в основном методом проб и ошибок. Подготовка такого расписания является очень трудоемкой работой, если ставить себе целью сделать его максимально эффективным.

Эффективность расписания оценивается отношением числа сборщиков, требующегося по часам работы пунктов, к числу сборщиков, предусмотренному расписанием. Например, по вторникам, средам и четвергам летом 1953 года у моста Дж. Вашингтона требовалось 344 часа работы. Приемлемое время работы сборщика равно 6,25 часа. Следовательно, минимально необходимое число сборщиков равно $344 : 6,25 = 55,04$. Если расписание предусматривает 57 человек, то его эффективность $55,04 : 57 = 97\%$. Первые расписания были не очень эффективными, причем всегда возникал вопрос, является ли полученное расписание наиболее эффективным, поскольку число работавших сборщиков было больше, чем первое целое число, превышающее минимальную потребность. Можно потерять много времени, пытаясь уменьшить число сборщиков, тогда как это действительно нельзя сделать при наложенных ограничениях.

Результаты анализа и данные опытов показывают, что эффективность расписания в большой степени зависит от амплитуды и продолжительности периодов максимальной интенсивности движения. Можно получить оценку числа сборщиков, необходимого для каждой смены, если учесть перерывы в утренний и вечерний периоды загрузки, а также в период после полуночи. Эта оценка находится путем подсчета требуемых часов работы пунктов для пика продолжительностью 3,5 часа и делением полученного числа на 3. Такая процедура позволяет учесть получасовой перерыв для каждого сборщика. Например, в утренний пик у моста Дж. Вашингтона по вторникам, средам и четвергам необходимы 70 часов работы пунктов, т. е. $70/3 = 23,3$, или 24 сборщика. В вечерний пик нужно 71,5 часа работы пунктов, т. е. $71,5/3 = 23,8$, или 24 сборщика. После полуночи нужен 21 час работы пунктов, или $21/3 = 7$ сборщиков. Общее количество сборщиков для трех смен составляет 55 человек; это указывает, что можно составить расписание, учитывающее действительные условия. Число сборщиков, предусмотренное действительным расписанием, составляло 56 человек, что дает эффективность $55,04/56 = 98,3\%$. В большинстве подобных задач можно получить эффективность расписания 95% и больше.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

С разработкой эффективного метода составления расписания решена последняя часть рассматривавшейся задачи. Однако, прежде чем рекомендовать полученные результаты для использования, нужно ответить на следующий вопрос: может ли предложенный метод действительно дать лучшие результаты, чем прежний метод, который в настоящее время позволяет сержанту приблизительно правильно определять число сборщиков платы и использовать свои собственные суждения по поводу того, сколько пунктов держать открытыми при изменении движения и когда отпустить сборщиков на перерыв? Единственный способ ответить на этот вопрос состоял в проверке предложенного метода. Если он даст хорошие результаты в течение недели, то его можно применять в течение любого времени.

Проверка проводилась у туннеля Линкольна. Для каждого получаса недели было предсказано необходимое число сборщиков в обоих направлениях движения. В результате было получено 512 предсказаний числа пунктов. Каждому сборщику был дан график, в котором было указано время работы его пункта и время перерыва. Сборщики получили указание строго следовать расписанию. Всю неделю работа шла точно по расписанию, без каких-либо изменений. За все это время ни разу не было чрезмерных очередей и ни разу не было необходимости уменьшать перерывы. Работа сборщиков, открытие и закрытие пунктов проходили без всякого вмешательства сержантов. Временами число открытых пунктов было слегка избыточным, но не до такой степени, как при прежнем методе. Нет необходимости говорить о чувстве удовлетворения, когда видишь, что такая большая работа успешно завершена.

ЛИТЕРАТУРА

1. Berkeley G. S., Traffic and Trunking Principles in Automatic Telephony, Ernest Benn, Ltd., London, 1949.
2. Brockmeyer E., Høstrome H. L. and Jensen Arne, The Life and Work of A. K. Erlang, Copenhagen Telephone Co., Copenhagen, 1948.
3. Crommelin C. D., P. O. Elect. Engrs' J., 26, pt. 4 (Jan. 1934).
4. Greenshields B. D., Shapiro D. and Erickson E. L., Traffic Performance at Urban Street Intersections, Yale University, New Haven, 1947.

Глава 16

МОДЕЛИ УПОРЯДОЧЕНИЯ

ВВЕДЕНИЕ

Одной из задач теории массового обслуживания, рассмотренных в 14-й главе, является задача определения числа обслуживающих единиц, необходимых для минимизации суммарных потерь, связанных с временем ожидания как обслуживающих единиц, так и абонентов. В этой главе рассматривается другая задача, в которой число обслуживающих единиц принимается фиксированным, а определяется или распределение поступающих требований, или последовательность обслуживания ожидающих абонентов, или обе зависимости. Задача сводится к нахождению календарного плана поступления требований или упорядочению операций, которые необходимо выполнить, чтобы сумма соответствующих затрат была минимальной.

Под «календарным планированием» здесь понимается определение времени поступления требований на обслуживание (или времени окончания обслуживания, или обеих величин). Например, расписание движения поездов или автобусов указывает время прибытия и отправления. «Упорядочение» означает определение «порядка», в котором обслуживаются ожидающие абоненты. Например, для деталей, поступающих на обработку, на «входе» станочного парка можно найти порядок (последовательность), в которой их следует обрабатывать. Часто путают понятия календарного планирования и упорядочения, тем самым забывая разницу между двумя типами задач. Задача календарного планирования решается с помощью теории массового обслуживания, т. е. структура этой задачи та же, что и задачи определения числа обслуживающих единиц, обсуждавшейся

в 14-й главе. Задача календарного планирования отличается только характером искомых величин, т. е. определяется число обслуживающих единиц или время (или скорость) поступления требований. Поскольку модели и методы, приведенные в 14-й главе, применимы к задаче календарного планирования, то в 16-й главе эти задачи не рассматриваются. В этой главе в основном решается задача упорядочения.

Модели упорядочения. Задачу упорядочения начали исследовать с помощью математического аппарата совсем недавно. К настоящему времени достигнуты сравнительно небольшие успехи. Даже нет полной формулировки задачи, так как существующая формулировка связана только с минимизацией некоторой функции времени, не учитывается наличие противоречивых целей, что характерно для задач исследования операций. В действительности же в задачах упорядочения необходимо учитывать такие противоречия. Например, при определении последовательности обработки деталей на ряде машин следует не только минимизировать общее время обработки (чтобы уменьшить стоимость материалов, находящихся в обработке, и ускорить получение конечного продукта при определенных капиталовложениях), но и учитывать материальную заинтересованность рабочих на разных рабочих местах. Существуют и другие условия, например, связанные с приоритетами сроков отгрузки отдельных изделий (и соответствующими потерями в связи с задержкой), которые обычно противоречат требованию минимизации некоторой функции времени обработки.

Эти противоречия следует учесть при окончательной формулировке задачи упорядочения. Уже достигнуты определенные успехи в этом направлении. Например, Роу и Джексон [17] сформулировали задачу упорядочения с учетом приоритетов и получили решения для некоторых частных случаев.

Эти недостатки не умаляют значения проделанной в этой области работы. Однако следует предостеречь читателя против механического применения описываемых моделей.

Задачи упорядочения чаще всего встречаются в производстве. Это не удивительно, так как более эффективное использование имеющегося оборудования — получение большего выхода — является главной задачей производственных организаций. Руководители предприятий охотно поддерживают различные виды исследований в области улучшения использования наличного оборудования с помощью методов планирования или упорядочения работ на этом оборудовании.

Многие планово-производственные отделы стремятся достигнуть эффективного использования оборудования с помощью таких средств, как ленточные диаграммы (диаграммы Ганта), циклограммы, и других наглядных методов планирования (см. гл. 10 в книге Мура [15]). Но какими бы полезными ни были эти средства, с их помощью часто не только не удастся получить оптимальную последовательность, но и нельзя определить, насколько полученная последовательность близка к оптимальной.

Чтобы оценить сложность задачи упорядочения, рассмотрим случай обработки четырех изделий, каждое из которых обрабатывается на каждой из пяти машин. Существует $(4!)^5$, или 7 962 624, различных вариантов обработки (последовательностей), некоторые из которых могут быть нереализуемы из-за того, что определенные операции следует выполнять в заданном порядке. Ясно, что заслуживает внимания любой способ, который даст нам возможность получить оптимальную или близкую к ней последовательность без перебора всех или большинства вариантов.

Как уже указывалось, часто «требования очередности», вытекающие из технологии процесса, ограничивают число возможных последовательно-

стей работ. Например, деталь перед окраской должна быть обезжирена, а отверстие сначала нужно просверлить, прежде чем нарезать резьбу. Во всех этих случаях необходимо, чтобы операция, которая должна следовать за другой операцией, выполнялась бы на том же самом рабочем месте, что и предшествующая, или же на рабочем месте, которое в выбранной последовательности стоит после первого.

ОБРАБОТКА n ДЕТАЛЕЙ НА ДВУХ МАШИНАХ БЕЗ ПРОПУСКОВ

Рассмотрим очень простой случай обработки n деталей на двух машинах A и B , причем каждая деталь требует одной и той же последовательности операций. Кроме того, не разрешаются пропуски, т. е. если какая-нибудь деталь обрабатывается на первой машине в первую очередь, то она

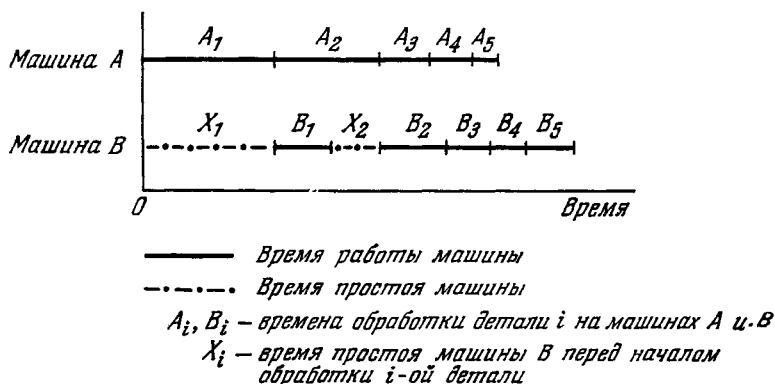


Рис. 16.1. Диаграмма Ганта.

должна обрабатываться и на второй машине тоже в первую очередь, а деталь, которая обрабатывается на машине A во вторую очередь, должна также обрабатываться во вторую очередь на машине B и т. д. Такое условие необходимо, например, для химических процессов, где рассматриваются материальные потоки от одного рабочего места до другого по конвейеру или по трубопроводу. Однако допускается существование промежуточных емкостей, т. е. материал может временно запасаться на конвейере, в трубопроводе или в резервуаре, пока не освободится следующее рабочее место. Тем временем предшествующее рабочее место освобождается для новой операции. Без потери общности можно принять, что все детали начинают обрабатываться на машине A , а затем на машине B . Обозначим A_i время, потребное на обработку i -й детали на машине A , B_i — время, потребное на обработку i -й детали на машине B , T — время обработки всех деталей $1, 2, \dots, n$, X_i — время простоя машины B от момента окончания обработки i — 1-й детали до момента начала обработки i -й детали.

Задача состоит в том, чтобы определить последовательность ($i_1, \dots, i_n$, где $i_1, \dots, i_n$ — перестановка чисел от 1 до n) такую, чтобы T было минимальным. Существует $n!$ возможных последовательностей. Одна из последовательностей для $n = 5$ показана на диаграмме Ганта* на рис. 16.1.

На рис. 16.1 представлена последовательность 1, 2, 3, 4, 5. Деталь 1 обрабатывается на машине A в течение A_1 часов, пока машина B простаивает.

* Использование диаграмм Ганта см. подробнее у Мура [15, стр. 228—235].

Как только заканчивается обработка детали 1 на машине A , на ней начинается обработка детали 2, а деталь 1 поступает на машину B и т. д. Время обработки всех деталей T определяется от момента начала обработки детали 1 на машине A до момента окончания обработки детали 5 на машине B . Машина B в любой момент времени или работает, или простаивает. Общее время работы машины B равно $\sum_{i=1}^5 B_i$; оно определяется технологией процесса, а не последовательностью. Очевидно, что

$$T = \sum_{i=1}^5 B_i + \sum_{i=1}^5 X_i. \quad (1)$$

Требуется минимизировать T , но так как $\sum_{i=1}^5 B_i$ — постоянная величина, то задача сводится к минимизации $\sum_{i=1}^5 X_i$. Из рис. 16.1 видно, что

$$\begin{aligned} X_1 &= A_1, \\ X_2 &= A_1 + A_2 - B_1 - X_1, \text{ если } A_1 + A_2 \geq X_1 + B_1, \\ X_2 &= 0, \text{ если } A_1 + A_2 < X_1 + B_1. \end{aligned}$$

Выражение для X_2 можно переписать в виде

$$X_2 = \max(A_1 + A_2 - B_1 - X_1, 0) = \max\left(\sum_{i=1}^2 A_i - \sum_{i=1}^1 B_i - \sum_{i=1}^1 X_i, 0\right).$$

Используя те же обозначения, заметим, что

$$X_1 + X_2 = \max(A_1 + A_2 - B_1, X_1) = \max\left(\sum_{i=1}^2 A_i - \sum_{i=1}^1 B_i, A_1\right).$$

Аналогично

$$\begin{aligned} X_3 &= \max\left(\sum_{i=1}^3 A_i - \sum_{i=1}^2 B_i - \sum_{i=1}^2 X_i, 0\right), \\ \sum_{i=1}^3 X_i &= \max\left(\sum_{i=1}^3 A_i - \sum_{i=1}^2 B_i, \sum_{i=1}^2 X_i\right) = \max\left(\sum_{i=1}^3 A_i - \sum_{i=1}^2 B_i, \sum_{i=1}^2 A_i - B_1, A_1\right). \end{aligned}$$

Обозначим

$$D_n(S) = \sum_{i=1}^n X_i,$$

где $D_n(S)$ есть функция от последовательности S . В общем виде

$$\begin{aligned} D_n(S) &= \sum_{i=1}^n X_i = \\ &= \max\left(\sum_{i=1}^n A_i - \sum_{i=1}^{n-1} B_i, \sum_{i=1}^{n-1} A_i - \sum_{i=1}^{n-2} B_i, \dots, A_1\right) = \max_{1 \leq u \leq n} \left(\sum_{i=1}^u A_i - \sum_{i=1}^{u-1} B_i\right). \quad (2) \end{aligned}$$

Это означает, что выражение в скобках оценивается отдельно для каждой положительной величины u (от 1 до n) и максимальная из всех величин равна $D_n(S)$.

Теперь задачу можно сформулировать следующим образом: выбрать такой порядок обработки деталей, чтобы минимизировать $D_n(S)$. Джонсон

[13] и Беллман [3] аналитически определили правила выбора оптимальной последовательности. Правила, разработанные Джонсоном, приводятся в приложении в конце этой главы. Эти правила принятия решения можно представить в виде следующего процесса, который иллюстрируется на примере, показанном в таблице 16.1.

Таблица 16.1
Время обработки
(в часах) пяти
деталей на двух
машинах

i	A_i	B_i
1	3	6
2	7	2
3	4	7
4	5	3
5	7	4

Процесс определения оптимальной последовательности [13].

1. Рассмотрим интервалы времени A_i и B_i и определим минимальную величину $[\min(A_i, B_i)]$. В приведенном примере эта величина равна $B_2 = 2$.

2. Если эта величина находится в столбце A_i , то i -ю деталь помещаем на машину A в первую очередь. Если эта величина находится в столбце B_i (как в нашем примере), то деталь i занимает последнюю очередь на машине A . Таким образом, деталь 2 на машине A обрабатывается последней.

3. Вычеркиваем выбранную деталь и продолжаем процедуру поиска, повторяя шаги 1 и 2. В случае двух одинаковых значений выбираем любую деталь. В нашем примере после выбора 2-й детали минимальное время равно 3, и оно соответствует A_1 и B_4 . Мы можем выбрать любую деталь, поэтому произвольно выбираем A_1 , т. е. помещаем на первое место деталь 1. Теперь



Рис. 16.2. Диаграмма Ганта оптимальной последовательности пяти деталей на двух машинах.

минимальное время соответствует B_4 . Следовательно, деталь 4 ставится на предпоследнее место. Следующая минимальная величина равна 4 (A_3 и B_5). Мы можем назначить 2-е место на машине A для детали 3 и 3-е место для детали 5. Полученная последовательность (1, 3, 5, 4, 2) будет оптимальной.

Эта последовательность представлена на диаграмме Ганта на рис. 16.2. Из рис. 16.2 видно, что время обработки всех деталей равно 28 часам и суммарное время простоев на машине B равно 6 часам.

Следует указать некоторые особенности этой задачи. Во-первых, считается, что не задан порядок окончания обработки отдельных деталей, т. е. не требуется получения одной детали быстрее другой. Задача усложняется, если ввести функцию предпочтения на сроки окончания обработки

отдельных деталей. Такие задачи рассматриваются ниже в данной главе. Во-вторых, считается, что существуют промежуточные емкости и стоимость материалов, находящихся в обработке, не зависит от вида деталей или слишком мала, так что ее можно не учитывать. Это справедливо для кратковременных процессов, но для длительных процессов иногда следует учитывать эти стоимости.

ОБРАБОТКА n ДЕТАЛЕЙ НА ТРЕХ МАШИНАХ БЕЗ ПРОПУСКОВ

Рассмотрим задачу, подобную предшествующей, с той разницей, что имеется три машины. Для каждой детали необходима та же последовательность операций и не допускаются пропуски. Введем новые обозначения с учетом обозначений, использованных в предыдущей задаче: Y_i — время простоя третьей машины перед началом обработки i -й детали, C_i — время обработки i -й детали на третьей машине. Тогда время обработки всех деталей выражается в виде

$$T = \sum_{i=1}^n C_i + \sum_{i=1}^n Y_i. \quad (3)$$

Следовательно, задача минимизации T сводится к задаче минимизации $\sum_{i=1}^n Y_i$, так как $\sum_{i=1}^n C_i$ — величина постоянная. Джонсон [13] нашел оптимальное решение * этой задачи для специального случая, когда или

(1) $\min A_i \geq \max B_i$ (наименьшее время, потребное для обработки на машине A любой из деталей, больше или равно наибольшему времени обработки на машине B любой детали), или

(2) $\min C_i \geq \max B_i$.

В примере, приведенном в таблице 16.2, первое условие выполняется в виде точного равенства.

Для получения оптимальной последовательности составляют новую таблицу (табл. 16.3). Из таблицы 16.3 получаем оптимальные последовательности с помощью описанного выше процесса определения оптимальной последовательности для двух машин. В этом случае оптимальными будут следующие последовательности:

3, 2, 1, 4, 5, 3, 2, 4, 5, 1, 3, 2, 4, 1, 5, ...

В тех случаях, когда условия $\min A_i \geq \max B_i$ или $\min C_i \geq \max B_i$ не выполняются, общего алгоритма для получения оптимальной последовательности еще не найдено. А следовательно, и для более общей задачи, в которой рассматриваются m машин и n деталей с одинаковыми (идентичными) маршрутами и без пропусков, тоже не найдено решения. Однако остается справедливым следующее положение: определяя оптимальные

* Джонсон показал, что для общего случая трех машин и n деталей

$$\sum_{i=1}^n Y_i = \max_{1 \leq u \leq v \leq n} (H_v + K_u),$$

где

$$H_v = \sum_{i=1}^v B_i - \sum_{i=1}^{v-1} C_i, \quad v = 1, 2, \dots, n,$$

$$K_u = \sum_{i=1}^u A_i - \sum_{i=1}^{u-1} B_i, \quad u = 1, 2, \dots, n.$$

Таблица 16.2

Время обработки (в часах)
пяти деталей на трех машинах

i	A_i	B_i	C_i
1	8	5	4
2	10	6	9
3	6	2	8
4	7	3	6
5	11	4	5

Таблица 16.3

Суммарное время обработки
(в часах) каждой из пяти
деталей на первой и средней
машинах и на средней
и последней машинах

i	$A_i + B_i$	$B_i + C_i$
1	13	9
2	16	15
3	8	10
4	10	9
5	15	9

последовательности (в смысле минимального времени обработки всех деталей), нужно минимизировать суммарное время простоев последней машины.

Идентичные маршруты, возможны пропуски. Хотя каждая из n деталей должна пройти через каждую из m машин в заданном технологическом порядке, не всегда необходимо с точки зрения процесса обработки, чтобы порядок обработки n деталей на каждой из машин был одинаков, т. е. допускаются пропуски. Беллман [3] и Джонсон [13] показали, однако, что для двух или трех машин одной из оптимальных последовательностей всегда будет последовательность с одинаковым порядком деталей на каждой машине. Этот результат несправедлив для случая $m > 3$.

РАЗЛИЧНЫЕ МАРШРУТЫ

Для многих производств, особенно механических, характерно, что изготовление различных продуктов требует различных маршрутов прохождения деталей по машинам или рабочим местам.

Две детали и m машин. Рассмотрим случай обработки двух деталей с различными маршрутами на m машинах. Для каждой детали допускается только один маршрут, и каждая машина может одновременно обрабатывать только одну деталь. Предполагается наличие промежуточных емкостей для полуфабрикатов. Задача заключается в определении последовательности, которая минимизирует время обработки всех деталей.

Существует 2^m возможных последовательностей, но не все они осуществимы (реализуемы) с точки зрения технологии процесса. В данной задаче может быть больше нереализуемых последовательностей, чем реализуемых. Поэтому следует устранить нереализуемые последовательности. Кроме того, среди оставшихся реализуемых последовательностей существуют заведомо неоптимальные, и их также нужно устранить. Акерс и Фридмен разработали метод для устранения таких последовательностей с использованием символической логики (в частности, булевой алгебры). Этот метод дает возможность получить подмножество последовательностей, среди которых обязательно содержатся оптимальные.

Рассмотрим этот метод на примере * двух деталей (изделий) и четырех машин a , b , c и d . Предположим, маршруты для каждого изделия заданы

* Пример взят из статьи Акерса и Фридмена [1].

в следующем виде:

Деталь 1	a, b, c, d
Деталь 2	d, b, a, c

Положим, что высказывание (правило) A означает, что на машине a деталь 1 обрабатывается перед деталью 2; высказывание $\bar{A}$ означает, что на машине a деталь 2 обрабатывается перед деталью 1. Аналогично $B, \bar{B}, C, \bar{C}, D$ и $\bar{D}$ — соответствующие высказывания для машин b, c и d . Например, в этом обозначении $A\bar{D}$ означает, что на машине a обрабатывается деталь 2 перед деталью 1. В таблице 16.4 представлены 16 возможных последовательностей (называемых «программами») обработки деталей на машинах.

Таблица 16.4

Все возможные программы обработки двух деталей на четырех машинах

Программа №	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$	A	$\bar{A}$
$\bar{B}$	$\bar{B}$	B	B	$\bar{B}$	$\bar{B}$	B	B	$\bar{B}$	$\bar{B}$	B	B	$\bar{B}$	$\bar{B}$	B	B	$\bar{B}$
$\bar{C}$	$\bar{C}$	$\bar{C}$	$\bar{C}$	C	C	C	C	$\bar{C}$	$\bar{C}$	$\bar{C}$	$\bar{C}$	C	C	C	C	$\bar{C}$
$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	D	D	D	D	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	D

Таблицу 16.4 можно представить в двоичном изображении, положив, что «1» означает выполнение правила и «0» означает невыполнение правила. Поместим четыре высказывания A, B, C и D в левом столбце и укажем на основании таблицы 16.4, выполняется ли каждое из правил для каждой из

Таблица 16.5

Двоичная таблица программ обработки двух деталей на четырех машинах

Программа №	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Высказывание	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
B	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
C	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
D	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

16 последовательностей. Результаты представлены в таблице 16.5. Если в таблице 16.4 приводится высказывание A , то в таблице 16.5 ему соответствует 1, и наоборот, если в таблице 16.4 дано высказывание $\bar{A}$, то в таблице 16.5 ставится 0, и т. д.

Следующим этапом является определение последовательностей, которые технологически нереализуемы. Для этого вернемся к заданным маршрутам для каждой детали:

Деталь 1	a, b, c, d
Деталь 2	d, b, a, c

Очевидно, что перед обработкой на машине d деталь 1 должна пройти обработку на машине a . Деталь 2, прежде чем обрабатываться на машине a , должна быть обработана на машине d . Следовательно, любая последовательность, которая включает $\bar{A}D$, является технологически нереализуемой. Это можно проследить следующим образом.

1. $\bar{A}$ означает, что деталь 2 должна обрабатываться перед деталью 1 на машине a .
2. Но деталь 2 не может поступить на машину a прежде, чем она не пройдет обработку на машине d .
3. Однако правило D означает, что деталь 2 не может поступить на машину d перед деталью 1. Но деталь 1 не может обрабатываться на машине d , пока она не пройдет обработку на машине a .
4. Следовательно, эту последовательность никогда нельзя начать. Акерс и Фридмен доказали следующую теорему: *чтобы программа для двух деталей была реализуема, необходимо и достаточно, чтобы для каждой пары машин x и y , где x предшествует y для детали 1 и x следует за y для детали 2, в программе не было высказывания $\bar{X}Y$.*

На основании этой теоремы мы можем в нашем примере исключить все программы, которые включают $\bar{A}D$, $\bar{A}B$, $\bar{B}D$ и $\bar{C}D$. Другие программы, включающие, например, высказывание $\bar{A}C$, не исключаются на основании этой теоремы, так как обе детали обрабатываются сначала на машине a , а затем на машине c .

Теперь вернемся к рассмотрению программ, в которых нет высказываний $\bar{A}D$, $\bar{A}B$, $\bar{B}D$ и $\bar{C}D$. Для этого просматривается каждый столбец в таблице 16.5 и проставляется 1, если «нет ($\bar{A}D$)» удовлетворяется, и 0, если не удовлетворяется. Например: каждая программа, в которой есть высказывание $\bar{A}D$, содержит 0 в первой строке и 1 в четвертой строке, так как 0 в первой строке таблицы 16.5 означает $\bar{A}$ и 1 в четвертой строке означает D . Аналогично в таблице 16.5

$\bar{A}B = 0$ в первой строке, 1 во второй строке,

$\bar{B}D = 0$ во второй строке, 1 в четвертой,

$\bar{C}D = 0$ в третьей строке, 1 в четвертой.

Таким образом, на основании таблицы 16.5 составляется таблица 16.6.

К реализуемым программам относятся программы, которые удовлетворяют всем четырем условиям и, следовательно, имеют 1 в каждой строке. Отбрасывается каждый столбец, в котором есть один или больше нулей. Эта операция эквивалентна получению логического произведения из строк и дает 1 в столбце, состоящем из одних 1, и 0 в противном случае. Из

Таблица 16.6

Определение технологически реализуемых программ с помощью двоичного представления высказываний

Программа № Высказывание	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Нет ($\bar{A}D$)	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1
Нет ($\bar{A}B$)	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1
Нет ($\bar{B}D$)	1	1	1	1	1	1	1	1	0	0	1	1	0	0	1	1
Нет ($\bar{C}D$)	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1
Логическое произведение	1	1	0	1	1	1	0	1	0	0	0	0	0	0	0	1

таблицы 16.6 видно, что из 16 возможных программ остается только 7, остальные 9 программ отбрасываются, как нереализуемые.

Акерс и Фридмен показали, что число возможных программ определяется из выражения

$$N = 1 + m + \sum_{k=2}^m i_k, \quad (4)$$

где i_k — число случаев, когда k определенных машин ($k = 2, 3, \dots, m$) имеют одинаковый порядок в технологическом маршруте обеих деталей, хотя между этими машинами могут стоять другие машины. Например, в нашем примере $m = 4$; $k = 2, 3, 4$, $i_2 = 2$ [(a, c) и (b, c)]; $i_3 = 0$, $i_4 = 0$. Следовательно,

$$N = 1 + 4 + 2 + 0 + 0 = 7.$$

Эти реализуемые программы показаны в таблице 16.7. Теперь надо исследовать их на оптимальность с помощью определенного критерия оптимальности.

Программа 16 (A, B, C, D) не может быть оптимальной, так как согласно ей нельзя начать обработку детали 2 до тех пор, пока не будет закончена обработка первой детали; обработка первой детали заканчивается на машине d , а вторая деталь начинает обрабатываться на машине d . Таким образом, две детали обрабатываются отдельно, и в каждый момент времени работает только одна из машин. Ясно, что любая программа, содержащая D , не может быть оптимальной. Наша задача заключается в том, чтобы исключить все такие программы.

Программу 16, которую мы только что рассмотрели, можно характеризовать следующим образом: в этой программе имеется машина, на которой обе детали обрабатываются последовательно во времени, причем пока эта машина работает, все другие машины простаивают. Такая машина называется *свободной машиной* для данной программы. Акерс и Фридмен обоб-

Таблица 16.7

Технологически реализуемые программы

Программа № Высказывание	1	2	4	5	6	8	16
A или $\bar{A}$	$\bar{A}$	A	A	$\bar{A}$	A	A	A
B или $\bar{B}$	$\bar{B}$	$\bar{B}$	B	$\bar{B}$	$\bar{B}$	B	B
C или $\bar{C}$	$\bar{C}$	$\bar{C}$	$\bar{C}$	C	C	C	C
D или $\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	$\bar{D}$	D

шили указанную особенность следующим образом: чтобы реализуемая программа для двух деталей принадлежала множеству оптимальных программ, необходимо и достаточно, чтобы в ней не было свободных машин *.

Для исключения программ, содержащих свободные машины, пользуются следующим правилом:

чтобы реализуемая программа не содержала свободных машин, необходимо и достаточно, чтобы для каждой машины y выполнялись следующие условия:

(1) если машины x и z расположены в следующем порядке:

Деталь 1: $\dots xy \dots z$,

Деталь 2: $\dots x \dots yz$,

то в программе не должно быть высказывания XYZ (когда y является первой машиной для детали 1, X опускается; когда же y является последней машиной для детали 2, опускается Z), и

(2) если машины u и v расположены в следующем порядке:

Деталь 1: $\dots u \dots uv$,

Деталь 2: $\dots uv \dots v$,

то в программе не должно быть высказывания $\bar{U}Y\bar{V}$ (когда y является первой машиной для детали 2, U опускается; когда y является последней машиной для детали 1, опускается V).

Используя (1) и полагая, что y означает машину a в нашем примере, а z означает машину c , получаем:

Деталь 1: $y \dots z$,

Деталь 2: $\dots yz$.

Следовательно, программы, содержащие $\bar{Y}Z$ (т. е. $\bar{A}C$), нужно исключить.

Снова применяя (1) и полагая на этот раз, что x означает машину b , а y означает машину c , получаем:

Деталь 1: $\dots xy$,

Деталь 2: $\dots x \dots y$.

Следовательно, надо исключить программы, содержащие $X\bar{Y}$ (т. е. $B\bar{C}$).

* Положения этого параграфа фактически представляют собой в сокращенном виде выдержки из работы Акерса и Фридмана [1]. Они только слегка видоизменены с тем, чтобы терминология всей главы была одна и та же.

Теперь вернемся к таблице 16.7 и определим, какие из реализуемых программ также удовлетворяют трем условиям оптимальности, т. е. не содержат указанных выше высказываний: D , $\bar{A}B$ и $B\bar{C}$. Результаты сведены в таблицу 16.8.

Таблица 16.8

Реализуемые программы, которые удовлетворяют необходимым условиям оптимальности

Программа № \ Высказывание	1	2	4	5	6	8	16
Нет (D)	1	1	1	1	1	1	0
Нет ($\bar{A}B$)	1	1	1	0	1	1	1
Нет ($B\bar{C}$)	1	1	0	1	1	1	1
Логическое произведение	1	1	0	0	1	1	0

Таким образом, окончательно получены четыре программы, которые удовлетворяют необходимым (но не обязательно достаточным) условиям как технологической реализуемости, так и оптимальности: 1, 2, 6 и 8. Теперь остается только поставить продолжительности выполнения всех операций и оценить каждую из этих четырех программ, например, с помощью диаграммы Ганта. Положим, что в предыдущем примере заданы следующие продолжительности различных операций (см. табл. 16.9).

Таблица 16.9

Время обработки в часах

Машина \ Деталь	a	b	c	d
1	2	4	5	1
2	2	5	3	6

Диаграммы Ганта для четырех программ даны на рис. 16.3.

Программа 8 дает минимальное время обработки всех деталей, а следовательно, является оптимальной для данного примера.

n деталей и m машин. Метод Акерса — Фридмена «сокращения» задачи

отыскания последовательности обработки деталей на m машинах можно применить в случае n деталей и m машин. В этом случае следует рассмотреть $(n!)^m$ различных программ. Теорема I реализуемости программ в этом случае заменяется теоремой II (доказанной Акерсом и Фридменом), которая формулируется следующим образом: чтобы программа обработки n деталей была технологически реализуемой, необходимо и достаточно, чтобы для каждой «петли», соответствующей α -последовательности ($2 \leq \alpha \leq n$), т. е.

Деталь 1: ... a_1 ... z_1

Деталь 2: ... b_2 ... a_2

Деталь 3: ... c_3 ... b_3

Деталь $\alpha - 1$: $y_{\alpha-1}$... $x_{\alpha-1}$

Деталь α : z_α ... y_α

в программе не было замкнутого высказывания

$$A_{(2<1)}B_{(3<2)}C_{(4<3)} \dots Y_{(\alpha<\alpha-1)}Z_{(1<\alpha)}.$$

Здесь высказывание $K_{(i<j)}$ означает, что на машине k деталь i обрабатывается перед деталью j .

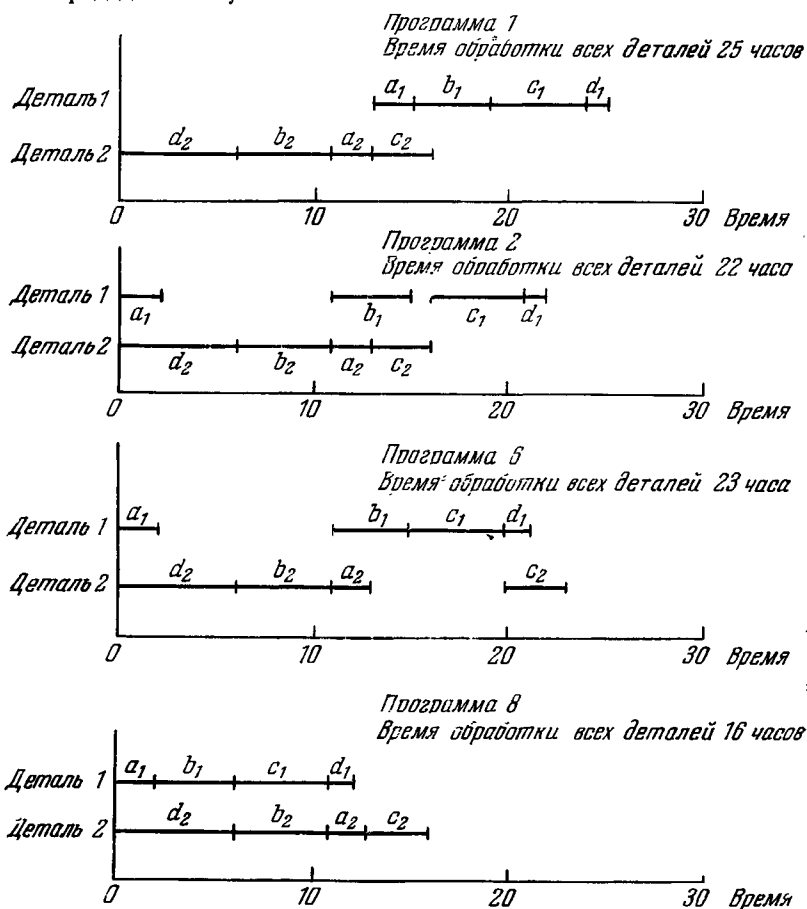


Рис. 16.3. Диаграмма Ганта для четырех реализуемых программ, близких к оптимальной программе.

Как было показано в предыдущем параграфе, небольшие задачи можно решать вручную, используя разработанный Акерсом и Фридменом метод решения без непосредственных вычислений. Преимуществом этого метода является возможность запрограммировать его на электронной вычислительной машине, что очень облегчает решение задач большого и среднего размера.

Следует отметить, что продолжительности операций необходимы только на последнем этапе, когда исключены все нереализуемые и заведомо неоп-

тимальные программы. Поэтому этот метод особенно ценен в тех случаях, когда технологические маршруты деталей фиксированы, а продолжительности операций меняются или из-за изменений в плане, или при изменении материалов (например, для различных сортов отливок необходимы различные тепловы режимы). В таких случаях при каждом изменении продолжительности операций проверяется на оптимальность только группа программ, которые выделены как реализуемые и близкие к оптимальным.

ВОЗМОЖНЫЕ ПУТИ РЕШЕНИЯ БОЛЕЕ СЛОЖНЫХ ЗАДАЧ УПОРЯДОЧЕНИЯ

Мы рассмотрели относительно простые задачи. Однако не существует ни аналитического решения, ни способов моделирования более сложных задач, в которых возможны различные маршруты для одной детали, существует несколько машин определенного типа, продолжительности обработки или затраты имеют вероятностный характер, возможны поломки машин или рабочие могут уклоняться от работы или мешать работе. Необходимо научиться решать каждую из этих задач и находить приемлемые решения.

Существует по крайней мере два возможных подхода к решению сложных задач. Один из них заключается в том, чтобы разбить задачу на подзадачи, которые можно решать с помощью методов, эффективных в соответствующей области. Этому методу присущи все недостатки, которые встречаются при управлении системой по частям, т. е. при оптимизации отдельных частей системы. Работа системы в целом может даже ухудшиться. Если же такая методика приводит к улучшению работы системы в целом, то она расценивается как «хорошая». Таким образом, даже когда приходится рассматривать систему по частям из-за отсутствия эффективных методов, следует иметь в виду работу системы в целом*.

Второй подход к решению задачи предполагает использование метода Монте-Карло. При этом используются распределения вероятностей таких случайных величин, которые характеризуют систему в терминах времени обработки на каждой машине и исправности машины. Кроме того, можно учесть вероятность различных маршрутов для изделий, наличие нескольких машин одного типа, изменение темпа работы в зависимости от рабочей смены и сезонные колебания всех этих величин. В результате можно создать модель системы, в которой по желанию можно учитывать различные особенности работы системы. На такой модели можно сравнивать различные методики упорядочения или отдельные последовательности. Например, мы можем сравнивать следующие две методики относительно определенного критерия эффективности системы.

1. Рассмотрим ряд операций, которые нужно выполнить на определенной машине, и назовем на первое место операцию, у которой до заданного срока завершения осталось меньше всего времени. В качестве примера рассмотрим случай с одной машиной, на которой с 1 марта необходимо выполнить 5 операций при условиях, заданных в столбцах (2), (3) и (4) таблицы 16.10.

В столбцах (1), (2), (3) и (4) таблицы 16.10 приведены исходные данные. Последовательность, указанная в столбце (5), составлена исходя из того же порядка выполнения операций, что и в столбце (4). Эффективность этой методики можно оценить в среднем, создав модель системы (при этом можно

* Более полную информацию по этому вопросу можно получить у Хитча [12].

Таблица 16.10

Пример планирования работ на машине

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Опера- ция №	Ожидаемые длительности операций на машине, в днях	Заданный срок испол- нения	Время до заданного срока испол- нения, в днях	Последова- тельность I	Время до заданного срока испол- нения за вы- четом дли- тельности операции	Последо- вательность II
593	3	$3/20$	20	4	17	4
465	10	$3/22$	22	3	12	2
607	1	$3/20$	20	1	19	5
305	2	$3/12$	12	2	10	1
336	7	$3/23$	23	5	16	3

учесть вероятностные характеристики таких величин, как длительности операций). Такая процедура иногда называется *операционной игрой* или *операционным экспериментом* *.

2. Вторую возможную методику упорядочения можно проследить на примере столбцов (6) и (7) таблицы 16.10. В этом случае последовательность определяется следующим образом: рассматриваются операции, которые необходимо выполнить на определенной машине, и назначается на первое место операция, у которой осталось меньше всего времени до заданного срока завершения за вычетом длительности этой операции.

Если мы многократно повторим опыт на модели **, то можем сравнить в вероятностном смысле результаты этих двух, а также других стратегий упорядочения и определить, какая из стратегий обладает в среднем наивысшей эффективностью.

Использование быстродействующих электронных вычислительных машин облегчает моделирование, особенно в сложных случаях.

Применение вычислительных машин для моделирования подробно излагается у Де Карло [8]. Предварительные опыты в этом направлении были проведены Роу и Джексоном [17]. Ротман [16] с помощью большой электронной вычислительной машины решал вопросы управления работой большого цеха.

Модель, использованная Ротманом ***, дает возможность исследовать влияние колебаний времени поступления требований, времени обслуживания, а также учесть возможность поломки оборудования, отсутствия обслуживающего персонала, нехватки материалов и т. д. Кроме того, дается правило изменения последовательности для различных опытов на вычислительной машине. Экспериментально было определено число перемен последовательностей, необходимое для получения наилучшего возможного решения. Каждая последовательность оценивалась по средней оценке, полученной при проведении восьми опытов для данной последовательности. Затем была выбрана последовательность с наименьшей полученной оценкой.

* Подробности относительно этого подхода к решению задачи см. у Кушена [6] и Торнвейта [19].

** О способах уменьшения количества опытов по методу Монте-Карло см. у Канана и Маршалла [14].

*** Приведена в лекции Гроша по краткому курсу исследования операций в Кейсовском технологическом институте (1954 г.). См. также [16].

Метод моделирования для решения больших задач упорядочения начали применять только недавно. В ближайшем будущем можно ожидать дальнейших успехов в этой работе.

СМЕЖНЫЕ ЗАДАЧИ

Среди задач, смежных с задачей упорядочения, особенно интересны две задачи. Первая — задача балансировки линии поточной сборки — также заключается в минимизации общего времени сборки при определенной последовательности операций. Вторая — задача коммивояжера — состоит в определении такого маршрута коммивояжера через ряд населенных пунктов, при котором минимизируется или пройденное расстояние, или время, или стоимость путешествия. В более общей постановке задача заключается в упорядочении ряда операций (например, посещений населенных пунктов) на оборудовании (например, коммивояжер) так, чтобы минимизировать некоторые характеристики движения от одной операции к следующей. Здесь не приводятся существующие точные или приближенные решения этих двух задач; они рассмотрены, чтобы познакомить читателя с их характером и указать источники, в которых содержится более подробная информация.

Задача балансировки линии поточной сборки. Сборочная линия состоит из последовательности рабочих мест (обслуживающих единиц), через которые изделие проходит в заранее установленном порядке. Рассмотрим сборочную линию, которая предназначена для обработки только одного вида изделия. Следует так спроектировать сборочную линию, чтобы минимизировать время обработки.

Процесс сборки разбит на ряд операций. Задача балансировки сборочной линии состоит в объединении этих операций, т. е. в такой группировке их, чтобы можно было поместить эту группу операций в одном пункте (рабочем месте). В идеальном случае на все рабочие места сборочной линии должны быть назначены операции с равным объемом работы (измеренным в единицах времени). Если бы это было возможным, то не было бы простоев ни на одном рабочем месте и сборочная линия была бы идеально сбалансирована. Если же объем работ на рабочих местах различен, то рабочее место с наибольшим объемом работы становится узким местом и производительность всей сборочной линии определяется производительностью этого рабочего места. В результате наблюдаются простои на некоторых рабочих местах и сборочный процесс замедляется. Такая линия называется *разбалансированной*.

Следовательно, при проектировании сборочной линии желательно минимизировать сумму простоев на рабочих местах, группируя определенным образом операции. Но в реальных задачах очень трудно добиться минимизации времени, особенно в связи с тем, что время выполнения отдельной операции или группы операций не остается постоянным даже для одного рабочего или машины и тем более для разных рабочих. Кроме того, обслуживающий персонал сборочной линии часто меняется. Чтобы упростить задачу, следует принять время выполнения каждой операции или группы операций постоянным. Это упрощение позволило получить первые результаты в решении задачи.

Если принять продолжительности выполнения каждой операции постоянными, то задача заключается в определении оптимальной комбинации операций, которые нужно так сгруппировать на каждом рабочем месте,

чтобы общая задержка была минимальной или сведена к нулю. Общая задержка равна сумме задержек за цикл на всех рабочих местах, где задержка определяется как время простоя на каждом рабочем месте из-за разбаланса линии.

В задаче учитываются ограничения (называемые требованиями очередности), которые вытекают из технологии процесса сборки. Например, нужно выполнить все работы внутри электрического прибора, прежде чем закрыть его корпусом. Во всех подобных случаях необходимо, чтобы операция, которая должна следовать за другой операцией, выполнялась бы на том же рабочем месте, что и предшествующая операция, или на рабочем месте, которое следует за первым.

Существуют два пути решения этой задачи. Первый, разработанный Брайтоном [5], предполагает количество рабочих мест постоянным; минимизируется общее время простоев за счет минимизации длительности цикла (т. е. времени, необходимого для выпуска единицы продукции на рабочем месте) на рабочем месте с наибольшим объемом работ. Брайтон разработал способ определения локального минимума, но не дал систематического метода нахождения абсолютного минимума.

Второй путь решения, развитый Селвсоном [18], предполагает длительность периода постоянной, определяется же оптимальное число рабочих мест. Селвсон разработал вычислительный метод, который приводит к оптимальному решению в большинстве случаев. Оба метода (и Селвсона и Брайтона) можно запрограммировать для решения на вычислительных машинах.

Задача балансировки сборочной линии встречается во многих областях, не связанных с производством. Например, в системе внутригородских грузовых перевозок до отправки поезда, состоящего из определенного числа груженых прицепов, которые необходимо перевезти со складов отправителей, расположенных в разных частях города, проходит определенный период времени, который может быть постоянным (аналогично циклу на сборочных линиях). Предполагаем, что все тягачи, потребные для перевозки прицепов, расположены на станции отправления и время для перевозки каждого прицепа известно (аналогично времени операции). Если мы потребуем минимизации числа тягачей (аналогично требованию минимального количества рабочих мест), то получим задачу, сходную с задачей балансировки сборочной линии: используя минимальное количество тягачей, минимизировать суммарное время простоев для всех тягачей. В этой задаче могут быть требования очередности в том смысле, что некоторые прицепы надо доставить раньше других.

Задача коммивояжера (выбора маршрута). Задача коммивояжера является «классической» математической задачей, которая состоит в определении маршрута через несколько населенных пунктов при условии, что каждый населенный пункт посещается один и только один раз и конечный пункт совпадает с начальным. Под оптимальным маршрутом понимается маршрут, при котором сумма расстояний (или стоимость путешествия или время) минимальна. Таким образом, эта задача предполагает такое упорядочение населенных пунктов, чтобы минимизировать некоторую функцию, характеризующую процесс передвижения между ними.

Аналогичная задача встречается в других областях. Например, рассмотрим производственный процесс или поточную линию, на которой производится определенное количество различных, но сходных изделий. У этих изделий могут быть общие детали; например, при производстве металли-

ческих кухонных шкафов могут быть общие дверцы, крышки, стенки и т. д. Таким образом, затраты на наладку линии для производства нового изделия зависят от того, какое изделие непосредственно предшествовало данному. Если два изделия имеют много общих деталей, то затраты на наладку относительно невелики, в противном случае они относительно высоки. Задача заключается в определении такой последовательности запуска изделий в производство, при которой сумма затрат на наладку линии была бы минимальной.

Задача коммивояжера по содержанию относится к задаче упорядочения, но в действительности она по характеру более сходна с задачей о назначении, рассмотренной в 12-й главе. Подобно задаче о назначении, ее можно представить в матричной форме, но, поскольку в ней не допускаются петли, требования к решению несколько иные. Однако иногда можно решать эту задачу как задачу о назначении [10], а затем проверить, содержится ли в решении петли. Если петель нет, то полученное решение будет оптимальным. Если же есть петли, то это решение используется в качестве начального приближения для дальнейшего решения задачи коммивояжера.

Таблица 16.11

Стоимостная матрица для задачи коммивояжера

Из \ В	A	B	C	D	E
A	∞	K_{AB*}	K_{AC}	K_{AD}	K_{AE}
B	K_{BA}	∞	K_{BC*}	K_{BD}	K_{BE}
C	K_{CA}	K_{CB}	∞	K_{CD*}	K_{CE}
D	K_{DA}	K_{DB}	K_{DC}	∞	K_{DE*}
E	K_{EA*}	K_{EB}	K_{EC}	K_{ED}	∞

Предположим, например, что нужно изготовить 5 изделий A, B, C, D и E. Построим матрицу, как указано в таблице 16.11. В каждую диагональную клетку таблицы проставляем бесконечную стоимость *, так как мы не хотим, чтобы какое-то изделие обрабатывалось само за собой. Все другие клетки таблицы заполняются затратами на наладку при условии, что изделия, стоящие в столбцах, обрабатываются после изделий, стоящих в строках. Например, K_{BA} — затраты на наладку производства изделия A, следующего за изделием B. Теперь можно использовать методы решения задачи о назначении. В результате получаем решение, при котором выбирается по одному элементу в каждой строке и столбце, причем сумма затрат в этих клетках минимальна.

* Здесь мы фактически используем условие (20) задачи о назначении, рассмотренной в главе 12.

Однако полученное решение может оказаться невозможным. Например, будут выбраны клетки ED и DE , что означает необходимость обработки D после E и в то же время E после D . Предположим, однако, что E обрабатывается в последнюю очередь и минимальные затраты получаются при выборе элементов, отмеченных звездочкой. Оптимальным решением будет A, B, C, D, E . Это решение называется «наклонным»*. В общем случае справедливо следующее положение [10]: *если наклон стоимостной матрицы является оптимальным решением задачи о назначении, то наклон является также оптимальным решением задачи коммивояжера, или задачи выбора маршрута.*

Хотя не существует общего аналитического решения таких задач, можно всегда получить решение, перебрав все возможные варианты. Однако это может потребовать слишком много времени**. Были достигнуты некоторые успехи в аналитическом решении задачи. Данциг, Фулкерсон и Джонсон [7] предложили четыре способа, иногда дающие решение в «симметрических» вариантах задачи, в которых расстояние (стоимость или время) между двумя элементами одинаково независимо от направления. Флад [9, 10] предложил метод решения симметрических задач.

Следует отметить, что в некоторых случаях значительно легче получить относительно грубое приближение к решению задачи. Такой случай рассмотрен Хером и Хагли [11].

ЗАКЛЮЧЕНИЕ

Задача упорядочения является одной из самых сложных задач исследования операций. В настоящее время найдено решение только для простых случаев.

Чрезвычайно простым является алгоритм Джонсона [13] определения оптимальной последовательности для n изделий, имеющих одинаковый маршрут следования через два рабочих места. Этот алгоритм легко можно использовать в цеховых условиях. Существует аналогичный алгоритм для случая трех рабочих мест и n изделий, имеющих идентичный маршрут следования.

В случае n изделий с различным маршрутом и m рабочих мест необходимо просмотреть $(n!)^m$ программ. Многие из них могут оказаться нереализуемыми из-за технологических ограничений, и их можно исключить. Среди оставшихся программ могут быть заведомо неоптимальные программы, и их также нужно исключить. Методика такого отбора программ без непосредственных вычислений разработана Акерсом и Фридменом [1].

Задача балансировки сборочной линии обычно возникает при наличии повторяющегося процесса и заключается в таком распределении операций на рабочих местах, чтобы сумма простоев на всех рабочих местах при заданном времени цикла была минимальной. Задача усложняется при наличии ограничений, связанных с требованиями очередности. Использование метода Селвсона [18] обычно приводит к оптимальному решению.

* *Наклоном матрицы $n \times n$ называется набор элементов, состоящий из элементов над главной диагональю и элемента в последней строке первого столбца.*

** Рассмотрим, например, задачу упорядочения 20 операций на одном рабочем месте. Существует $20! = 2\,432\,902\,008\,176\,640\,000$ различных последовательностей. Чтобы найти решение на быстродействующей электронной вычислительной машине, оценивающей каждую последовательность за одну микросекунду и работающей в течение 8 часов 365 дней в году, потребовалось бы почти четверть миллиона лет.

Задача коммивояжера, или выбора маршрута, заключается в таком упорядочении объектов, чтобы минимизировать сумму стоимостей, времен, расстояний или других величин, связанных с переходом от одного объекта к другому, при условии, что каждый объект встречается в последовательности один и только один раз. Общего аналитического решения этой задачи еще не найдено. В некоторых случаях ее можно решить как задачу о назначении [10]. В других случаях можно использовать методы, разработанные Данцигом, Фулкерсоном, Джонсоном [7] и Фладом [9, 10].

В настоящее время наиболее эффективным средством решения сложных задач упорядочения, особенно благодаря быстродействующим вычислительным машинам, является моделирование (операционная игра или операционный эксперимент), часто с использованием метода Монте-Карло.

П р и л о ж е н и е

ДОКАЗАТЕЛЬСТВО АЛГОРИТМА ОПРЕДЕЛЕНИЯ ОПТИМАЛЬНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ ДЛЯ n ДЕТАЛЕЙ НА ДВУХ МАШИНАХ

Рассматривается задача упорядочения n деталей с одинаковым маршрутом следования на двух машинах A и B без пропусков. Используются следующие обозначения: A_i — время обработки детали i на машине A , B_i — время обработки детали i на машине B , X_i — время простоя машины B от момента окончания обработки детали $i - 1$ до начала обработки детали i , $D_n(S) = \sum_{i=1}^n X_i$ — суммарное время простоев на машине B для последовательности S .

Задача заключается в определении такой последовательности S^* для деталей $(1, \dots, n)$, чтобы $D_n(S^*) \leq D_n(S_0)$ для любой последовательности S_0 .

Джонсон [13] и Беллман [3] показали, что оптимальную последовательность можно получить следующим образом: деталь j располагается перед деталью $j + 1$, если

$$\min(A_j, B_{j+1}) < \min(A_{j+1}, B_j); \quad (5)$$

порядок деталей j и $j + 1$ не влияет на время обработки, если

$$\min(A_j, B_{j+1}) = \min(A_{j+1}, B_j). \quad (6)$$

Джонсон [13] приводит следующее доказательство этого правила.

Начнем с последовательности S' и получим из нее другую последовательность S'' , меняя местами детали j и $j + 1$. Имеем две последовательности

$$S' = 1, 2, 3, \dots, j-1, j, j+1, j+2, \dots, n,$$

$$S'' = 1, 2, 3, \dots, j-1, j+1, j, j+2, \dots, n.$$

Обозначим

$$K_u = \sum_{i=1}^u A_i - \sum_{i=1}^{u-1} B_i,$$

причем K'_u соответствует K_u для последовательности S' и K''_u — для S'' . Из уравнения (2) следует

$$D_n(S) = \max_{1 \leq u \leq n} \left(\sum_{i=1}^u A_i - \sum_{i=1}^{u-1} B_i \right).$$

Таким образом, получаем

$$D_n(S) = \max_{1 \leq u \leq n} \sum K_u.$$

Заметим, что $K'_u = K''_u$ для $u = 1, 2, \dots, j-1, j+2, \dots, n$, но K'_j и K'_{j+1} могут отличаться от K''_j и K''_{j+1} соответственно. При этом $D_n(S')$ отличается от $D_n(S'')$.

Справедливы следующие два положения:

(а) Если $\max(K'_j, K'_{j+1}) = \max(K''_j, K''_{j+1})$, то $D_n(S') = D_n(S'')$ и, согласно критерию минимизации всего времени обработки, не существует разницы между последовательностями S' и S'' .

(б) Если $\max(K'_j, K'_{j+1}) < \max(K''_j, K''_{j+1})$, то следует выбрать последовательность S' , а не S'' , т. е. деталь j должна предшествовать детали $j+1$. Указанное соотношение можно выразить иначе, если заменить K'_j , т. е.

$$K'_j = \sum_{i=1}^j A_i - \sum_{i=1}^{j-1} B_i, \quad K'_{j+1} = \sum_{i=1}^{j+1} A_i - \sum_{i=1}^j B_i.$$

Следовательно,

$$\max(K'_j, K'_{j+1}) = \max\left(\sum_{i=1}^j A_i - \sum_{i=1}^{j-1} B_i, \sum_{i=1}^{j+1} A_i - \sum_{i=1}^j B_i\right). \quad (7)$$

Аналогично

$$K''_j = \sum_{i=1}^{j-1} A_i + A_{j+1} - \sum_{i=1}^{j-1} B_i, \quad K''_{j+1} = \sum_{i=1}^{j+1} A_i - \sum_{i=1}^{j-1} B_i - B_{j+1}.$$

Следовательно,

$$\max(K''_j, K''_{j+1}) = \max\left(\sum_{i=1}^{j-1} A_i + A_{j+1} - \sum_{i=1}^{j-1} B_i, \sum_{i=1}^{j+1} A_i - \sum_{i=1}^{j-1} B_i - B_{j+1}\right). \quad (8)$$

Вычитая $\sum_{i=1}^{j+1} A_i - \sum_{i=1}^{j-1} B_i$ из правых частей выражений (7) и (8), получим следующий результат:

если

$$\max(-A_{j+1}, -B_j) < \max(-A_j, -B_{j+1}), \quad (9)$$

то

$$\max(K'_j, K'_{j+1}) < \max(K''_j, K''_{j+1}). \quad (10)$$

Умножая выражение (9) на -1 , что означает также изменение знака неравенства, получим

$$\min(A_{j+1}, B_j) > \min(A_j, B_{j+1})$$

или, что то же самое,

$$\min(A_j, B_{j+1}) < \min(A_{j+1}, B_j).$$

Теперь можно выразить положение (б) в виде следующего правила:

деталь j предшествует детали $j+1$, если

$$\min(A_j, B_{j+1}) < \min(A_{j+1}, B_j).$$

Джонсон показал, что это соотношение транзитивно. Можно получить оптимальную последовательность S^* из любой начальной последовательности S_0 , применяя вышеуказанное правило последовательной замены следующих друг за другом деталей. Каждая такая замена дает значение $D_n(S)$ не большее, чем в предыдущей последовательности.

Л И Т Е Р А Т У Р А

1. A k e r s S. B., Jr., and F r i e d m a n J., A Non-Numerical Approach to Production Scheduling Problems, J. Opns. Res. Soc. Am., 3, no. 4, 429—442 (Nov. 1955).
2. B a r a n k i n E. W., The Scheduling Problem as an Algebraic Generalization of Ordinary Linear Programming, Discussion Paper No. 9, Logistic Research Project, University of California, Los Angeles, Aug. 28, 1952.
3. B e l l m a n R., Mathematical Aspects of Scheduling Theory, RAND Report P-651, RAND Corporation, Santa Monica, Apr. 11, 1955.
4. B e l l m a n R. and G r o s s O., Some Combinatorial Problems Arising in the Theory of Multi-stage Processes, J. Soc. Indust. Appl. Math., 2, no. 3, 175—183 (Sept. 1954).
5. B r y t o n B., Balancing of a Continuous Production Line, M. S. Thesis, Northwestern University, Evanston, June 1954.
6. C u s h e n W. E., Operational Gaming in Industry, in J. F. McCloskey and J. M. Copping (eds.), Operations Research for Management, II, The Johns Hopkins Press, Baltimore, 1956.
7. D a n t z i g G., F u l k e r s o n R. and J o h n s o n S., Solution of a Large-Scale Travelling-Salesman Problem, J. Opns. Res. Soc. Am., 2, 393—410 (1954).
8. D e C a r l o C. R., The Use of Automatic and Semi-Automatic Processing Equipment in Production and Inventory Control, Proceedings of the Conference on Operations Research in Production and Inventory Control, Case Institute of Technology, Cleveland, 1954.
9. F l o o d M. M., Operations Research and Logistics, Proceedings: First Ordnance Conference on OR, Office of Ordnance Research, Durham, pp. 3—25, Jan. 1955.
10. F l o o d M. M., The Traveling-Salesman Problem, J. Opns. Res. Soc. Am., 4, no. 1, 61—75 (Feb. 1956).
11. H a r e V. C. and H u g l i W. C., Applications of Operations Research to Production Scheduling and Inventory Control, II, Proceedings of the Conference on «What Is Operations Research Accomplishing in Industry?», Case Institute of Technology, Cleveland, 1955.
12. H i t c h C., Sub-Optimization in Operations Problems, J. Opns. Res. Soc. Am., 1, no. 3, 87—99 (May 1953).
13. J o h n s o n S. M., Optimal Two- and Three-Stage Production Schedules with Setup Times Included, Nav. Res. Log. Quart., 1, no. 1, 61—68 (Mar. 1954).
14. K a h n H. and M a r s h a l l A. W., Methods of Reducing Sample Size in Monte Carlo Computations, J. Opns. Res. Soc. Am., 1, no. 5, 263—278 (Nov. 1953).
15. M o o r e F. G., Production Control, McGraw-Hill Book Co., New York, 1951.
16. R o t h m a n S., A Problem in Production Scheduling, General Electric Co., Evendale, Ohio, 1953 (privately circulated).
17. R o w e A. J. and J a c k s o n J. R., Research Problems in Production Routing and Scheduling, Research Report No. 46, Management Sciences Research Project, University of California, Los Angeles, Oct. 26, 1955.
18. S a l v e s o n M. E., The Assembly Line Balancing Problem, J. Indust. Eng., 6, no. 3, 18—25 (May—June 1955).
19. T h o r n t h w a i t e C. W., Operations Research in Agriculture, J. Opns. Res. Soc. Am., 1, no. 2, 33—38 (Feb. 1953).

Эта часть состоит из одной лишь 17-й главы, в которой освещаются некоторые вопросы теории и применения моделей замен (их называют также моделями восстановления). Представленная здесь работа является частью результатов, полученных при осуществлении программы фундаментальных исследований, субсидированных и возглавляемых Кейсовским технологическим институтом.

Большинство исследований, связанных с процессами замен, было выполнено за пределами исследования операций *. Теория исследования операций занималась приложением этих исследований к неизученным ранее явлениям и лишь теперь начинает самостоятельные теоретические изыскания.

В зависимости от характера оборудования процессы замен делятся на два класса. Первый связан с оборудованием, которое, в процессе эксплуатации, устаревая, становится менее производительным физически вследствие износа или морально в результате появления новых, более совершенных машин. Примером такого вида оборудования служат металлорежущие станки. Второй класс связан с оборудованием, которое не устаревает в процессе эксплуатации, но зато в некоторый момент приходит в полную негодность. Здесь в качестве примера можно назвать электрические лампочки.

При решении задач первого класса приходится находить равновесие между затратами на приобретение нового оборудования и издержками, связанными с эксплуатацией уже имеющихся машин. Эти издержки вызваны либо потерей в производительности работающего оборудования, либо необходимостью поддерживать эту производительность на определенном уровне, либо и тем и другим вместе.

В общем виде задачи первого класса пока еще не решены. Тем не менее для ряда частных случаев удалось найти решения и построить модели замен.

Гранту [20] удалось решить задачу замены оборудования при следующих предположениях: (а) новые, более эффективные машины доступны лишь в момент замены, (б) величина процента на капитал остается постоянной в течение всего срока службы оборудования, (в) ежегодные издержки на эксплуатацию оборудования не снижаются.

Еще одну модель замены построил Терборг [41], предположив, что технология производства улучшается с постоянной скоростью. Пользуясь этой гипотезой, он вычислил скорость старения оборудования в прошлом и предсказал ее значение в будущем. Кроме того, он использовал в расчетах предсказанную на будущее цену нового оборудования. Возможные ошибки прогноза Терборг не учитывает.

Дин [11, 12] подверг критике Гранта за то, что при расчете инвестиций он полагал процентную ставку постоянной. Дин предложил метод, позволяющий сравнивать эффекты от различных инвестиций, и в его модели величина капиталовложений зависит от условий и возможностей коммерческой деятельности.

* Наибольший вклад в развитие этой области науки внесли две организации: Machinery and Allied Products Institute (MAPI) и National Center for Education and Research in Dynamic Equipment Policy.

Математические основы теории замен разрабатываются уже весьма давно. Эта проблема привлекла внимание многих выдающихся математиков, статистиков, экономистов и актуариев. Достаточно назвать имена Блекуэлла [4], Брауна [6], Чанга и Полларда [8], Чанга и Вольфовица [9], Дуба [14], Феллера [18], Карлина [24] и Принрейча [37].

Вслед за Алчианом [1] Беллман [2] предложил использовать для решения задачи замен метод функциональных уравнений теории динамического программирования. Беллман нашел стратегию, которая максимизирует общий доход при следующих предположениях: (а) задан объем производимой продукции, (б) известно, как меняется со временем стоимость ухода за оборудованием, (в) замена оборудования возможна лишь в определенные моменты времени, (г) установленное оборудование немедленно начинает выдавать продукцию.

При решении задач второго класса приходится определять, какие именно единицы оборудования следует заменить (например, все, за исключением установленных в прошлую неделю) и как часто следует проводить замену с тем, чтобы минимизировать общие затраты, которые складываются из следующих частей:

1. Стоимость устанавливаемого оборудования (т. е. закупочная цена или стоимость изготовления).

2. Затраты, связанные с заменой отдельных деталей.

3. Затраты, связанные с повреждением оборудования (например, потери заработка или дохода в связи с неисправностью оборудования).

Одной из крайностей может быть стратегия, в соответствии с которой оборудование заменяется лишь после его выхода из строя. Такая стратегия минимизирует затраты на оборудование (поскольку максимизируется срок его службы). Вместе с тем издержки, связанные с отдельными заменами и неисправностями, могут оказаться весьма значительными.

Другой крайностью является замена всех деталей после первой же поломки одной из них (или даже раньше). Такая стратегия приводит к высокой стоимости оборудования, но зато к малым затратам на некомплектность. При этом могут снизиться затраты и на замену в силу ее массовости. Оптимальная стратегия обычно лежит где-то посередине.

Срок службы оборудования носит обычно случайный характер. Для того чтобы найти закон распределения вероятностей повреждения (отказа) оборудования в зависимости от срока его службы, было предпринято значительное число исследований. Литература по этому вопросу обширна.

Гудман [19] разработал методику измерения и сравнения долговечности узлов, не требующую наблюдений за их отдельными деталями. Эпштейн и Собел [15, 16] проделали значительную работу, связанную со статистикой долговечности оборудования, условная вероятность выхода из строя которого постоянна. Дэвис [10] и другие установили, что вакуумные лампы имеют именно такую характеристику поломок.

На важность этой характеристики указал Будман [5] в своей работе «Надежность авиационного радарного оборудования». Идеи этой работы были далее развиты Шеллардом [40], который предложил методы расчета, позволяющие находить вероятность отказа оборудования в виде функции времени (Шеллард рассматривает сложную систему, состоящую из многих узлов, причем выход любого узла из строя приводит к повреждению всей системы в целом). Шеллард также исследовал, насколько можно повысить надежность такой системы путем замены отдельных ее узлов после определенного срока их службы.

После того как распределение вероятностей для срока службы оборудования найдено, обычно требуется составить выражение для среднего числа отказов в зависимости от времени его работы. В большинстве реальных ситуаций повреждения устраняются по мере их возникновения, т. е. в промежутках времени между групповыми

заменами. Поэтому выражение для среднего числа отказов аналитически найти трудно, и приходится разрабатывать различные процедуры, позволяющие вычислить это среднее число приближенно. Одна из таких процедур была предложена в МАРІ [39]. Она основана на предположении о том, что срок службы изделий подчиняется нормальному закону. В главе 17 используются дискретные аппроксимации непрерывного распределения длительности сроков службы. Кроме того, для нахождения среднего числа отказов применялся метод Монте-Карло с ускоренным масштабом времени.

В отделе ламп компании «General Electric» разработан алгоритм замены осветительных ламп. В соответствии с этим алгоритмом в период между двумя групповыми заменами в системе возможна максимум одна поломка. В главе 17 эта задача замены решается без учета возможности одной поломки.

Американской ассоциацией предпринимателей (American Management Association) опубликован обзор по алгоритмам замены оборудования. Он может оказаться полезным для промышленности.

Глава 17

МОДЕЛИ ЗАМЕН

ВВЕДЕНИЕ

Основными задачами теории замен являются прогноз затрат, связанных с обновлением оборудования, и выработка наиболее экономичной стратегии замен. Прогноз затрат, связанных с эксплуатацией группы изделий со случайной длительностью срока службы (например, лампы освещения, радиолампы), включает в себя два момента. Вначале находится оценка для закона распределения срока службы изделий. Далее на основании этой оценки вычисляется предполагаемое число отказов как функция срока эксплуатации всей группы изделий. Весьма обширная литература по этому вопросу содержит целый ряд схем, позволяющих приближенно найти число отказов [6, 7, 17, 18, 30, 37].

Прогноз затрат, связанных с оборудованием, производительность которого падает в процессе эксплуатации (сюда относятся, например, металлорежущие станки, автомобили и т. д.), отличается рядом особенностей. Здесь анализ издержек должен учитывать те факторы, которые приводят к росту производственных затрат, удлиняют время простоя, увеличивают число отказов и длительность ремонта и т. д.

Эксплуатация устаревшего оборудования связана с ростом производственных затрат. Вместе с тем замена старого оборудования новым также сопряжена с расходами. Существует некоторый срок службы, после которого замена устаревшего оборудования новым целесообразна, — с этого момента экономия за счет нового оборудования начинает превышать компенсацию его первоначальной стоимости.

В данной главе излагаются методы сравнения альтернативных стратегий замены. Здесь выводятся некоторые соотношения, связанные с задачей минимизации затрат, и указываются методы, позволяющие предсказать издержки с помощью законов распределения вероятностей сроков службы изделий.

ЗАМЕНА ОБОРУДОВАНИЯ, ПОДВЕРЖЕННОГО СТАРЕНИЮ

Для сравнения эффективности альтернативных стратегий замен в данной работе используется величина всех будущих расходов, приведенная к текущему моменту, связанных с каждой отдельной стратегией. Эти приведенные издержки складываются из всех затрат, связанных с применением той или иной стратегии, и погашаются за счет специально созданного фонда.

Затраты, связанные с заменой. В общем случае при решении задач замены необходимо учитывать все расходы, связанные с выбором машины и сроком ее эксплуатации. Хотя для погашения расходов ассигнуются наличные средства, тем не менее иногда приходится учитывать и дополнительные издержки, связанные с эффектом накопления, особенно если они влияют на количество наличных средств. В этом смысле наиболее ярким примером является уменьшение скидки на уплату подоходного налога.

В отдельных случаях некоторые издержки могут не включаться в общую калькуляцию. Например, при расчете оптимального срока замены одной отдельно взятой машины не следует учитывать затрат, которые не изменяются с ростом срока службы машины. К таким затратам можно отнести непосредственные производственные издержки на живую силу, энергию

и т. д. Пусть теперь стоит задача выбора одной из двух машин. Если постоянные (во времени) расходы, связанные с одной из них, не равны аналогичным расходам, связанным со второй, то такие издержки уже должны учитываться. Но зато теперь можно опустить другие виды расходов, одинаковые для обеих машин.

Особенно затруднителен учет затрат, связанных с техническим обслуживанием оборудования. Правда, можно предположить, что существует некоторая оптимальная стратегия обслуживания, и использовать в расчетах затраты, связанные именно с нею. Поиск оптимальной стратегии обслуживания относится уже к другому классу задач.

Таблица 17.1

Две модели затрат

Год	Издержки на начало года, долл		Приведенные издержки, долл (учетный процент=10%)	
	машина 1	машина 2	машина 1	машина 2
1	900	1400	900	1400
2	600	100	545,45	90,91
3	700	700	578,52	578,52
Итого	2200	2200	2023,97	2069,43
Разница	\$0		\$45,46	

Уровень издержек и цена денег, ввиду наличия процентов на капитал, меняются со временем. Поэтому ни один из двух нижеследующих критериев оптимизации: (1) критерий минимума неприведенных затрат и (2) критерий минимума средних приведенных затрат за время между двумя заменами — не может быть признан удовлетворительным. Это положение иллюстрируется таблицами 17.1 и 17.2.

В таблице 17.1 приведены издержки*, связанные с эксплуатацией двух машин (1 и 2) в течение 3 лет. Суммарные издержки за 3 года на каждую из машин составляют 2200 долларов. Однако машина 2 потребовала более высоких первоначальных затрат, нежели машина 1, и потому в действительности с учетом процентов на капитал обошлась дороже. В самом деле, хотя на первом году на машину 2 дополнительно израсходовали 500 долларов, а на втором году ту же сумму сэкономили, однако достаточно было ассигновать в первом году лишь 455 долларов, чтобы дополнительно израсходовать 500 долларов, нужных для машины 1 на втором году. (Здесь принято, что учетный процент составляет 10%.)

Таблица 17.2

Ежегодные расходы на машину с учетом замены

Год	Машина 1, долл	Машина 2, долл
1	1000	1700
2	200	100
3	400	200
4		300
5		400
6		500

* Приведенные издержки равны сумме истинных затрат. Они рассчитываются по формуле $C_n / (1 + r)^{n-1}$, где C_n — расходы на начало n -го года, r — годовой учетный процент, n — количество лет.

Таким образом, хотя суммарные расходы за 3 года на каждую из машин совершенно одинаковы, однако с учетом процентов на капитал машина 1 обошлась на 45,46 доллара дешевле.

Предположим снова, что годовой учетный процент составляет 10%. Пусть машина 1 подлежит замене через каждые 3 года, а машина 2 — через каждые 6 лет, и пусть ежегодные расходы на них соответствуют таблице 17.2.

Общие приведенные издержки на машину 1 за 3 года составляют

$$\$1000 + \frac{\$200}{1,10} + \frac{\$400}{(1,10)^2} = \$1512,$$

или 504 доллара в год.

Приведенные издержки на машину 2 за 6 лет равны

$$\$170 + \frac{\$100}{1,10} + \frac{\$200}{(1,10)^2} + \frac{\$300}{(1,10)^3} + \frac{\$400}{(1,10)^4} + \frac{\$500}{(1,10)^5} = \$2765,$$

или 461 доллар в год. Вполне очевидно преимущество машины 2. Однако это сравнение не объективно, ибо в первом случае во внимание принимались 3 года работы машины, а во втором — 6 лет. Общие приведенные издержки на машину 1 за 6 лет равны

$$\frac{\$1000}{1} + \frac{\$200}{1,10} + \frac{\$400}{(1,10)^2} + \frac{\$1000}{(1,10)^3} + \frac{\$200}{(1,10)^4} + \frac{\$400}{(1,10)^5} = \$2647,$$

что на 118 долларов меньше, чем на машину 2 за тот же самый период.

Сравнение этих двух машин оказалось возможным потому, что удвоенный срок службы машины 1 равен сроку службы машины 2. В общем случае такое сравнение можно производить лишь на бесконечных интервалах времени. Метод этого сравнения обсуждается в следующем разделе.

Уравнение затрат. Рассмотрим последовательность периодов времени одинаковой длины. Обозначим их 1, 2, 3, 4, ... и сопоставим каждому из них некоторую величину затрат $C_1, C_2, C_3, C_4, \dots$ соответственно.

Впредь при изучении вопросов, связанных с заменой оборудования, подверженного старению, будем предполагать, что эти затраты монотонно возрастают. Пусть выплата соответствующего количества денег осуществляется в начале каждого периода, начальная стоимость нового оборудования равна A , а учетный процент в течение каждого периода составляет 100%. Тогда приведенная сумма K_n всех будущих затрат, связанных со стратегией, по которой оборудование заменяется после каждых n периодов, составит

$$K_n = \left(A + C_1 + \frac{C_2}{1+r} + \frac{C_3}{(1+r)^2} + \dots + \frac{C_n}{(1+r)^{n-1}} \right) + \\ + \left(\frac{A+C_1}{(1+r)^n} + \frac{C_2}{(1+r)^{n+1}} + \dots + \frac{C_n}{(1+r)^{2n-1}} \right) + \dots \quad (1)$$

Это выражение можно переписать в виде

$$K_n = \left(A + \sum_{i=1}^n \frac{C_i}{(1+r)^{i-1}} \right) + \frac{1}{(1+r)^n} \left(A + \sum_{i=1}^n \frac{C_i}{(1+r)^{i-1}} \right) + \\ + \frac{1}{(1+r)^{2n}} \left(A + \sum_{i=1}^n \frac{C_i}{(1+r)^{i-1}} \right) + \dots \quad (2)$$

Представим правую часть уравнения (2) в виде произведения двух сомножителей, одним из которых является выражение в больших круглых

скобках, а вторым — сумма бесконечной сходящейся геометрической прогрессии. Тогда

$$K_n = \frac{A + \sum_{i=1}^n [C_i / (1+r)^{i-1}]}{1 - [1/(1+r)]^n}. \quad (3)$$

Таким образом, K_n — это общее количество денег, необходимых в данный момент времени для погашения всех будущих затрат, связанных с покупкой и эксплуатацией оборудования, при условии, что замена его осуществляется через каждые n лет. Это еще не означает, что какая-либо компания действительно учредит такого размера фонд. Однако если K_n меньше, чем K_{n+1} , то замена оборудования через каждые n лет выгоднее, чем замена после каждого $n+1$ года. Более того, если стратегия замены оборудования после каждых n лет является наилучшей, то должны выполняться два неравенства:

$$K_{n+1} - K_n > 0 \quad \text{и} \quad K_{n-1} - K_n > 0.$$

Можно показать*, что условие $K_{n-1} - K_n > 0$ эквивалентно неравенству

$$\frac{C_n}{1 - [1/(1+r)]} < K_{n-1}, \quad (4)$$

а условие $K_{n+1} - K_n > 0$ — неравенству

$$\frac{C_{n+1}}{1 - [1/(1+r)]} > K_n. \quad (5)$$

Неравенства (4) и (5) являются необходимым условием** оптимальности стратегии, состоящей в замене оборудования после каждых n периодов. Этой стратегии соответствуют минимальные затраты K_n . Из неравенств (4) и (5) можно сделать ряд полезных выводов. Рассмотрим вначале неравенство (4). Обозначим $X = 1/(1+r)$. Тогда

$$\frac{C_n}{1-X} < K_{n-1}. \quad (4a)$$

Из выражения (4a) следует, что $C_n < (1-X) K_{n-1}$. Подставляя в последнюю формулу значение K_{n-1} , полученное из (3) заменой n на $(n-1)$, находим

$$C_n < (1-X) \frac{(A + C_1 + C_2X + \dots + C_{n-1}X^{n-2})}{1 - X^{n-1}} \quad (46)$$

или

$$C_n < \frac{(A + C_1) + C_2X + \dots + C_{n-1}X^{n-2}}{1 + X + X^2 + \dots + X^{n-2}}. \quad (4b)$$

Правая часть неравенства (4b) является средневзвешенной суммой затрат за все предшествующие периоды времени, до $(n-1)$ -го включительно. Веса $1, X, X^2, \dots, X^{n-2}$ представляют собой коэффициенты приведения, соответствующие затратам в каждом периоде. Аналогичным образом можно представить и второе неравенство:

$$C_{n+1} > K_n (1-X), \quad (5a)$$

$$C_{n+1} > \frac{1-X}{1-X^n} (A + C_1 + C_2X + \dots + C_nX^{n-1}) \quad (5b)$$

* См. приложение в конце этой главы.

** Хотя неравенства (4) и (5) дают необходимое условие оптимальности стратегии, однако можно показать, что при монотонном возрастании C_n ($C_{n+1} > C_n$ при всех n) они являются и достаточным условием.

или

$$C_{n+1} > \frac{(A + C_1) + C_2X + \dots + C_nX^{n-1}}{1 + X + X^2 + \dots + X^{n-1}}. \quad (5в)$$

На основе анализа этих неравенств можно предложить следующие правила минимизации затрат.

1. Если издержки в следующем периоде ниже средневзвешенных прошлых затрат, то оборудование заменять не следует.

2. Если же издержки в следующем периоде превосходят величину средневзвешенных затрат, то оборудование следует заменить.

Геометрическая интерпретация этих правил представлена на графике рис. 17.1. Здесь по оси абсцисс отложена сумма коэффициентов приведения,

а по оси ординат — сумма приведенных затрат. Очевидно, средневзвешенные затраты равны тангенсу угла наклона прямой, проходящей через начало координат и нанесенную на чертеже точку P_i ($P_1, P_2, \dots$). Рассмотрим две последовательные точки P_n и P_{n+1} . По оси ординат эти точки разделяет расстояние $C_{n+1}X^n$, а по оси абсцисс — X^n . Поэтому тангенс угла наклона прямой, проходящей через точки P_n и P_{n+1} , равен $(C_{n+1}X^n)/X^n = C_{n+1}$. Если значение C_{n+1} меньше величины углового коэффициента линии OP_n , то угловой коэффициент прямой OP_{n+1} меньше углового коэффициента OP_n . (На рис. 17.1 это справедливо, например, для точки P_1 .) В этом случае замену не следует производить. Если же значение C_{n+1} больше величины углового коэффициента OP_n , то и угловой коэффициент OP_{n+1} больше углового коэффициента OP_n . (На рис. 17.1 это справедливо, например, для точки P_6 .) Здесь замена целесообразна. Рис. 17.1 построен по данным таблицы 17.3.

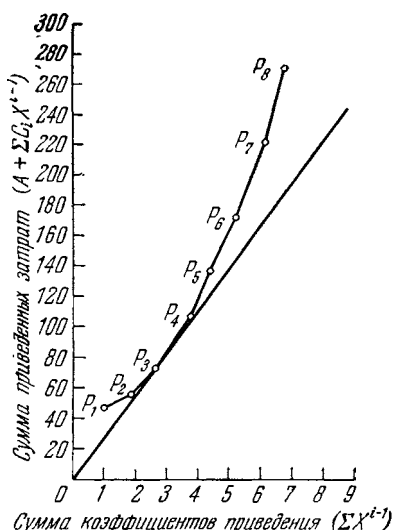


Рис. 17.1. Соотношение суммы приведенных затрат и суммы коэффициентов приведения (по данным табл. 17.3).

Из таблицы 17.3 видно, что минимальное значение средневзвешенных затрат (колонка (7)) достигается после трех периодов. Этот минимум находится путем сравнения предсказанных затрат в четвертом периоде (составляющих 30, колонка (2)) со средневзвешенными расходами за три предыдущих периода (т. е. соответственно 27,16, колонка (7)).

Из рис. 17.1 видно, что минимальный угловой коэффициент соответствует прямой OP_3 . Из рисунка также видно, что угловой коэффициент прямой OP_1 больше, чем угловой коэффициент OP_2 .

В общем случае угловой коэффициент прямой, проходящей через начало координат и произвольную точку P_d , равен

$$\frac{A + \sum_{i=1}^d C_i X^{i-1}}{1 + X + X^2 + \dots + X^{d-1}} = \frac{(1-X)(A + \sum_{i=1}^d C_i X^{i-1})}{(1-X)(1 + X + X^2 + \dots + X^{d-1})} = (1-X) K_d. \quad (6)$$

Так как при любом заданном $r(1-X) > 0$, то минимальный угловой коэффициент прямой соответствует наименьшему значению K_d , т. е. той стратегии, которая минимизирует затраты.

Таблица 17.3

Расходы, связанные с заменой

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Период (i)	Затраты (C _i)	$X^{i-1} \equiv \left(\frac{1}{1+r}\right)^{i-1}$	$C_i X^{i-1}$	$A + \sum C_i X^{i-1}$	$\sum X^{i-1}$	$\frac{A + \sum C_i X^{i-1}}{\sum X^{i-1}} = \frac{1}{(1-X)K_n}$
1	0 *	1,0000	0,00	50,00	1,0000	50,00
2	10	0,9524	9,52	59,52	1,9524	30,49
3	20	0,9070	18,14	77,66	2,8594	27,16 Замена
4	30	0,8638	25,92	103,58	3,7232	28,82
5	40	0,8227	32,91	136,49	4,5459	30,02
6	50	0,7835	39,18	175,67	5,3294	32,96
7	60	0,7462	44,77	220,44	6,0756	36,28
8	70	0,7107	49,75	270,19	6,7863	39,81
9	80	0,6768	54,14	324,33	7,4631	43,46
10	90	0,6446	58,01	382,34	8,1077	47,16

* $A + C_1 = 50$, где $A = 50$, $C_1 = 0$.
 Обозначения:
 Колонка (1): число прошедших периодов.
 Колонка (2): затраты в каждом периоде.
 Колонка (3): коэффициент приведения, принятый в соответствующем периоде.
 Колонка (4): затраты в каждом периоде, приведенные к началу первого периода (данные колонки (2) $\times$ данные колонки (3)).
 Колонка (5): сумма членов колонки (4).
 Колонка (6): сумма членов колонки (3).
 Колонка (7): отношение членов колонки (5) к членам колонки (6).

Применение уравнения затрат. В предыдущем разделе было выведено уравнение затрат — уравнение (3). Из этого уравнения были получены некоторые неравенства и с их помощью для данного вида оборудования были найдены оптимальные сроки замены. Эти соотношения иллюстрировались графиком. Точно так же с помощью графика или уравнения можно найти момент замены приобретенного оборудования — для этого надо подвергнуть анализу уже произведенные фактические расходы. При этом знание самого минимального значения K_n не обязательно; выше было показано, как найти число n , минимизирующее затраты. Эта процедура до некоторой степени напоминает нахождение минимума функции многих переменных. Значения независимых переменных, доставляющие минимум функции, можно найти и не зная, чему равно это минимальное значение, — достаточно лишь приравнять производные нулю и решить получившуюся систему уравнений.

Однако в том случае, когда эксплуатируемое оборудование сравнивается с новым или иным видом машин, знание минимального значения K_n для каждого типа оборудования уже необходимо. Затраты, связанные с работающим оборудованием, можно установить по записям бухгалтерии, а расходы, связанные с новым оборудованием, надо «предсказать» с помощью какого-либо метода. Простейший из них связан с формулой МАРІ, выведенной при следующих весьма жестких предположениях [41].

1. В течение всего срока эксплуатации расходы на машину возрастают по линейному закону.

2. Скорость роста затрат непосредственно не измеряется; ее значение определяется некоторым произвольным распределением «срока службы», полученным предприятием по результатам эксплуатации этого вида оборудования.

В данном разделе эти упрощающие предположения не учитываются. Более того, каждому, кто использует этот метод, представляется возможность самому выбрать способ прогноза затрат, наиболее подходящий для его задачи. В некоторых случаях будущие затраты можно оценить довольно точно, опираясь на достаточное количество информации чисто инженерного характера. С другой стороны, недостаток информации может потребовать крайне упрощенных предположений, например таких, которые были использованы при выводе формулы МАРІ.

Оценив затраты, можно с помощью описанного здесь метода определить «наилучшую стратегию» и далее найти минимум ожидаемых расходов, связанных с новым оборудованием. Сравнение этих данных с затратами на действующее оборудование уже не вызывает затруднений.

Предположим, что по отношению к приведенным затратам новое оборудование оказалось более экономичным. Это еще не означает, что необходима немедленная замена. Работающее оборудование может быть еще весьма новым и обладать достаточной эффективностью. Можно показать, что замена оборудования не должна осуществляться до тех пор, пока производственные затраты, приходящиеся на один период при старом оборудовании, не достигнут средневзвешенных затрат, связанных с новым оборудованием. Пусть K'_n — минимальная учетная величина всех будущих затрат на новое оборудование, $D_1, D_2, D_3, \dots, D_m$ — затраты, приходящиеся на каждый будущий период при условии, что эксплуатируется уже имеющееся оборудование, $X = 1/(1+r)$ — коэффициент приведения, где r — учетный процент, Π_m — приведенная величина всех будущих расходов при условии, что работающее оборудование заменяется после m периодов. Как и ранее, запишем:

$$\left. \begin{aligned} \Pi_m &= D_1 + D_2X + D_3X^2 + \dots + D_mX^{m-1} + K'_nX^m, \\ \Pi_{m+1} &= D_1 + D_2X + D_3X^2 + \dots + D_mX^{m-1} + D_{m+1}X^m + K'_nX^{m+1}, \\ \Pi_{m-1} &= D_1 + D_2X + D_3X^2 + \dots + D_{m-1}X^{m-2} + K'_nX^{m-1}. \end{aligned} \right\} \quad (7)$$

После вычитания имеем:

$$\Pi_{m+1} - \Pi_m = D_{m+1}X^m + K'_n(X^{m+1} - X^m), \quad (8)$$

$$\Pi_m - \Pi_{m-1} = D_mX^{m-1} + K'_n(X^m - X^{m-1}). \quad (9)$$

Из уравнения (8) видно, что неравенство $\Pi_{m+1} > \Pi_m$ эквивалентно соотношению

$$[D_{m+1}X^m + K'_n(X^{m+1} - X^m)] > 0,$$

т. е.

$$[D_{m+1} + K'_n(X - 1)] > 0$$

или, что то же самое.

$$D_{m+1} > (1 - X) K'_n. \quad (10)$$

Аналогично из уравнения (9) вытекает, что неравенство $\Pi_m < \Pi_{m-1}$ эквивалентно

$$D_m < (1 - X) K'_n. \quad (11)$$

Из неравенств (10) и (11) следует, что эксплуатация старого оборудования обеспечивает минимум затрат до тех пор, пока расходы в следующем периоде не превысят величины $(1 - X) K'_n$. Напомним, что значение $(1 - X) K'_n$ (см. уравнение (4б)) представляет собою средневзвешенные затраты, связанные с эксплуатацией оборудования, при условии, что замена осуществляется один раз за n периодов.

В некоторых случаях ожидание улучшенных процессов или улучшенного оборудования приводит к задержке в сроках замены. Причем предпочитают ожидать, но не заменять работающее оборудование на аналогичные или даже улучшенные, но иного образца машины. Изложенные в данной главе идеи могут помочь оценить благоразумие такой стратегии в тех отдельных случаях, когда известны характеристики затрат и момент ввода в строй нового процесса.

ЗАМЕНА ОБОРУДОВАНИЯ, ПОДВЕРЖЕННОГО ПОЛОМКАМ

Второй большой класс задач теории замен связан с оборудованием, которое в процессе эксплуатации мало меняет свои свойства, но зато по истечении некоторого срока службы приходит в полную негодность. Период времени между началом эксплуатации оборудования и его выходом из строя (отказом) для каждого вида оборудования не является постоянным, но подчиняется некоторому распределению.

Пусть закон распределения вероятностей для долговечности изделий известен. Тогда можно вычислить условную вероятность того, что отказ произойдет в течение некоторого малого интервала времени, например в промежутке от t до $t + \Delta t$. С ростом t эта условная вероятность может уменьшаться, оставаться неизменной или возрасти. Убывающий характер условной вероятности справедлив на начальной стадии эксплуатации. В этом случае исправное оборудование в начальный период срока службы увеличивает общее среднее время его эксплуатации. Примером промышленного оборудования с таким распределением долговечности служат авиационные двигатели. Вот почему для удлинения среднего срока службы двигателя и уменьшения вероятности его выхода из строя в начальный период времени в производственный процесс включают искусственное (в том смысле, что полезная работа не выполняется) старение двигателей. По истечении начального периода эксплуатации вероятность отказа, конечно, возрастает с ростом длительности срока службы.

Постоянная вероятность отказа характерна для изделий, которые приходят в негодность по случайным причинам, не связанным со сроком службы, например из-за сотрясений. В этом случае все изделия выходят из строя фактически независимо от срока их службы. Так, например, было показано, что скорость, с которой приходят в негодность вакуумные лампы авиационного оборудования, не зависит от срока их службы [5, 10].

В данном разделе главным образом будет изучаться лишь то оборудование, для которого вероятность выхода из строя увеличивается со сроком службы. Этот тип изделий приводит к наиболее интересным задачам стратегии замен. В большинстве случаев сама замена сломанного оборудования является тривиальной задачей.

Сущность вопроса обычно связана с проблемой капиталовложений либо со сравнением производительности оборудования с затратами на него. Если данное изделие жизненно необходимо для нормальной работы сложной системы, частью которой оно служит, то от стратегии замен зависит производительность всей системы в целом. Например, выход из строя насоса на нефтеочистительном заводе может остановить все производство, а затраты, связанные с заменой насоса, ничтожны по сравнению со стоимостью процессов. В дальнейшем изложении, до конца этой главы, мы будем предполагать, что все отказы должны быть устранены.

Теперь возникает задача о планировании замен неповрежденных изделий. До сих пор мы рассматривали изделия, которые заменялись по мере

ухудшения качества их работы. В данном разделе это ухудшение не является решающим фактором. Замена работающего, но все еще годного изделия новым имеет смысл только в том случае, если стоимость замены после поломки выше, чем до поломки, и если новая деталь уменьшает вероятность выхода оборудования из строя. Понятно, что в том случае, когда вероятность отказа уменьшается со временем или остается неизменной, второе условие уже не учитывается.

Из предыдущего изложения видно, что стратегия замен зависит от вероятности повреждений (отказов). Поэтому весьма важно оценить закон распределения вероятностей отказов. Статистические методы, позволяющие оценить долговечность деталей, были разработаны довольно быстро, и, как следствие, появилась обширная литература по этому вопросу [15, 16, 19].

Знание закона распределения вероятностей отказов ценно само по себе. Но еще важнее дополнить это знание некоторым методом, позволяющим определять надвигающиеся отказы. Иными словами, полезно знать, что на следующей неделе ожидается отказ четырех изделий из каждых десяти, но куда лучше было бы знать, какие именно четыре изделия выйдут из строя.

Может оказаться, что в предвидении возможного повреждения четырех изделий из десяти экономически оправдана замена всех десяти. Однако если известно, какие именно четыре изделия потенциально могут выйти из строя, то только их и надо заменять, а остальные шесть можно оставить на месте по крайней мере на следующую неделю. Очевидно, знание метода, позволяющего предсказать отказы, уменьшает количество замен, необходимых для того, чтобы уровень отказов не превышал заданный. Сокращение числа замен и отказов — это выигрыш, связанный с применением контроля. Этот выигрыш увеличивается по мере улучшения метода, позволяющего отделить еще годные изделия от тех, которым угрожает близкий отказ. Получаемая от этого прибыль может с лихвой перекрыть затраты, связанные с контролем. В действительности результатом контроля является скорее оценка для вероятности отказа, а не конкретный прогноз того, будет работать изделие в ближайший период или же сломается. Остается определить еще максимальную величину вероятности отказа, при которой изделие все же не подлежит замене.

В следующем разделе мы в явном виде введем затраты, связанные с заменой изделия и его оставлением, и построим стратегию, минимизирующую средние расходы в зависимости от стоимости замены, стоимости отказа и вероятности отказа.

Проблема замены действующего оборудования будет иллюстрироваться задачей о групповой замене всех электрических ламп в некоторой установке через определенные промежутки времени. Существует некоторый интервал, обеспечивающий минимум затрат при условии, что общая стратегия складывается из групповых замен и ликвидации отдельных повреждений ламп в промежутке времени между групповыми заменами. Наша задача — найти длительность этого оптимального интервала. Мы будем рассматривать вероятность отказа, затраты, связанные с отказом, и расходы, связанные с заменой, выведем общее уравнение затрат и найдем оптимальную длину интервала времени между групповыми заменами.

Кривые выхода из строя. Начальная информация о характеристиках срока службы электрической лампы может быть изображена в виде кривой выхода из строя. Пусть установленная группа из N ламп, и пусть в конце промежутка времени, состоящего из t равных интервалов, $S(t)$ ламп еще пригодны для эксплуатации. Эти годные лампы составляют некоторую долю

Таблица 17.4

**Характеристики срока службы электрической лампы:
исходная популяция состоит из 100 000 единиц**

(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Число прошедших интервалов t	Число ламп, оставшихся годными, $S(t)$	Количество ламп, испортившихся за 1 период, $S(t-1) - S(t)$	Вероятность поломки $p(t)$	Условная вероятность поломки $v_t, 0$	Число прошедших интервалов t	Число ламп, оставшихся годными, $S(t)$	Количество ламп, испортившихся за 1 период, $S(t-1) - S(t)$	Вероятность поломки $p(t)$	Условная вероятность поломки $v_t, 0$
0	100 000				10	48 000	15 000	0,15	0,2381
1	100 000	0	0	0	11	32 000	16 000	0,16	0,3333
2	99 000	1 000	0,01	0,0100	12	18 000	14 000	0,14	0,4375
3	98 000	1 000	0,01	0,0101	13	10 000	8 000	0,08	0,4444
4	97 000	1 000	0,01	0,0102	14	6 000	4 000	0,04	0,4000
5	96 000	1 000	0,01	0,0103	15	3 000	3 000	0,03	0,5000
6	93 000	3 000	0,03	0,312	16	2 000	1 000	0,01	0,3333
7	87 000	6 000	0,06	0,645	17	1 000	1 000	0,01	0,5000
8	77 000	10 000	0,10	0,1149	18	0	1 000	0,01	1,0000
9	63 000	14 000	0,14	0,1818					

Колонка (1): число прошедших периодов.

Колонка (2): количество ламп, годных к концу периода. Эти числа найдены на основе данных, представленных главными производителями ламп.

Колонка (3): скорость изменения данных колонки (2).

Колонка (4): частное от деления числа из колонки (3) на 100 000.

Колонка (5): число из колонки (3), деленное на число из колонки (2) за предыдущий период.

$s(t) = \frac{S(t)}{N}$ от первоначального их числа. Таблица 17.4 является типичным примером таблицы выхода из строя. В ней приведены данные, показывающие, как изменяется количество годных ламп с течением времени.

Исходное число ламп равно 100 000, интервалы времени имеют одинаковую длительность.

Этой таблице соответствует график рис. 17.2, на котором по оси абсцисс отложено число прошедших интервалов времени, а по оси ординат — количество годных ламп. Получившаяся кривая $S(t)$ и есть кривая выхода из строя. Не вдаваясь в тонкости изучаемых здесь задач математической статистики, будем полагать, что относительное число ламп, годных к моменту t , равно вероятности того, что единственная новая лампа, выбранная случайным образом, остается годной в течение всего интервала времени.

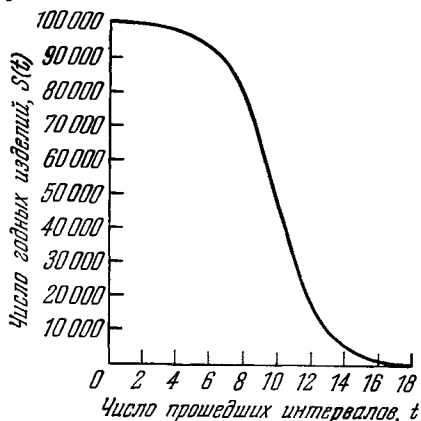


Рис. 17.2. Число годных изделий после t интервалов времени (по данным табл. 17.4).

Длительность срока службы. Наиболее удобной эксплуатационной характеристикой группы ламп, пожалуй, является закон распределения вероятностей для срока службы. Это распределение может быть найдено по

таблице выхода из строя, если положить вероятность отказа равной отношению числа изделий, вышедших из строя в течение t -го периода времени, т. е.

$$p(t) = \frac{1}{N} [S(t-1) - S(t)].$$

Эти значения легко подсчитать. Сначала находят первые разности, т. е. число деталей, сломавшихся в течение каждого периода времени (колонка (3) табл. 17.4). Затем эти числа делят на 100 000 и полученные результаты (колонка (4)) полагают равными соответствующим вероятностям поломки. График зависимости вероятности отказа от времени t приведен на рис. 17.3.

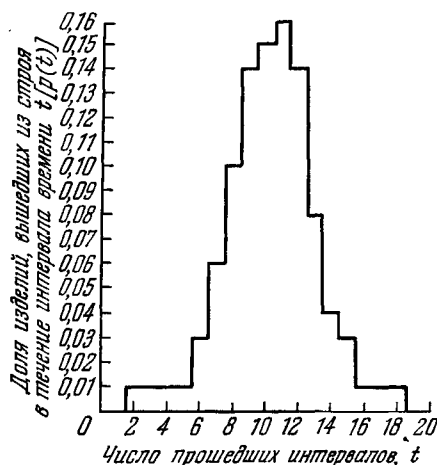


Рис. 17.3. Вероятность того, что лампа, поставленная в начале первого интервала времени, выйдет из строя в течение t -го интервала (по данным табл. 17.4).

Условная вероятность поломки.

Другой важной характеристикой срока службы изделия является условная вероятность поломки либо его дополнение — вероятность того, что изделие, годное к моменту времени t , останется годным и к моменту $t + 1$. Эта величина также может быть получена из таблицы отказов, если вычислить отношение числа изделий, сломавшихся в течение некоторого периода, к числу изделий, годных к началу этого периода. Иными словами,

$$v_{t,0} = \frac{S(t-1) - S(t)}{S(t-1)} = 1 - \frac{S(t)}{S(t-1)}.$$

График условной вероятности приведен на рис. 17.4.

Процесс замены. Впредь, до конца этой главы, будем предполагать, что отказы приходится на конец периода. Следовательно, детали, которые были поставлены, скажем, в конце третьего периода времени вместо сломанных, имеют к началу четвертого периода нулевую длительность срока службы. Это предположение на данном этапе избавляет нас от трудностей, связанных с учетом частичных периодов или непрерывных изменений во времени. Как отмечалось выше, в течение первых t интервалов времени все отказы устранялись по мере их появления. В конце t -го интервала происходит замена всех изделий независимо от их срока службы. Задача состоит в том, чтобы найти значение t , минимизирующее общие затраты.

Скорость замены. I метод. Процесс замены происходит в несколько стадий. Вначале подлежат замене сломанные изделия из числа первоначально установленных, затем некоторые из изделий-заменителей* и, наконец (когда все первоначальные изделия будут исчерпаны), те из заменителей, которые успели выйти из строя в течение самого процесса замены.

* Назовем «заменителем» (или «изделием-заменителем») некоторое изделие, поставленное взамен эксплуатируемого в данный момент. Будем называть «заменителем первого порядка» изделие, поставленное на первом этапе процесса замены, «заменителем второго порядка» — поставленное на втором этапе и т. д. до любого порядка. (Прим. перев.)

Общее выражение для числа деталей, вышедших из строя в течение t -го промежутка времени, имеет вид

$$f(t) = N \left\{ p(t) + \sum_{x=1}^{t-1} p(x) p(t-x) + \sum_{b=2}^{t-1} \left[\sum_{x=1}^{b-1} p(x) p(b-x) \right] p(t-b) + \dots \right\}, \quad (12)$$

где N — общее число деталей в установке, $p(x)$ — вероятность отказа детали, срок службы которой равен x . Выражение, стоящее в формуле (12) в квадратных скобках, содержит члены, определяющие вероятность первого, второго, третьего и т. д. отказов любого из изделий в течение t -го интервала времени. Очевидно, $p(t)$ — вероятность первого отказа, найденная ранее с помощью данных, аналогичных данным таблицы 17.4. Второй член в правой части выражения, $\sum_{x=1}^{t-1} p(x) p(t-x)$, есть вероятность того, что в течение t -го интервала изделие, поставленное взамен отказавшего ранее, также выйдет из строя. Она равна совместной вероятности двух независимых событий, состоящих в том, что исходное изделие выйдет из строя в момент x , а заменившее его — в момент $t-x$.

Для того чтобы лучше уяснить смысл этой суммы, распишем ее подробней. Пусть, например, $t=6$, т. е. рассматривается шестой интервал. Тогда

$$\sum_{x=1}^5 p(x) p(6-x) = p(1) p(5) + p(2) p(4) + p(3) p(3) + p(4) p(2) + p(5) p(1).$$

Каждый из этих членов представляет собою вероятность события, приводящего ко второму отказу в пределах именно шестого периода.

Вероятность третьего отказа в течение t -го периода времени определяется третьим членом в квадратных скобках в формуле (12) и вычисляется следующим образом. Вероятность второго отказа в $(t-b)$ -м промежутке умножается на вероятность отказа в b -м промежутке времени, и эти произведения суммируются по всем b , меньшим, чем t . Справедливость этого доказывается с помощью тех же рассуждений, что и в предыдущем случае. Аналогично могут быть вычислены вероятности четвертого, пятого и т. д. отказов. В результате находится общее выражение для $f(t)$. Теперь целесообразно было бы рассмотреть численный пример, — мы сделаем это, опираясь на значения $p(t)$, приведенные в таблице 17.4.

Пусть в начальный момент времени имеется 100 000 изделий. Из колонки (4) таблицы 17.4 следует, что в первом периоде отказов нет (в этой таблице приведено распределение вероятностей для срока службы изделий). Во втором периоде из строя выходит 1000 изделий, или 1% исходного их числа. В третьем периоде ломается еще 1% первоначального числа ламп. Однако для 1000 заменителей этот период фактически является первым, и потому ни один из них не выходит из строя. В четвертом периоде ломается 1% перво-

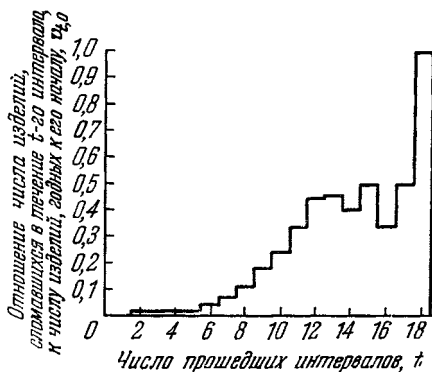


Рис. 17.4. Условная вероятность выхода из строя в течение t -го интервала (по данным табл. 17.4).

начальных изделий и 1% заменителей первого порядка. Следовательно, всего понадобится 1010 заменителей второго порядка.

Таким образом, изделия, установленные в течение любого интервала времени, будут заменяться в каждом из 18 периодов, следующих за установкой, и, таким образом, увеличивать общую скорость отказов в каждом из этих периодов.

Для изучаемого случая данные о скорости отказов в течение каждого из интервалов времени приведены в таблице 17.5. Так, например, число

Таблица 17.5 *

Общее число отказов (замен) в каждом из периодов

(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Пе- риод	Теку- щие заме- ны [$f(t)$]	Общие замены [$\Sigma f(t)$]	Пе- риод	Теку- щие заме- ны [$f(t)$]	Общие замены [$\Sigma f(t)$]	Пе- риод	Теку- щие заме- ны [$f(t)$]	Общие замены [$\Sigma f(t)$]	Пе- риод	Теку- щие заме- ны [$f(t)$]	Общие замены [$\Sigma f(t)$]
1	0	0	11	16 665	69 448	21	12 047	162 167	31	10 413	258 689
2	1 000	1 000	12	15 000	84 448	22	11 706	173 873	32	10 507	269 196
3	1 000	2 000	13	9 480	93 928	23	10 820	184 693	33	10 348	279 544
4	1 010	3 010	14	6 175	100 103	24	9 697	194 390	34	9 999	289 543
5	1 020	4 030	15	6 160	106 263	25	8 700	203 090	35	9 636	299 179
6	3 030	7 060	16	5 521	111 784	26	8 288	211 378	36	9 079	308 258
7	6 040	13 100	17	7 309	119 093	27	8 413	219 791	37	9 220	317 478
8	10 090	23 190	18	9 317	128 410	28	8 862	228 653	38	9 271	326 749
9	14 201	37 391	19	10 181	138 591	29	9 523	238 176	39	9 447	336 196
10	15 392	52 783	20	11 529	150 120	30	10 100	248 276	40	9 669	345 865

Колонка (1): периоды после первоначальной установки.

Колонка (2): рассчитывается так, как описано в тексте.

Колонка (3): общая сумма значений колонки (3).

* Данные заимствованы из таблицы 17.4.

отказов в течение двадцатого периода складывается из 1% числа отказов в восемнадцатом периоде, 1% — в семнадцатом, 1% — в шестнадцатом, 1% — в пятнадцатом, 3% — в четырнадцатом, 6% — в тринадцатом и т. д.

Этот расчет выполнен с учетом вероятности поломки в течение времени, прошедшего от момента установки изделия до двадцатого интервала.

На рис. 17.5 изображен график, на котором нанесены кривые зависимости числа отказов от номера интервала. Одна из них построена с учетом замен, другая относится лишь к первоначально установленным изделиям. Интересно отметить, что в течение первых периодов эти кривые очень близки друг к другу. Однако по мере того, как все большую роль начинают играть заменители второго и более высоких порядков, расхождение между ними резко увеличивается.

Известный интерес представляют также колебательный характер кривой скорости отказов и ее постепенная сходимость к некоторому постоянному значению. Как доказано в [17, стр. 276], предельная скорость отказов равна частному от деления числа установленных деталей на среднюю длительность срока службы. Можно показать, что в данном случае средняя продолжительность срока службы составляет 10,3 единицы времени. Поэтому предельная скорость равна $100\,000/10,3 = 9709$ заменам в течение одного

периода. Эта скорость используется далее для нахождения условий, при которых групповая замена вообще нецелесообразна.

При расчете числа отказов в данном случае (табл. 17.5) предполагалось, что повреждение, случившееся в t -м периоде, устраняется лишь в начале $(t+1)$ -го периода. Это предположение ведет к очень малому смещению

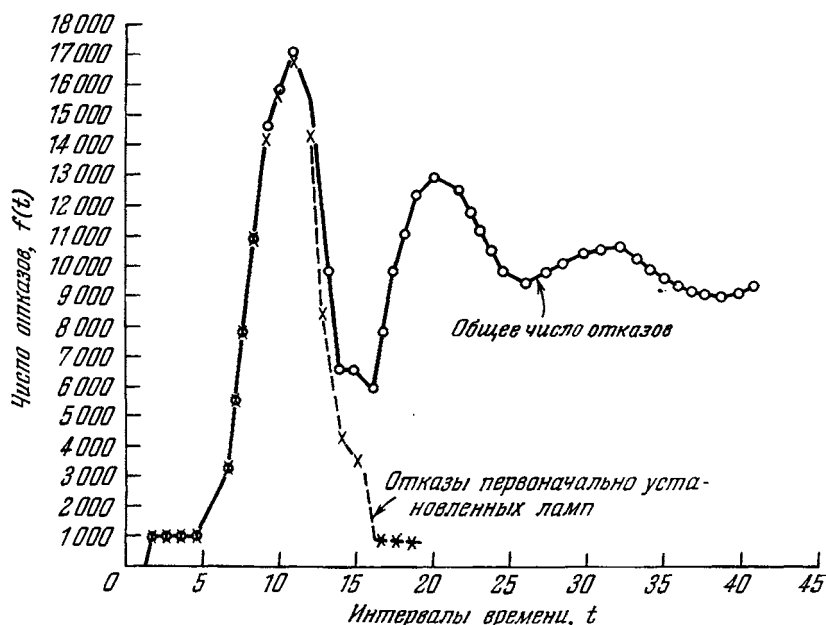


Рис. 17.5. Количество вышедших из строя ламп из числа 100 000 первоначально установленных и общее число отказов в течение каждого интервала t (по данным таблиц 17.4 и 17.5).

в оценке общего числа отказов даже в том случае, когда периоды времени обладают значительной длительностью. Более точные оценки могут быть получены с помощью метода Монте-Карло. Однако это потребовало бы огромного объема выборки, что, конечно, неосуществимо без применения электронных вычислительных машин. На рис. 17.6 для сравнения изображены две кривые распределения числа отказов (по оси ординат отложено количество поломок, выраженное в процентах от числа первоначально установленных изделий). Одна из них найдена теоретически, относящиеся к ней данные приведены в колонке (3) таблицы 17.5, а сама она изображена на рис. 17.5. Вторая вычислена методом Монте-Карло по 300 реализациям.

Скорость замены. II метод. Задачу определения скорости замены целесообразно рассмотреть и с несколько иной точки зрения. Напомним, что мы вычисляли условную вероятность отказа в определенные моменты времени.

Предположим, что закон, по которому 100 000 изделий распределяются в зависимости от срока их службы, известен для каждого момента времени. Тогда можно найти среднее число отказов в следующем периоде. Для этого все изделия надо разбить на группы в соответствии со сроком их службы, умножить число деталей в каждой группе на вероятность отказа, соответствующую этой группе, и все произведения сложить. С точки зрения вычислений этот подход является громоздким и нерациональным, поскольку он требует знания закона распределения в конце каждого периода времени.

(Метод, рассмотренный в предыдущем разделе, требовал лишь знания числа отказов в каждом периоде.) Однако он дает возможность уяснить суть некоторых фундаментальных понятий, с которыми нам пришлось ранее столкнуться.

Рассмотрим вектор-строку A_i , состоящую из 18 элементов, каждый из которых равен числу изделий со сроком службы 0, 1, 2, ..., 17 соответственно к началу i -го периода. Можно построить матрицу, которая преобразует этот временной вектор в начале любого периода в аналогичный вектор в конце того же периода. Обозначим эту матрицу $V = \{v_{ij}\}$, где v_{ij} — вероятность того, что изделие со сроком службы i превратится в течение следующего периода в изделие со сроком службы j , где i, j пробегают значения от 0 до 17. Ясно, что изделие либо сломается, либо срок службы его увеличится на один период. Сломанное изделие заменяется новым и потому, в известном смысле, превращается в изделие с нулевым сроком службы. Таким образом, при $j \neq 0$ и $j \neq i + 1$ $v_{ij} \equiv 0$ для всех i . С другой стороны, при $j = 0$ или $j = i + 1$ имеем $v_{ij} \geq 0$.

В течение i -го периода изделие остается годным либо выходит из строя. Так как эти исходы образуют полную группу несовместных событий, то $v_{i,0} + v_{i,i+1} = 1$. Физический смысл $v_{i,0}$ — вероятность отказа изделия со сроком службы i в течение i -го периода времени, т. е. это не что иное, как условная вероятность, приведенная в таблице 17.4. Используя эту таблицу, в нашем примере можно построить следующую матрицу преобразований:

$$V = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,0100 & 0 & 0,9900 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,0101 & 0 & 0 & 0,9899 & 0 & 0 & 0 & 0 & 0 \\ 0,0102 & 0 & 0 & 0 & 0,9898 & 0 & 0 & 0 & 0 \\ 0,0103 & 0 & 0 & 0 & 0 & 0,9897 & 0 & 0 & 0 \\ 0,0312 & 0 & 0 & 0 & 0 & 0 & 0,9688 & 0 & 0 \\ 0,0645 & 0 & 0 & 0 & 0 & 0 & 0 & 0,9355 & 0 \\ 0,1149 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,8851 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}.$$

Пусть A_1 — вектор-строка, задающая распределение по строкам служб изделий в начале первого периода времени, тогда

$$A_2 = A_1 V$$

задает это распределение в начале второго периода,

$$A_3 = A_2 V = (A_1 V) V = A_1 V^2$$

— в начале третьего периода и т. д. В общем случае распределение числа деталей в зависимости от срока их службы к началу n -го периода задается выражением

$$A_n = A_1 V^{n-1}. \quad (13)$$

Пусть, например, в начальный момент времени выборка составляет 100 000 новых деталей. Тогда

$$A_1 = \| 100\,000 \quad 0 \quad 0 \quad 0 \dots 0 \|,$$

т. е. к началу первого периода имеется 100 000 деталей с нулевым сроком службы.

По правилам умножения матриц

$$A_2 = A_1 V = \parallel 0 \ 100\ 000 \ 0 \ 0 \ \dots \ 0 \parallel,$$

т. е. к началу второго периода времени все 100 000 деталей отработали один период и ни одна из них не сломалась.

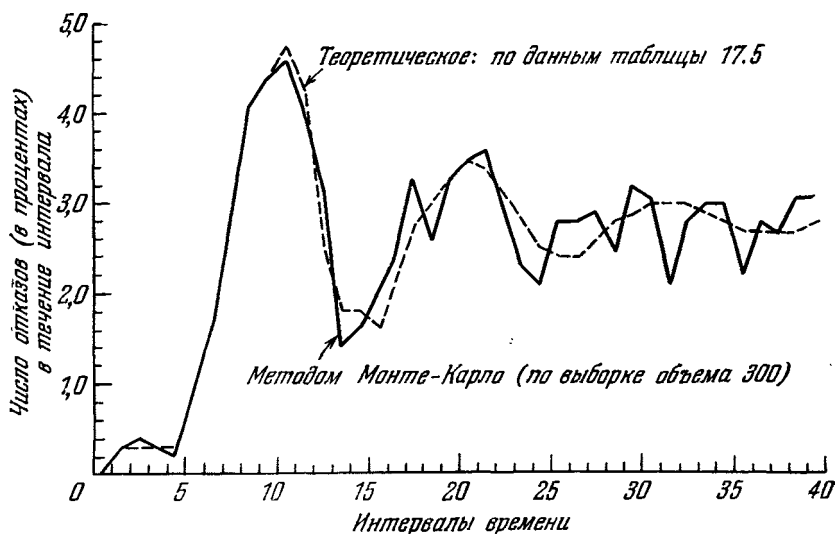


Рис. 17.6. Распределение числа отказов, найденное теоретически и с помощью метода Монте-Карло.

Продолжая далее, находим

$$A_3 = A_2 V = \parallel 1000 \ 0 \ 99\ 000 \ 0 \ 0 \ \dots \ 0 \parallel,$$

т. е. к началу третьего периода 99 000 деталей обладают сроком службы длительностью в два периода, а остальные 1000 пришли в негодность и были заменены деталями с нулевым сроком службы.

Далее находится A_4

$$A_4 = A_3 V = \parallel 1010 \ 1000 \ 0 \ 97900 \ 0 \ \dots \ 0 \parallel \text{ и т. д.}$$

Квадратная матрица V состоит из неотрицательных элементов, обладающих тем свойством, что сумма членов, стоящих в одной строке, равна единице (так как $v_{i,0} + v_{i,i+1} = 1$). Поэтому высшие степени матрицы V сходятся к некоторой матрице V^* [21]. Следовательно, для достаточно больших n можно записать

$$A_n = A_1 V^* \text{ и } A_{n+1} = A_1 V^*,$$

так что

$$A_n = A_{n+1} \text{ (для достаточно больших } n\text{).} \quad (14)$$

Из этого результата следует существование стационарного предельного распределения изделий по группам в зависимости от срока их службы. В это распределение входит также и стационарное количество изделий с нулевым сроком службы, равное установившейся скорости отказов. Можно отметить еще одну интересную связь между этим стационарным распреде-

нием и кривой выхода из строя. Выше было показано, что для достаточно больших n $A_n = A_{n+1}$. Однако по определению $A_{n+1} = A_n V$. Поэтому для достаточно больших n

$$A_n = A_n V,$$

так что

$$A_n [I - V] = 0,$$

где I — единичная матрица, диагональные элементы которой равны единице, а все остальные — нулю. Эта матрица обладает следующим свойством:

$$IV = VI = V.$$

В первом столбце матрицы V стоят элементы $v_{i,0}$ — условная вероятность поломки в i -м периоде. В каждой строке, кроме первой, имеется еще лишь один ненулевой элемент $v_{i,i+1}$ — условная вероятность того, что в течение i -го периода времени изделие остается годным. Ясно, что

$$v_{i,i+1} + v_{i,0} = 1 \quad (15)$$

или

$$v_{i,i+1} = 1 - v_{i,0}.$$

Следовательно, матрица $I - V$ имеет вид

$$I - V = \begin{pmatrix} 1 - v_{1,0} & -(1 - v_{1,0}) & 0 & 0 & 0 \\ -v_{2,0} & 1 & -(1 - v_{2,0}) & 0 & 0 \\ -v_{3,0} & 0 & 1 & -(1 - v_{3,0}) & 0 \\ -v_{4,0} & 0 & 0 & 1 & -(1 - v_{4,0}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \quad (16)$$

Расписывая матричное уравнение $A_n (I - V) = 0$, получаем систему уравнений

$$\left. \begin{aligned} a_0 (1 - v_{1,0}) - a_1 v_{2,0} - a_2 v_{3,0} - \dots &= a_i v_{i+1,0} - \dots = 0, \\ -a_0 (1 - v_{1,0}) + a_1 &= 0, \\ -a_1 (1 - v_{2,0}) + a_2 &= 0, \\ \vdots & \\ -a_{i-1} (1 - v_{i,0}) + a_i &= 0, \\ \dots & \end{aligned} \right\} \quad (17)$$

где a_i — элементы вектора A_n . Перепишем первое уравнение этой системы следующим образом:

$$a_0 = a_0 v_{1,0} + a_1 v_{2,0} + \dots + a_i v_{i+1,0} + \dots \quad (18)$$

Из этого уравнения еще раз вытекает, что среднее число замен в течение некоторого периода времени может быть найдено, если число изделий с данным сроком службы умножить на вероятность их отказа в течение следующего периода и полученные произведения суммировать по всем значениям срока службы.

Остальные уравнения этой системы могут быть переписаны в следующем виде:

$$a_1/a_0 = 1 - v_{1,0}, \quad a_2/a_1 = 1 - v_{2,0}, \quad \dots, \quad a_i/a_{i-1} = 1 - v_{i,0}, \quad \dots \quad (19)$$

В предыдущем изложении условная вероятность $v_{i,0}$ вычислялась как частное от деления числа отказов в i -м периоде на общее количество

годных деталей со сроком службы $i-1$ к началу этого периода, т. е.

$$v_{i,0} = \frac{S(i-1) - S(i)}{S(i-1)}. \quad (20)$$

Отсюда

$$1 - v_{i,0} = 1 - \frac{S(i-1) - S(i)}{S(i-1)} = \frac{S(i)}{S(i-1)}. \quad (21)$$

Однако выше было показано, что

$$1 - v_{i,0} = a/a_{i-1}$$

и потому

$$a_i/a_{i-1} = S(i)/S(i-1). \quad (22)$$

Из этого соотношения следует, что относительное предельное распределение группы деталей, каждое из которых заменяется лишь после его отказа, может быть предсказано по таблице выхода из строя.

Рассмотрим электрические лампы. Известно, что уровень светового потока отдельной лампы падает по мере ее эксплуатации. Поэтому для нахождения светового потока, создаваемого группой ламп с различными сроками службы, необходимо знать их предельное распределение. Таким образом, эта связь между предельным распределением числа ламп в зависимости от срока их службы и кривой выхода из строя позволяет оценить уровень светового потока для случая, когда групповые замены отсутствуют.

Стоимость замены. Второе основное требование, предъявляемое к любой полезной стратегии замен, состоит в том, что затраты, связанные с заменой оборудования после его отказа, должны превышать затраты, связанные с его заменой до отказа. Эта разница в издержках и оправдывает замену еще годного оборудования с целью уменьшить вероятность его выхода из строя. Существует целый ряд причин, по которым в различных конкретных производственных ситуациях выгоднее заменять еще работающее оборудование, а не ждать его отказа. Например, издержки, связанные с заменой сломавшегося оборудования, складываются из стоимости самого оборудования, стоимости труда, стоимости продукции, потерянной в результате простоя, а также затрат, вызванных повреждением материала и оборудования в результате пожара, аварии или других случайных причин. Более того, общая сумма этих затрат может не быть постоянной для каждого отказа, а зависеть от числа отказов в течение каждого интервала времени. Это имеет место, например, при простоях обслуживающего оборудования, когда число отказов, приходящееся на один период времени, достаточно велико. Будем, однако, предполагать, что в нашем примере групповой замены электрических ламп расходы, связанные с ликвидацией отказов, остаются постоянными.

Может оказаться, что благодаря экономии труда, скидкам при закупке партий материалов, а также по другим причинам групповая замена обходится дешевле, нежели ликвидация поломок. Достаточно лишь определить, что затраты, приходящиеся на одно изделие при групповой замене, постоянны и меньше аналогичных расходов, связанных с ликвидацией отдельных поломок.

Уравнение затрат. Выведем уравнение, определяющее затраты на эксплуатацию системы как функцию от управляемой величины t — числа периодов между двумя групповыми заменами. Пусть $K(t)$ — общие затраты,

приходящиеся на весь интервал времени между моментом групповой установки и концом t -го периода. Если в течение интервала длиной t заменяется вся группа изделий, то в среднем на один период времени приходится $K(t)/t$ затрат. Пусть, далее, C_1 — затраты, приходящиеся на одно изделие при групповой замене, C_2 — затраты, связанные с ликвидацией отдельного отказа, $f(X)$ — число отказов в течение X -го периода, N — число изделий в группе. Тогда общие затраты $K(t)$ определяются формулой

$$K(t) = NC_1 + C_2 \sum_{x=1}^t f(X),$$

где NC_1 — стоимость групповой замены ламп, а $C_2 \sum_{x=1}^{t-1} f(X)$ — затраты, связанные с ликвидацией отдельных отказов в конце каждого из $t-1$ периодов между групповыми заменами. Таким образом, затраты, приходящиеся на один период, составляют

$$\frac{K(t)}{t} = \frac{N(C_1)}{t} + \frac{C_2}{t} \sum_{x=1}^{t-1} f(X). \quad (23)$$

В отличие от предыдущих разделов данной главы при выводе этого уравнения мы использовали затраты, приходящиеся на один период, а не прямую сумму всех будущих расходов. Вообще говоря, особой необходимостью это не вызывалось, так как длительность интервала времени здесь предполагалась относительно небольшой, и потому эффект, например, учетного процента был бы, соответственно, незначительным. В тех задачах, где срок службы оборудования имеет большую длительность, в уравнение должны быть введены соответствующим образом все факторы, влияющие на затраты.

Минимизация затрат. Стратегия, приводящая к групповым заменам после $\hat{t}$ периодов, обеспечивает минимум затрат в том случае *, если

$$\left. \begin{aligned} & \frac{K(t)}{t} < \frac{K(\hat{t}-1)}{\hat{t}+1} \\ \text{или, что эквивалентно,} & \frac{K(\hat{t}+1)}{\hat{t}+1} - \frac{K(\hat{t})}{\hat{t}} > 0 \\ \text{и если} & \frac{K(\hat{t})}{\hat{t}} < \frac{K(\hat{t}-1)}{\hat{t}-1} \\ \text{или, что эквивалентно,} & \frac{K(\hat{t}-1)}{\hat{t}-1} - \frac{K(\hat{t})}{\hat{t}} > 0. \end{aligned} \right\} \quad (24)$$

Перепишем уравнение затрат (23) в следующем виде:

$$\frac{K(\hat{t}+1)}{\hat{t}+1} = \frac{NC_1}{\hat{t}+1} + \frac{C_2}{\hat{t}+1} \sum_{x=1}^{\hat{t}} f(X), \quad (25a)$$

* Эти условия являются необходимыми, но недостаточными. Рассмотрим, например, функцию $F(t) = t \sin t$, $0 \leq t \leq 4\pi$, удовлетворяющую этим условиям даже при двух значениях t , хотя в действительности она обладает лишь одной точкой глобального минимума.

где $t = \hat{t} + 1$. Тогда при $t = \hat{t}$

$$\frac{K(\hat{t})}{\hat{t}} = \frac{NC_1}{\hat{t}} + \frac{C_2}{\hat{t}} \sum_{X=1}^{\hat{t}-1} f(X). \quad (256)$$

Вычитая одно уравнение из другого, находим

$$\begin{aligned} \frac{K(\hat{t}+1)}{\hat{t}+1} - \frac{K(\hat{t})}{\hat{t}} &= NC_1 \left(\frac{1}{\hat{t}+1} - \frac{1}{\hat{t}} \right) + C_2 \sum_{X=1}^{\hat{t}-1} f(X) \left(\frac{1}{\hat{t}+1} - \frac{1}{\hat{t}} \right) + C_2 \frac{f(\hat{t})}{\hat{t}+1} = \\ &= \frac{-NC_1 - C_2 \sum_{X=1}^{\hat{t}-1} f(X) + C_2 \hat{t} f(\hat{t})}{(\hat{t}+1)\hat{t}}. \end{aligned} \quad (26)$$

Для того чтобы левая часть уравнения (26) была положительной, необходимо выполнение следующего условия:

$$C_2 \hat{t} f(\hat{t}) > NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)$$

или

$$C_2 f(\hat{t}) > \frac{NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)}{\hat{t}}. \quad (27)$$

Точно так же можно найти, что требование

$$K(\hat{t}-1)/(\hat{t}-1) > K(\hat{t})/\hat{t}$$

приводит к неравенству

$$C_2 f(\hat{t}-1) < \frac{NC_1 + C_2 \sum_{X=1}^{\hat{t}-2} f(X)}{\hat{t}-1}. \quad (28)$$

Неравенства (27) и (28) определяют необходимые условия, которым должна отвечать оптимальная стратегия групповой замены. Им можно дать следующее толкование: $[NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)]/\hat{t}$ — затраты, приходящиеся в среднем на один период при условии, что все лампы заменяются по окончании $\hat{t}$ периодов; $C_2 f(\hat{t})$ — затраты, приходящиеся на $\hat{t}$ -й период при условии, что в конце его не производится групповая замена. Иными словами, ликвидация $f(\hat{t})$ отдельных отказов с затратами C_2 на каждую из них осуществляется в том и только в том случае, если на этот период не приходится групповая замена.

Таким образом, из неравенства (27) следует, что групповая замена в конце $\hat{t}$ -го периода целесообразна в том случае, если затраты по ликвидации отдельных отказов в течение $\hat{t}$ -го периода превзойдут средние расходы на один период в течение всего промежутка времени длиной $\hat{t}$.

Аналогично из неравенства (28) вытекает, это в том случае, когда затраты по ликвидации отдельных отказов в конце $\hat{t}$ -го периода ниже, чем средние расходы на один период в течение всего времени $\hat{t}$, групповая замена в конце $\hat{t}$ -го промежутка нецелесообразна.

Сейчас будет показано, каким образом можно найти оптимальное для групповой замены значение $\hat{t}$ при условии, что отношение $C_1/C_2 = 0,25$. Знание конкретных численных значений для отдельных групп затрат несущественно. В этом легко убедиться, если разделить соответствующие неравенства на C_2 . С учетом этого неравенства (27) и (28) принимают вид

$$f(\hat{t}) > \frac{N(C_1/C_2) + \sum_{X=1}^{\hat{t}-1} f(X)}{\hat{t}} \quad (29)$$

и

$$f(\hat{t}-1) < \frac{N(C_1/C_2) + \sum_{X=1}^{\hat{t}-2} f(X)}{\hat{t}-1}. \quad (30)$$

В данном случае $N(C_1/C_2) = 25\,000$. Следовательно, ищется такое значение t , для которого

$$\left. \begin{aligned} \hat{t}f(\hat{t}) - \sum_{X=1}^{\hat{t}-1} f(X) &> 25\,000, \\ (\hat{t}-1)f(\hat{t}-1) - \sum_{X=1}^{\hat{t}-2} f(X) &< 25\,000. \end{aligned} \right\} \quad (31)$$

В таблице 17.6 выписаны значения, которые принимают при различных t слагаемые, стоящие в левых частях обоих неравенств (31). Кроме

Таблица 17.6

Средние расходы при различных стратегиях групповой замены:
 $C_1/C_2 = 0,25$ (данные заимствованы из таблицы 17.5).

(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	0	0	0	0	25 000C ₂	25 000C ₂
2	1 000	1 000	0	2 000	25 000C ₂	12 500C ₂
3	1 000	2 000	2 000	2 000	26 000C ₂	86 667C ₂
4	1 010	3 010	2 000	2 040	27 000C ₂	6 750C ₂
5	1 020	4 030	2 040	2 090	28 010C ₂	5 602C ₂
6	3 030	7 060	2 090	14 150	29 030C ₂	4 839C ₂
7	6 040	13 100	14 150	35 220	32 060C ₂	4 580C ₂ *
8	10 090	23 190	35 220	67 620	38 100C ₂	4 762C ₂
9	14 201	37 391	67 620	104 619	48 190C ₂	5 354C ₂
10	15 392	52 783	104 619	116 529	62 391C ₂	6 239C ₂

Колонка (1): число периодов между групповыми заменами.

Колонка (2): число замен (табл. 17.5).

Колонка (3), (4), (5): вычислены по формулам, приведенным в начале колонок.

Колонка (6): вычислена по формуле при $C_1 = 0,25 C_2$.

Колонка (7): частное от деления данных колонки (6) на данные колонки (1).

* Поэтому $\hat{t} = 7$.

того, здесь также приведены затраты, вычисленные с помощью функции $K(t)$ при $C_1 = 0,25C_2$ и выраженные в единицах C_2 . Из таблицы видно,

что в конце седьмого периода групповая замена обеспечивает минимум средних расходов ($4580C_2$) и что оба неравенства (31) действительно удовлетворяются лишь при этом значении t , т. е. при $\hat{t} = 7$.

На рис. 17.7 изображен график зависимости общих расходов от числа t периодов между двумя групповыми заменами. Средние расходы вычисляются путем деления общих расходов на число периодов и равны угловому коэффициенту прямой, соединяющей начало координат с любой точкой кривой затрат. Поэтому стратегия замен, минимизирующая расходы, может быть найдена непосредственно из графика. Рис. 17.7 идентичен рис. 17.1 с точностью до будущих затрат, которые в данном примере не принимались во внимание.

Полученное здесь решение справедливо лишь при $C_1/C_2 = 0,25$. Было бы полезным найти его для любых значений C_1/C_2 . Отметим вначале, что для определения оптимального значения t в данном случае надо было сравнить числа в колонках (4) и (5) таблицы 17.6 со значением N (C_1/C_2). Разделив их на $N = 100\,000$, мы найдем новые величины, которые надо сравнить для нахождения оптимального t уже с C_1/C_2 . Эти величины приведены в таблице 17.7.

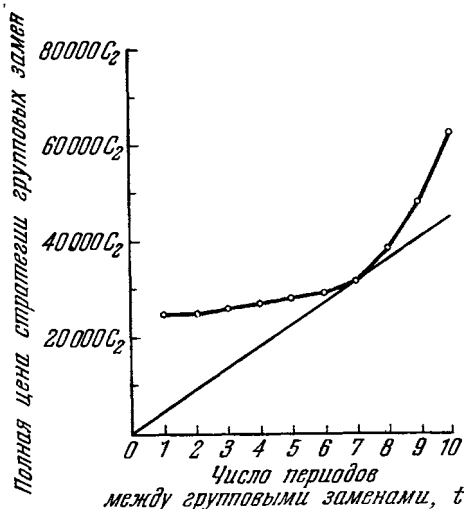


Рис. 17.7. Зависимость величины затрат от числа интервалов между заменами при различных стратегиях групповых замен $C_1/C_2 = 0,25$ (по данным табл. 17.6).

Таблица 17.7

Оптимальные значения t при различных отношениях затрат C_1/C_2

Величина отношения	$\hat{t}^*$	Величина отношения	$\hat{t}^*$
0	1	От 0,0209 до 0,1415	6
От 0 до 0,02	2	От 0,1415 до 0,3522	7
0,02	3	От 0,3522 до 0,6762	8
От 0,02 до 0,0204	4	От 0,6762 до 1,0462	9
От 0,0204 до 0,0209	5	От 1,0462 до 1,1653	10

* $\hat{t}$ — то значение t , которое обеспечивает минимум затрат.

Выше было показано, как определяется оптимальное число периодов между групповыми заменами. Однако еще не рассматривалась стратегия, в соответствии с которой устраняются лишь отдельные отказы, а групповые замены отсутствуют вообще.

Предположим, что число отказов $f(t)$, приходящееся на один период, стремится к $\bar{f}(t)$. Если при этих условиях групповые замены отсутствуют, то затраты, приходящиеся на один период, стремятся к $C_2 \bar{f}(t)$. Выше отмечалось (неравенство (27)), что для оптимального значения $\hat{t}$ имеет место соотношение

$$G_2 f(\hat{t}) > \frac{NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)}{\hat{t}}.$$

Очевидно, если $\bar{f}(t) > f(\hat{t})$, то $C_2 \bar{f}(t) > C_2 f(\hat{t})$, поэтому

$$C_2 \bar{f}(t) > \frac{NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)}{\hat{t}},$$

и групповая замена экономически целесообразна. Выведем необходимое условие нецелесообразности групповой замены. Из неравенства (28) следует, что *

$$C_2 f(\hat{t}-1) < \frac{NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)}{\hat{t}}.$$

Тогда при $\bar{f}(t) < f(\hat{t}-1)$

$$C_2 \bar{f}(t) < C_2 f(\hat{t}-1) < \frac{NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)}{\hat{t}},$$

и групповая замена нецелесообразна. Иными словами, издержки, связанные с заменой еще годного оборудования, никогда не окупятся за счет снижения числа отказов.

Наконец, если $\bar{f}(t) < f(\hat{t})$, а $f(\hat{t}-1) < \bar{f}(t)$, то результаты сравнения становятся менее очевидными и требуется более глубокое исследование экономической целесообразности групповой замены. В частности, величина $\hat{t}$ определяется кривой выживания и затратами C_1 и C_2 . Поэтому для одних значений C_1 и C_2 $f(\hat{t})$ может оказаться меньше, чем $\bar{f}(t)$, а для других — больше. Однако можно найти наибольшее значение C_1/C_2 , при котором групповая замена еще целесообразна. Это делается следующим образом.

Пусть затраты по ликвидации отдельных отказов и оптимальная стоимость групповой замены равны между собой. Тогда

$$\frac{NC_1 + C_2 \sum_{X=1}^{\hat{t}-1} f(X)}{\hat{t}} = C_2 \bar{f}(t).$$

* Это вытекает из неравенства: если $a/b < r < s$, то $a/b < (r+a)(s+b)$.

Отсюда

$$\left. \begin{aligned} NC_1 + C_2 \sum_1^{\hat{t}-1} f(X) &= \hat{t} C_2 \bar{f}(t), \\ NC_1 &= \hat{t} C_2 \bar{f}(t) - C_2 \sum_1^{\hat{t}-1} f(X), \\ \frac{C_1}{C_2} &= \frac{\hat{t} \bar{f}(t) - \sum_1^{\hat{t}-1} f(X)}{N}. \end{aligned} \right\} \quad (32)$$

Поскольку в данном числовом примере $\bar{f}(t) = 9709$, то оба условия $f(\hat{t}-1) < \bar{f}(t)$ и $\bar{f}(t) < f(\hat{t})$ удовлетворяются при $\hat{t} = 8$. Обращаясь к таблице 17.5, находим $6040 < 9709$ и $9709 < 10\,090$.

Подставляя в (32) численные значения, имеем

$$\frac{C_1}{C_2} = \frac{8 \cdot 9709 - 13\,100}{100\,000} = 0,64572 \approx 0,65.$$

Таким образом, при $C_1/C_2 > 0,65$ групповая замена вообще нецелесообразна.

Значения $\hat{t}$, отвечающие конкретным величинам C_1/C_2 , могут быть найдены из рис. 17.8. Для построения этой диаграммы требуется знание лишь $s(t)$ — кривой выхода из строя, и в случае электрических лампочек она легко может быть построена предпринимателем. Тогда каждый потребитель, зная свои затраты C_1 и C_2 , сумеет с помощью графика рис. 17.8 определить свою оптимальную стратегию.

Оценим полезность рис. 17.8 на примере расчета оптимального $\hat{t}$ при $C_1/C_2 = 0,10$. С помощью уравнения средних затрат (23)

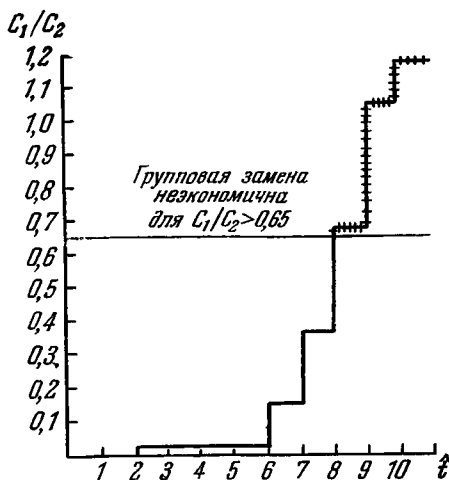


Рис. 17.8. Оптимальная стратегия замены при различных значениях C_1/C_2 (по данным табл. 17.7).

$$\frac{K(t)}{t} = \frac{NC_1 + C_2 \sum_{X=1}^{t-1} f(X)}{t} = \frac{100\,000 \cdot 0,10 C_2 + C_2 \sum_{X=1}^{t-1} f(X)}{t}$$

можно вычислить непосредственно расходы, связанные с каждой из стратегий, и выбрать наилучшую. Эти затраты приведены в таблице 17.8. Таким образом, в результате непосредственного вычисления находим, что при $\frac{C_1}{C_2} = 0,10$ $\hat{t} = 6$. Справедливость этого решения можно проверить с помощью таблицы 17.7 или рис. 17.8. Из обоих следует, что при $\frac{C_1}{C_2} = 0,10$ оптимальное число $\hat{t}$ периодов между групповыми заменами равно шести.

Таблица 17.8

Затраты $K(t)$ и $\frac{K(t)}{t}$ при различных t

t	$K(t)$	$K(t)/t$	t	$K(t)$	$K(t)/t$
1	$10\,000C_2$	$10\,000C_2$	6	$10\,000C_2 + 4\,030C_2$	$2\,338C_2$
2	$10\,000C_2 + 0C_2$	$5\,000C_2$	7	$10\,000C_2 + 7\,060C_2$	$2\,437C_2$
3	$10\,000C_2 + 1\,000C_2$	$3\,667C_2$	8	$10\,000C_2 + 13\,100C_2$	$2\,887C_2$
4	$10\,000C_2 + 2\,000C_2$	$3\,000C_2$	9	$10\,000C_2 + 23\,190C_2$	$3\,688C_2$
5	$10\,000C_2 + 3\,010C_2$	$2\,600C_2$			

Другие модели. Рассмотренные здесь решения относятся лишь к одной частной, описанной ранее модели. Однако точно так же исследуются и иные, отличные от нее типы моделей. Например, можно рассматривать модель, в соответствии с которой новые лампы используются лишь при групповой замене, а уже работавшие — для ликвидации отказов в интервалах между групповыми заменами. Иная модель нужна для случая, когда работающие лампы заменяются не в фиксированные моменты времени, а по истечении определенного срока службы.

Содержание данной главы было ограничено рассмотрением двух основных задач теории замен, одна из которых связана со старением изделия, а другая — со случайным характером срока службы оборудования.

Приложение

В этом приложении мы выведем соотношения (4) и (5). Пусть $X = 1/(1+r)$, так что

$$K_n = \frac{A + \sum_{i=1}^n C_i X^{i-1}}{1 - X^n}.$$

Заменяя в этом выражении n на $n+1$, находим

$$\begin{aligned} K_{n+1} &= \frac{A + \sum_{i=1}^{n+1} C_i X^{i-1}}{1 - X^{n+1}} = \frac{A + \sum_{i=1}^n C_i X^{i-1} + C_{n+1} X^n}{1 - X^{n+1}} = \\ &= \frac{(1 - X^n) K_n + C_{n+1} X^n}{1 - X^{n+1}} = \frac{1 - X^n}{1 - X^{n+1}} K_n + \frac{C_{n+1} X^n}{1 - X^{n+1}}. \end{aligned}$$

Поэтому

$$K_{n+1} - K_n = K_n \left(\frac{1 - X^n}{1 - X^{n+1}} - 1 \right) + \frac{C_{n+1} X^n}{1 - X^{n+1}} = \frac{K_n (X^{n+1} - X^n) + C_{n+1} X^n}{1 - X^{n+1}}.$$

Далее, если $K_{n+1} - K_n > 0$, то

$$[K_n (X^{n+1} - X^n) + C_{n+1} X^n] > 0,$$

откуда $X < 1$ или, что эквивалентно,

$$(1 - X^{n+1}) > 0.$$

Разделив последние выражения на X^n , находим

$$K_n (X - 1) + C_{n+1} > 0.$$

Отсюда

$$C_{n+1} > (1-X) K_n,$$

т. е.

$$C_{n+1}/(1-X) > K_n$$

или, что эквивалентно,

$$\frac{C_{n+1}}{1-1/(1+r)} > K_n,$$

что и требовалось доказать. Для того чтобы доказать справедливость второго неравенства, отметим, что из предыдущего следует

$$K_n - K_{n+1} = \frac{K_n(X^n - X^{n+1}) - C_{n+1}X^n}{1 - X^{n+1}}.$$

Заменяя n на $n-1$, находим

$$K_{n-1} - K_n = \frac{K_{n-1}(X^{n-1} - X^n) - C_n X^{n-1}}{1 - X^n}.$$

Поэтому, если $K_{n-1} - K_n > 0$, то

$$K_{n-1}(X^{n-1} - X^n) - C_n X^{n-1} > 0,$$

так как $X > 1$. Разделив все на X^{n-1} , имеем

$$(1-X) K_{n-1} - C_n > 0.$$

Отсюда

$$C_n/(1-X) < K_{n-1}.$$

Так как $X = 1/(1+r)$, то это эквивалентно

$$\frac{C_n}{1-1/(1+r)} < K_{n-1},$$

что и требовалось доказать.

ЛИТЕРАТУРА

1. Alchian A., Economic Replacement Policy, RAND Report No. R-224, Apr. 1952.
2. Bellman R., Notes in the Theory of Dynamic Programming — III: Equipment Replacement Policy, RAND Report No. P-632, Jan. 1955.
3. Benson C. B. and Kimball B. F., Mortality Characteristics of Physical Property Based Upon Location Life Table and Re-Use Ratios, *Econometrica*, 13, 214 (1945).
4. Blackwell D., Extension of a Renewal Theorem, *Pac. J. Math.*, 3, 315—332 (1953).
5. Boodman D. M., The Reliability of Airborne Radar Equipment, *J. Opns. Res. Soc. Am.*, 1, 39—45 (Feb. 1953).
6. Brown A. W., A Note on the Use of a Pearson Type III Function in Renewal Theory, *Ann. math. Statist.*, 11, 448—453 (1940).
7. Campbell N. R., The Replacement of Perishable Members of an Operating System, *J. R. Statist. Soc.*, VII, Sec. B, 110—130 (1941).
8. Chung K. L. and Pollard H., An Extension of Renewal Theory, *Proc. Amer. math. Soc.*, 3, 303—309 (1952).
9. Chung K. L. and Wolfowitz J., On a Limit Theorem in Renewal Theory, *Ann. Math.*, 55—56, 1—6 (1952).
10. Davis D. S., An Analysis of Some Failure Data, *J. Amer. Statist. Ass.*, 47, 113—150 (June 1952).
11. Dean Joel, Capital Budgeting, Columbia University Press, New York, 1951.
12. Dean Joel, Capital Expenditures, Management and the Replacement of Milk Trucks, Joel Dean Associates, New York (pamphlet).

13. Dean Joel, Replacement Investments, Chap. VI in «Capital Budgeting», Columbia University Press, New York, 1951.
14. Doob J. L., Renewal Theory from the Point of View of Probability, Trans. Amer. math. Soc., 63, 422—438 (1948).
15. Epstein B. and Sobel M., Life Testing — I, J. Amer. Statist. Ass., 48, 486—502 (1953).
16. Epstein B. and Sobel M., Some Theorems Relevant to Life Testing from an Exponential Distribution, Ann. math. Statist., 25, 373—381 (June 1954).
17. Feller W., An Introduction to Probability Theory and Its Applications, Vol. I, John Wiley & Sons, New York, 1950.
18. Feller W., On the Integral Equation of Renewal Theory, Ann. math. Statist., 13, 243—267 (Sept. 1941).
19. Goodman L., Methods of Measuring Useful Life of Equipment under Operational Conditions, J. Amer. Statist. Ass., 48, 503—530 (Sept. 1953).
20. Grant E., Principles of Engineering Economy, 3rd ed., Ronald Press Company, New York, 1950.
21. Herstein I. and Debreu G., Non-Negative Square Matrices, Econometrica, 21, 597—607 (Oct. 1953).
22. Jeming Joseph, Estimates of Average Service Life and Life Expectancies and the Standard Deviation of Such Estimates, Econometrica, 11, 141—150 (1943).
23. Kai Lai Chung, On the Renewal Theorem in Higher Dimensions, Parts 1 and 2, Skand. Aktuar Tidskr. (1952).
24. Karlin S., On the Renewal Equation, Pac. J. Math., 5, 229—257 (1955).
25. Kendall D. G., Random Fluctuations in the Age-Distribution of a Population Whose Development Is Controlled by the Simple Birth and Death Process, J. R. Statist. Soc., 12, Sec. B, 278 (1950).
26. Kendall D. G., Stochastic Processes and Population Growth, J.R. Statist. Soc., 11, Sec. B, 230 (1949).
27. Kimball Bradford F., A System of Life Tables for Physical Property Based on the Truncated Normal Distribution, Econometrica, 15, 342 (1947).
28. Kurtz E. B., Life Expectancy of Physical Property, Ronald Press Co., New York, 1930.
29. Leslie P. H., On the Use of Matrices in Certain Population Problems, Biometrika, 33, 183—212 (1945).
30. Lotka A. J., A Contribution to the Theory of Self-Renewing Aggregates, with Special Reference to Industrial Replacement, Ann. math. Statist., 10, 1—25 (1939).
31. Lotka A. J., Industrial Replacement, Skand. Aktuar Tidskr., 51 (1933).
32. Lotka A. J., On an Integral Equation in Population Analysis, Ann. math. Statist., 10, 144—161 (1939).
33. Lotka A. J., Population Analysis: A Theorem Regarding the Stable Age Distribution, J. Wash. Acad. Sci., 27, 299 (1937).
34. Lotka A. J., The Stability of the Normal Age Distribution, Proc. nat. Acad. Sci., 8, 339 (1922).
35. Lotka A. J., The Theory of Industrial Replacement, Skand. Aktuar Tidskr., 1—14 (1940).
36. Preinreich G. A. D., The Economic Life of Industrial Equipment, Econometrica, 8, 12 (1940).
37. Preinreich G.A.D., The Present Status of Renewal Theory, Waverly Press, Baltimore, 1940.
38. Preinreich G. A. D., The Theory of Industrial Replacement, Skand. Aktuar Tidskr., 1—19 (1939).
39. Problems for Short Course in Operations Research, Summer, 1953, Massachusetts Institute of Technology, Cambridge (hctographed).
40. Shellard G. G., Failure of Complex Equipment, J. Opns. Res. Soc. Am., 1, 130—136 (1953).
41. Terborgh B., Dynamic Equipment Policy, McGraw-Hill Book Co., New York, 1949.
42. Tested Approaches to Capital Equipment Replacement, Special Report No. 1, American Management Association, New York, 1954.
43. Winfrey R. and Kurtz E. B., Life Characteristics of Physical Property, Bulletin 103, Iowa State College, Ames, 30, no. 3 (1931).

Во всех обсуждавшихся до сих пор моделях рассматривались противоречия интересов, внутренних по отношению к данной организации, скажем противоречие между стремлением минимизировать издержки производства и стремлением минимизировать объем капиталовложений в запасы. Сейчас мы приступаем к исследованию моделей, учитывающих противоречия внешние по отношению к данной организации либо по крайней мере внешнее противоречие, называемое «конкуренцией». Конкуренция проявляется в описываемых ниже задачах в том смысле, что эффективность решений, принятых одной стороной, зависит от решений другой стороны. Исследуются состязательные задачи двух типов: игры и торги.

Глава 18 представляет собой введение в теорию игр. Эта теория привлекает к себе в последнее время большое внимание, но пока что она не нашла широкого практического применения. Иными словами, математический аппарат теории игр не используется в широких масштабах в исследовании операций, однако логические основы этой теории и введенные в ней понятия играют существенную роль для исследования операций.

Почти все без исключения интересуются играми и стремятся научиться выигрывать, «не прибегая к явному обману». Поэтому на теорию игр обращено всеобщее внимание. Однако теория игр ставит гораздо более серьезные цели, а ее развитие направлено на то, чтобы глубже постигнуть сущность состязательного экономического поведения. Пока что удалось формализовать лишь относительно простые игровые ситуации, но математический аппарат описания даже элементарных игр чрезвычайно сложен. Тем не менее, как уже указывалось, теория игр представляет ценность, не связанную в сущности с ее математическим аппаратом. Она продемонстрировала возможность рационального выбора стратегии из ряда альтернатив как в состязательных, так и в несостязательных ситуациях.

В главе 19 рассматривается несколько моделей конкурентных ситуаций, возникающих в условиях торгов. Эти исследования начали проводить в самое последнее время, и они представляют собой лишь первые шаги в новом важном направлении изучения состязательного поведения; однако эти модели уже успешно применялись и при решении промышленных задач. К сожалению, нельзя описать эти задачи по сообщениям соблюдения коммерческой тайны, но мы разберем примеры, аналогичные реальным условиям.

Глава 18

ТЕОРИЯ ИГР

ВВЕДЕНИЕ

Анализ математической стороны и основных принципов теории игр был дан Джоном фон Нейманом еще в 1928 году. В этой ранней работе Нейман не разбирал практические приложения рассматриваемой задачи, сосредоточив основное внимание на логических основаниях квантовой механики. В 1944 году фон Нейман и Моргенштерн опубликовали свою хорошо известную работу «Теория игр и экономического поведения», положившую начало бурному развитию математического исследования игр. Эта работа явилась основным толчком для развития линейного программирования и теории статистических решений Вальда *. Она открыла также новый подход к задачам выбора решений в конкурентных ситуациях.

ПРИРОДА ИГР

В каждой игре есть цель, или конечное состояние, к которому стремятся игроки, выбирая направления допустимых по правилам игры действий. В некоторых случаях смысл игры заключается в достижении цели с наибольшей эффективностью. Эффективность может измеряться счетом, как в гольфе или бейсболе. В гольфе, например, цель — заполнить 18 (или какое-то другое определенное число) лунок с как можно меньшим счетом. В некоторых случаях эффективность измеряется временем или числом ходов до конца игры и цель игрока — прийти к концу состязания первым.

В других играх только одно лицо или команда может достигнуть цели; предмет игры — добиться этой цели. Например, в шахматах такая цель — объявить мат.

Игра может иметь альтернативные цели. Например, при игре в кости: выбросить 7 или 11 при первом бросании и не выбросить 2, 3 или 12; или если при первом бросании получено 4, 5, 6, 8, 9 или 10, то целью игрока будет выбросить то же число прежде, чем будет получена семерка.

Игры с одним участником, такие, как пасьянс или гольф (при некоторых условиях), — игры без конкуренции. Участник играет на счет или для достижения цели. Это очень похоже на один из видов ситуации по принятию решений, рассмотренной в главе 5, где задача состоит в том, чтобы достигнуть одной или большего количества целей с наибольшей эффективностью. Приведенный там подход (стратегия по принятию решения) может быть применен к подобным играм с одним участником.

Нас будут интересовать игры с конкуренцией, в которых есть действия и противодействия. Поэтому термин *игра* будет использоваться только по отношению к конкурентным играм.

Конкурентная игра — это такая игра, где существует конечное состояние (выигрыш), которого добивается каждый игрок, но не каждый может его добиться. Таким образом, по отношению к этой цели игроки находятся в противоречии. Но благодаря правилам игры это противоречие приводит к общему направлению действий. Каждый игрок имеет множество возможных ходов. Выбрать один из них — значит *сделать ход*. *Партия* — это последовательность или множество ходов, которые приводят игру к конечному состоянию.

* См. ссылки в главах 20 и 21.

Во многих играх достижение цели сопровождается каким-нибудь выигрышем, в частности денежным. Эти выигрыши и проигрыши (отрицательные выигрыши) являются в некотором смысле способом счета игры, т. е. служат выражением эффективности. Мы предположим, что выигрыш в игре может быть всегда переведен в денежное выражение, поскольку мы интересуемся экономическими играми, или, другими словами, коммерческой конкуренцией.

Игра с нулевой суммой — это такая игра, в которой сумма выигрышей участников после конца игры равна нулю. Так, если *A* платит *B* 1 доллар, то *B* платит *A* —1 доллар (т. е. получает 1 доллар от *A*). Сумма этих выигрышей равна нулю. Но предположим, что *A* должен дополнительно платить 1 доллар в «банк» или в «игорный дом», тогда сумма выигрышей игроков не равна нулю. Это была бы игра с *ненулевой суммой*.

Стратегия — это установленный игроком метод выбора ходов в течение игры. Таким образом, стратегия — это совокупность правил выбора решения.

Наконец, платежная *матрица* — это таблица, которая определяет, какие платежи должны быть сделаны после завершения игры.

Прежде чем мы перейдем непосредственно к теории игры, сделаем одно последнее замечание: эта теория не пытается описывать, как могла бы быть проведена игра. Она содержит процедуру и принципы, при помощи которых можно отбирать партии. В действительности теория игр является теорией принятия решений, применимой к конкурентным ситуациям.

ИГРА ДВУХ ЛИЦ С НУЛЕВОЙ СУММОЙ

Рассмотрим следующую игру для двух игроков. Игрок *A* передвигает на рис. 18.1 фишки с ноликами; игрок *B* передвигает фишки с крестиками. По правилам игры фишки могут быть передвинуты из одного белого квадрата в любой другой соединенный с ним белый квадрат при условии, что этот квадрат не занят. За один ход может и должна быть передвинута одна фишка. Игроки ходят по очереди. Партия кончается, если игрок должен ходить, но не может сделать хода, не нарушив правил игры. Игрок, не имеющий возможности сделать очередной ход, проигрывает. Цель игры — выиграть в наименьшее число ходов.

Допустим, что в этом примере игрок *A* делает ход, исходя из данной позиции, и выбранный им ход не сообщается игроку *B*. Игрок *B*, не знающий, какой ход сделал *A*, может наметить следующую стратегию. Первый ход *B*: передвинуть фишку с *d4* на *c3* (*A* делает свой второй ход). Второй ход *B*: передвинуть фишку с *c3* на *b2*. Из рис. 18.1 сразу видно, что (1) эти ходы возможны независимо от того, какие два первых хода сделает игрок *A*, и (2) игрок *A* проигрывает игру после второго хода *B*.

Таким образом, в этом случае у нас есть стратегия для *B*, которая обеспечивает ему выигрыш. Итак, можно показать, что в этой игре одна стратегия игрока *B* обеспечит ему выигрыш независимо от того, какие ходы собирается сделать игрок *A*, если *A* ходит первым. В этом случае можно сказать, что *B* нашел решение игры. (Заметим, что для *B* в этом случае не существует лучшей стратегии.)

	1	2	3	4
a	○		○	
b				*
c	○			
d		*		*
	1	2	3	4

Рис. 18.1. Игра двух лиц с нулевой суммой.

Вернемся к примеру. Какова лучшая стратегия игрока *B*, если он не знает, кто будет делать первый ход? В этом примере мы можем выбрать для игрока *B* следующую стратегию:

№ хода	Ход
1	С <i>d4</i> на <i>c3</i> .
2	С <i>c3</i> на <i>b2</i> , или с <i>d2</i> на <i>c1</i> или с <i>b4</i> на <i>a3</i> , и партия будет окончена выигрышем <i>B</i> . Если ни один из этих ходов не возможен, то выбирается ход с <i>c3</i> на <i>d4</i> .
3	Если хода нет, партия окончена выигрышем <i>A</i> .
4	Такой же, как 2-й ход.
5	Такой же, как 3-й ход.
и т. д.	

Легко видеть, что это лучшая стратегия игрока *B*, независимо от того, кто ходит первым. Подобным образом мы можем составить стратегию для игрока *A*. Найдя лучшие (оптимальные) стратегии для обоих игроков, мы получили полное *решение* для игры. Определение оптимальных стратегий для игроков составляет *решение* игры.

Эти понятия интуитивно принимаются на веру. Позже мы увидим, что, например, в игре трех лиц с нулевой суммой мы не сможем применить те же понятия для определения решения. Появятся некоторые трудности, так как надо будет принимать в расчет то, что игроки могут образовывать коалиции. Еще больше трудностей в определении решения встретится в играх *n* лиц с ненулевой суммой.

Теперь, однако, вернемся к игре двух лиц с нулевой суммой и примем следующее определение: игра решена, если мы можем найти оптимальные стратегии для каждого из игроков.

ПРЯМОУГОЛЬНЫЕ ИГРЫ (ДВУХ ЛИЦ С НУЛЕВОЙ СУММОЙ И С СЕДЛОВОЙ ТОЧКОЙ)

Рассмотрим следующую игру. Игрок *A* имеет три возможных плана игры *: *P*, *Q*, *R*. Игрок *B* имеет два возможных плана игры: *S*, *T*. Правила игры устанавливают, что в соответствии с выбранными планами производятся следующие платежи:

Выбран- ные планы	Платежи	Выбран- ные планы	Платежи
<i>P</i> , <i>S</i>	<i>A</i> платит <i>B</i> 2 доллара	<i>Q</i> , <i>T</i>	<i>B</i> платит <i>A</i> 3 доллара
<i>P</i> , <i>T</i>	<i>B</i> платит <i>A</i> 2 доллара	<i>R</i> , <i>S</i>	<i>B</i> платит <i>A</i> 1 доллар
<i>Q</i> , <i>S</i>	<i>A</i> платит <i>B</i> 1 доллар	<i>R</i> , <i>T</i>	<i>B</i> платит <i>A</i> 2 доллара

* В специальной литературе в этом случае вместо термина *план игры* употребляется термин *чистая стратегия*. (Прим. перев.)

Какова оптимальная стратегия для игроков A и B в этой игре?

Правила платежей удобно записать в матричной форме. Пусть положительное число показывает выигрыш игрока A , а отрицательное число показывает выигрыш игрока B . Тогда мы имеем платежную «матрицу», показанную в таблице 18.1.

Таблица 18.1

		Игрок B	
		План	
Игрок A	P	—2	2
	Q	—1	3
	R	1	2

Рассмотрим игрока B . Очевидно, что план T для него невыгоден. Если он выбирает этот план, он всегда проигрывает. Таким образом, его оптимальная стратегия — всегда выбирать план S . В худшем случае (когда A выберет план R) он проиграет 1 доллар.

Теперь обратимся к игроку A . Ему достанется наибольший выигрыш, если он выберет план Q , а B выберет план T . Но это вряд ли произойдет *, так как из-за предыдущих рассуждений B никогда не выберет T . Таким образом, лучшее, что может сделать A (если B выберет S), — это выбрать план R ; в этом случае игрок A выиграет 1 доллар.

Таким образом, мы нашли полное решение игры. Кроме того, при этом решении игрок A выигрывает 1 доллар, а игрок B проигрывает 1 доллар. В этом случае 1 доллар является *ценой игры*.

Игра, только что рассмотренная нами, называется *прямоугольной игрой*, так как ее матрица выигрышей прямоугольная. Чтобы получить решение прямоугольной игры, нам надо найти:

1. Оптимальные стратегии для двух игроков.
2. Цену игры.

Принцип минимакса и максимина. Рассмотрим платежную матрицу прямоугольной игры, приведенную в таблице 18.2. Сначала мы решим эту игру, пользуясь рассуждениями, данными в предыдущей части. Затем мы введем метод решения, основанный на принципах минимакса. Рассуждения в обоих случаях одинаковы и, следовательно, ведут к одному решению.

Таблица 18.2

		B	
		План	
A	P	—2	—4
	Q	—1	3
	R	1	2

* В этой части мы предполагаем, что все игроки разумны. Поэтому здесь, например, игрок B не выбрал бы план T . Если противник действует неразумно, игрок может извлечь из этого пользу.

Метод 1. Игрок A никогда не выберет план P , так как он всегда с большим успехом может выбрать план Q или план R . Учитывая это, игрок B не может вообще принимать в расчет план P . В этом случае, очевидно, он не выберет T , так как для него всегда выбор S выгоднее. В свою очередь A основывается на том, что B выберет S , и, таким образом, его лучшая политика в игре — план R . Итак, мы пришли к решению:

Оптимальная стратегия игрока A : план R .

Оптимальная стратегия игрока B : план S .

Цена игры для A : 1 доллар (выигрыш).

Цена игры для B : —1 доллар (проигрыш).

Метод 2. Теперь рассмотрим следующие рассуждения.

Игрок A :

При плане P его наименьший (минимум) выигрыш —4 доллара.

При плане Q его наименьший (минимум) выигрыш —1 доллар.

При плане R его наименьший (минимум) выигрыш +1 доллар.

Наибольший (максимум) из наименьших (минимум) возможных выигрышей 1 доллар. Значит, мы можем сказать, что «максимин для A » равен одному доллару (что соответствует выборам R , S).

Игрок B :

При плане S его наибольший (максимум) проигрыш 1 доллар.

При плане T его наибольший (максимум) проигрыш 3 доллара.

Таким образом, наименьший (минимум) из наибольших проигрышей — 1 доллар. Мы говорим, что «минимакс для B » равен 1 доллару (что опять соответствует выборам R , S).

В математических обозначениях «максимин для A » записывается выражением

$$\max_i \min_j a_{ij},$$

где a_{ij} является элементом i -й строки и j -го столбца платежной матрицы. Точно так же «минимакс для B » обозначается выражением

$$\min_j \max_i a_{ij}.$$

Если для данной игры

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = g,$$

то игрок A может выиграть g , но больше чем g игрок B может не дать ему выиграть.

В математических символах платежная матрица нашего примера представляет собой матрицу $[a_{ij}]$, $i = 1, 2, 3$; $j = 1, 2$. Элементы матрицы нашего примера следующие:

$$a_{11} = -2, \quad a_{12} = -4,$$

$$a_{21} = -1, \quad a_{22} = 3,$$

$$a_{31} = 1, \quad a_{32} = 2,$$

а результаты задаются равенствами

$$\max_i \min_j a_{ij} = a_{31} = 1,$$

$$\min_j \max_i a_{ij} = a_{31} = 1.$$

Таким образом, решение задается выбором планов R и S (соответствующих a_{31}) игроками A и B .

Заметим, что метод 2 (использование принципов минимакса и максимина) дает то же решение, что и метод 1.

Седловые точки. Не всякая прямоугольная игра приводит к решениям с единственным оптимальным выбором для обоих игроков A и B . Рассмотрим, например, игру, платежная матрица которой приведена в таблице 18.3.

Таблица 18.3

		B	
		S	T
A	P	-2	1
	Q	2	-1

Если A выберет план P , то B , очевидно, выберет план S . Если A выберет план Q , то B выберет T . Мы видим, для A нет определенного лучшего плана. То же можно сказать и про игрока B .

Используя принципы минимакса, находим:

максимин для $A = -1$ доллар (выбор Q, T),
минимакс для $B = 1$ доллар (выбор P, T),

и в этом случае мы имеем:

максимин для A не равен минимаксу для B , т. е.

$$\max_i \min_j a_{ij} \neq \min_j \max_i a_{ij}.$$

Такие игры мы рассмотрим в следующем параграфе.

Игры, для которых максимин для A равен минимаксу для B , называются *играми с седловой точкой*. В примере, приведенном в предыдущем параграфе, седловая точка состояла из выборов R, S .

Вообще, если

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = a_{i_0 j_0},$$

то

оптимальная стратегия для A — план i_0 ,
оптимальная стратегия для B — план j_0 ,
цена игры (для A) — $a_{i_0 j_0}$.

Наиболее легкий прием отыскания седловой точки заключается в определении числа, *наименьшего из всех чисел своей строки и наибольшего из чисел своего столбца*. Если такого числа нет, то нет и седловой точки. Стратегии, соответствующие найденному числу, — оптимальные стратегии игроков, а найденное число — цена игры. Если существует два или больше таких чисел, то имеется два или более решений. Каждое решение соответствует седловой точке.

В качестве примера рассмотрим игру, платежная матрица которой приведена в таблице 18.4. В этом примере (III, III) — седловая точка, так как 2 — наименьшее число в своей строке и наибольшее в своем столбце. Следовательно, III — оптимальный план для A , III — оптимальный план для B , 2 доллара — цена игры (A выигрывает, а B проигрывает).

Таблица 18.4

		<i>B</i>				
		План	I	II	III	IV
<i>A</i>	I	3	−5	0	6	
	II	−4	−2	1	2	
	III	5	4	2	3	

Платежная матрица, показанная в таблице 18.5, a , имеет две седловые точки (P, S) и (P, U), а в платежной матрице таблицы 18.5, b нет седловой точки.

Допустим, мы хотим построить платежную матрицу и установить, имеет ли седловую точку следующая игра: каждый из двух игроков одновременно кладет на стол одной рукой

монету в 10 центов, а другой — монету в один цент. Если у игроков в левой руке были одинаковые монеты, то все четыре монеты забирает игрок A .

Таблица 18.5

	S	T	U
P	1	2	1
Q	0	-4	-1
R	1	3	-2

(а)

	S	T	U
P	1	2	1
Q	0	-4	-1
R	2	1	2

(б)

В противном случае все четыре монеты забирает игрок B . Платежная матрица приведена в таблице 18.6, где D обозначает десять центов, а P — один цент. В этой игре нет седловой точки.

Таблица 18.6
Платежная матрица (в центах)

		B		
		План	PD	DP
A	PD	11	-11	
	DP	-11	11	

Таблица 18.7

		B		
		План	S	T
A	P	-3	7	
	Q	6	1	

ПРЯМОУГОЛЬНЫЕ ИГРЫ (БЕЗ СЕДЛОВЫХ ТОЧЕК): СМЕШАННЫЕ СТРАТЕГИИ

Рассмотрим игру с платежной матрицей, показанной в таблице 18.7. Платежная матрица не имеет седловой точки, т. е. в этой игре ни один из участников игры не может выбрать один план в качестве своей оптимальной стратегии. Следовательно, каждый игрок должен найти некоторую «смешанную стратегию» для максимизации своего выигрыша или минимизации своего проигрыша.

Допустим, например, что A решает играть половину времени по плану P , а половину по плану Q^* . Если B при этом будет все время играть по плану S , то ожидаемый выигрыш для A будет

$$\frac{1}{2}(-3) + \frac{1}{2}(6) = -1,5 + 3 = \$1,50,$$

а если B будет все время играть по плану T , то ожидаемый выигрыш игрока A будет

$$\frac{1}{2}(7) + \frac{1}{2}(1) = 3,5 + 0,5 = \$4,0.$$

Допустим далее, что B также применяет смешанную стратегию и в половине случаев выбирает план S , а в половине T , причем каждый выбор определяет случайно. В этом случае ожидаемый выигрыш игрока A будет

$$\frac{1}{2}[\frac{1}{2}(-3) + \frac{1}{2}(6)] + \frac{1}{2}[\frac{1}{2}(7) + \frac{1}{2}(1)] = \$2,75.$$

Подобным образом мы можем подсчитать, каков ожидаемый выигрыш A при других смешанных стратегиях. Например, если A играет в четверти случаев план P , а Q — в трех четвертях, в то время как B играет одну треть раз S , а две трети T , то ожидаемый выигрыш A будет

$$\begin{aligned} \frac{1}{3}[\frac{1}{4}(-3) + \frac{3}{4}(6)] + \frac{2}{3}[\frac{1}{4}(7) + \frac{3}{4}(1)] = \\ = \frac{1}{12}(-3 + 18 + 14 + 6) = \frac{35}{12} = \$2,92. \end{aligned}$$

Возникает вопрос: какова оптимальная смешанная стратегия для игроков? Из приведенных примеров мы видим, что выигрыш A меняется от 1,5 долларов до 4 долларов. Мы хотим узнать, не может ли игрок A обеспечить себе некоторый минимум выигрыша и каков он? Подобным образом, может ли B застраховаться от проигрыша больше некоторого максимального?

На эти вопросы можно ответить положительно. Математическая теория игр дает нам доказательство того, что всегда существуют оптимальные стратегии и способы их нахождения.

Мы рассмотрим некоторые математические результаты в следующих двух параграфах. Рассмотрим теперь, как могут быть найдены оптимальные стратегии этой игры и каковы ожидаемые значения выигрыша и проигрыша для игроков.

Пусть A играет P с частотой x , а Q с частотой $1 - x$. Тогда, если B будет играть все время S , выигрыш A будет

$$g(A, S) = x(-3) + (1-x)6 = 6 - 9x.$$

Если B будет все время играть T , выигрыш A будет

$$g(A, T) = x(7) + (1-x)1 = 1 + 6x.$$

* Игрок A мог бы делать выбор плана случайно так, что план P выбирается с частотой $\frac{1}{2}$ и план Q выбирается с частотой $\frac{1}{2}$. Пригодным способом для случайного выбора в этом случае было бы бросание симметричной монетки.

Математически можно показать, что если A выберет x так, что $g(A, S) = g(A, T)$, то для него это будет оптимальной стратегией. Итак,

$$6 - 9x = 1 + 6x, \quad 5 = 15x, \quad x = 1/3.$$

В результате мы получаем

$$g(A) = 1/3(-3) + 2/3(6) = \$3,0.$$

Таким образом, независимо от частот, с которыми B играет S или T , A всегда выигрывает 3 доллара. (Если, например, B играет S с частотой $1/4$ и T с частотой $3/4$, то ожидаемый выигрыш для A будет

$$g(A, 1/3, 2/3) = 1/4[1/3(-3) + 2/3(6)] + 3/4[1/3(7) + 2/3(1)] = \$3,0,$$

и так же при любом выборе частот для B .) Следовательно, выбором частот $1/3$ и $2/3$ A получает обеспеченный выигрыш в 3 доллара.

Тот же самый метод можно применить к игроку B . Обозначим частоту выбора плана S буквой y , тогда частота выбора T будет $1 - y$. Для оптимальной стратегии мы имеем:

$$g(B, P) = y(-3) + (1-y)7 = y(6) + (1-y)1 = g(B, Q),$$

$$7 - 10y = 1 + 5y, \quad 6 = 15y, \quad y = 2/5, \quad 1 - y = 3/5,$$

$$g(B) = 2/5(-3) + 3/5(7) = \$3,0.$$

Заметим, что $g(A) = g(B)$, как и следовало ожидать для игры с нулевой суммой. Таким образом, полным решением данной игры будет: (1) A следует играть P и Q соответственно с частотами $1/3$ и $2/3$, (2) B следует играть S и T с частотами $2/5$ и $3/5$ соответственно, (3) цена игры — 3 доллара.

Основные теоремы для прямоугольных игр. Платежи для прямоугольных игр всегда могут быть заданы в виде матрицы $m \times n$, где игрок A имеет m возможных планов, B имеет n возможных планов и платежная матрица есть матрица $[a_{ij}]$ (см. табл. 18.8).

Таблица 18.8

		B			
A	План	I	II	...	n
	I	a_{11}	a_{12}	...	a_{1n}
	II	a_{21}	a_{22}	...	a_{2n}
	.	.	.	...	.
	m	a_{m1}	a_{m2}	...	a_{mn}

Математически можно показать, что (1) каждой прямоугольной игре соответствует определенное значение цены g ; это значение единственно; (2) существует оптимальная стратегия игрока A , иначе говоря, существуют

частоты $x_1, x_2, \dots, x_m$ такие, что $x_1 + x_2 + \dots + x_m = 1$, и если игрок A играет, применяя план I с частотой x_1 , план II с частотой $x_2, \dots$, план m с частотой x_m , то он может обеспечить себе средний выигрыш не менее g , где g — цена игры; (3) также и для игрока B существует оптимальная стратегия

$$Y = [y_1, y_2, \dots, y_n], \quad \sum_{j=1}^n y_j = 1$$

такая, что, если он применяет планы I, II, $\dots$, n с частотами $y_1, y_2, \dots, y_n$ соответственно, он обеспечивает себе проигрыш не больше g .

Отметим, что для прямоугольной игры, матрица которой имеет седловую точку (i_0, j_0) , мы имеем следующее решение:

$$x_i = \begin{cases} 0 & \text{при } i \neq i_0, \\ 1 & \text{при } i = i_0, \end{cases} \quad y_j = \begin{cases} 0 & \text{при } j \neq j_0, \\ 1 & \text{при } j = j_0, \end{cases} \quad g = a_{i_0 j_0}.$$

Общее решение прямоугольных игр. Можно показать, что неизвестные $x_1, x_2, \dots, x_m; y_1, y_2, \dots, y_n$ и g могут быть найдены из следующих соотношений:

$$x_1 + x_2 + \dots + x_m = 1, \quad x_i \geq 0, \quad (1)$$

$$y_1 + y_2 + \dots + y_n = 1, \quad y_j \geq 0, \quad (2)$$

$$x_1 a_{1j} + x_2 a_{2j} + \dots + x_m a_{mj} \geq g, \quad \text{где } j = 1, 2, \dots, n, \quad (3)$$

$$y_1 a_{i1} + y_2 a_{i2} + \dots + y_n a_{in} \leq g, \quad \text{где } i = 1, 2, \dots, m. \quad (4)$$

Соотношение (3) в действительности представляет собой n неравенств, по одному неравенству для каждого j . Соотношение (4) представляет собой m неравенств. Таким образом, мы имеем $m + n + 1$ неизвестных с $m + n + 2$ соотношениями (с дополнительными ограничениями $x_i \geq 0, y_j \geq 0$, так как отрицательные частоты не имеют смысла). (Заметим, что соотношения (3) и (4) могут быть как равенствами, так и неравенствами.) Теоремы предыдущей части гарантируют существование решения, удовлетворяющего этим соотношениям. Они гарантируют также единственность g . Однако для x_i и y_j игра может иметь несколько или даже бесконечное множество решений.

Пример. Для игры, представленной в таблице 18.7, неизвестными будут x_1, x_2, y_1, y_2 и g . Соотношения следующие:

$$x_1 + x_2 = 1, \quad x_1 \geq 0, \quad x_2 \geq 0, \quad (5)$$

$$y_1 + y_2 = 1, \quad y_1 \geq 0, \quad y_2 \geq 0, \quad (6)$$

$$x_1(-3) + x_2(6) \geq g, \quad (7)$$

$$x_1(7) + x_2(1) \geq g \quad (8)$$

$$y_1(-3) + y_2(7) \leq g, \quad (9)$$

$$y_1(6) + y_2(1) \leq g. \quad (10)$$

Такие задачи могут быть решены при помощи «обычной алгебры», графическим способом, применением матричной алгебры или итеративным методом. Рассмотрим подробнее эти методы.

Алгебраические решения. Алгебраический метод — способ прямого нахождения неизвестных из соотношений (1), (2), (3) и (4). Решение основано на указанном ранее факте: каждая прямоугольная игра имеет цену, а именно цену g , которая существует и единственна. Мы постараемся найти одну такую цену g , которая удовлетворяет нашим соотношениям. Мы строим

процедуру поиска цены g , опираясь на то, что она единственна. Итак, первый очевидный шаг — предположить, что соотношения (3) и (4) суть равенства.

Посмотрим, к чему приведет это допущение при решении рассмотренного ранее примера. В этом случае получается следующая система уравнений:

$$x_1 + x_2 = 1, \quad (11)$$

$$y_1 + y_2 = 1, \quad (12)$$

$$-3x_1 + 6x_2 = g, \quad (13)$$

$$7x_1 + x_2 = g, \quad (14)$$

$$-3y_1 + 7y_2 = g, \quad (15)$$

$$6y_1 + y_2 = g. \quad (16)$$

У нас имеется пять неизвестных и шесть уравнений. Поэтому мы решим только пять уравнений и проверим, удовлетворяет ли найденное решение шестому. Проверим также, удовлетворяют ли найденные значения неизвестных соотношениям $x_1 \geq 0$, $x_2 \geq 0$, $y_1 \geq 0$, $y_2 \geq 0$. Если неизвестные удовлетворяют всем этим условиям, то они составляют полное решение.

Уравнения (11) и (12) приводят к равенствам $x_2 = 1 - x_1$, $y_2 = 1 - y_1$. Следовательно, (13), (14) и (15) переходят в следующие уравнения:

$$-3x_1 + 6(1 - x_1) = g, \text{ отсюда } -9x_1 - g = -6 \text{ или } 9x_1 + g = 6, \quad (17)$$

$$7x_1 + (1 - x_1) = g, \text{ отсюда } 6x_1 + 1 = g \text{ или } 6x_1 - g = -1, \quad (18)$$

$$-3y_1 + 7(1 - y_1) = g, \text{ отсюда } -3y_1 + 7 - 7y_1 = g \text{ или } 10y_1 + g = 7. \quad (19)$$

Складывая уравнения (17) и (18), находим $15x_1 = 5$, т. е. $x_1 = 1/3$, так что из уравнений (17), (18) и (11) получаем:

$$x_1 = 1/3, \quad x_2 = 2/3, \quad g = 3.$$

Подставляя эти значения в уравнение (16), получаем $10y_1 = 7 - 3 = 4$, откуда

$$y_1 = 2/5, \quad y_2 = 3/5.$$

Итак, поскольку $1/3 > 0$, $2/3 > 0$, $2/5 > 0$, $3/5 > 0$, мы имеем полное решение, а именно:

$$x_1 = 1/3, \quad x_2 = 2/3, \quad y_1 = 2/5, \quad y_2 = 3/5.$$

С другой стороны, если мы попытаемся тем же методом получить решение игры, представленной в таблице 18.2 (которую мы перенесли

Таблица 18.9

		B		
		Частота	y_1	y_2
			S	T
A	x_1	P	-2	-4
	x_2	Q	-1	3
	x_3	R	1	2

в таблицу 18.9), то получим:

$$x_1 + x_2 + x_3 = 1, \quad (20)$$

$$y_1 + y_2 = 1, \quad (21)$$

$$x_1(-2) + x_2(-1) + x_3(1) = g, \quad (22)$$

$$x_1(-4) + x_2(3) + x_3(2) = g, \quad (23)$$

$$y_1(-2) + y_2(-4) = g, \quad (24)$$

$$y_1(-1) + y_2(3) = g, \quad (25)$$

$$y_1(1) + y_2(2) = g. \quad (26)$$

Мы имеем шесть неизвестных: x_1, x_2, x_3, y_1, y_2 , цену игры g — и семь уравнений. Рассмотрим, например, уравнения (21), (24) и (25):

$$\begin{aligned} y_2 = 1 - y_1, \quad -2y_1 + (1 - y_1)(-4) &= g, \quad -2y_1 - 4 + 4y_1 = g, \\ 2y_1 - 4 = g, \quad -y_1 + (1 - y_1)3 &= g, \end{aligned}$$

т. е.

$$-4y_1 + 3 = g.$$

Отсюда

$$2y_1 - 4 = -4y_1 + 3, \quad y_1 = 7/6,$$

что противоречит ограничению $0 \leq y_1 \leq 1$. Таким образом, допустив, что соотношения переходят в равенства, мы не находим решения.

Общая процедура дальнейшего поиска решения следующая: допускаем, что одно соотношение переходит в неравенство, а остальные оставляем равенствами. Если мы опять приходим к противоречию, допускаем другие неравенства до тех пор, пока в конце концов не достигнем решения.

Имеется, однако, теорема, которая позволяет облегчить эти вычисления. Она устанавливает, что если

$$x_1 a_{1j} + x_2 a_{2j} + \dots + x_m a_{mj} > g, \quad \text{то } y_j = 0,$$

и таким же образом, если

$$y_1 a_{i1} + y_2 a_{i2} + \dots + y_n a_{in} < g, \quad \text{то } x_i = 0.$$

Применяя эту теорему к только что приведенному примеру, получаем:

(а) Если $x_1(-2) + x_2(-1) + x_3(1) > g$, то $y_1 = 0$. Тогда уравнения (24), (25) и (26) приводят к $y_2 = 0, g = 0$, так что $y_1 + y_2 = 0$, что противоречит условию $y_1 + y_2 = 1$.

(б) Подобно этому получается противоречие, если допустить

$$x_1(-4) + x_2(3) + x_3(2) > g.$$

(в) Если $y_1(-2) + y_2(-4) < g$, то $x_1 = 0$, и уравнения (20), (22) и (23) приводят к противоречию.

Продолжая таким же способом, в конце концов мы находим, что только следующие соотношения дают решения:

$$x_1 + x_2 + x_3 = 1, \quad (20)$$

$$y_1 + y_2 = 1, \quad (21)$$

$$x_1(-2) + x_2(-1) + x_3(1) = g, \quad (22)$$

$$x_1(-4) + x_2(3) + x_3(2) > g, \quad (23a)$$

$$y_1(-2) + y_2(-4) < g, \quad (24a)$$

$$y_1(-1) + y_2(3) < g, \quad (25a)$$

$$y_1(1) + y_2(2) = g. \quad (26)$$

Неравенство (23а) приводит к $y_2 = 0$, так что $y_1 = 1$. Далее, неравенство (24а) влечет $x_1 = 0$, и неравенство (25а) влечет $x_2 = 0$, что дает $x_3 = 1$. Следовательно, окончательным решением будет

$$x_1 = 0, \quad x_2 = 0, \quad x_3 = 1, \quad y_1 = 1, \quad y_2 = 0, \quad g = 1.$$

Это и есть алгебраическое решение игры с седловой точкой. Мы видим, что алгебраическое решение игры очень громоздко. Этот метод становится во много раз более запутанным при увеличении планов для A и B .

Графическое решение. Графическое решение может быть использовано, когда один из игроков имеет только два возможных плана игры. Мы

Таблица 18.10

		B			
		Частота	y_1	y_2	y_3
		План	R	S	T
A	x_1	P	2	-2	3
	x_2	Q	-3	5	-1

проиллюстрируем этот метод примером, показанным в таблице 18.10. Общие соотношения следующие:

$$x_1 + x_2 = 1, \quad (27)$$

$$y_1 + y_2 + y_3 = 1, \quad (28)$$

$$2x_1 - 3x_2 \geq g, \quad (29)$$

$$-2x_1 + 5x_2 \geq g, \quad (30)$$

$$3x_1 - x_2 \geq g, \quad (31)$$

$$2y_1 - 2y_2 + 3y_3 \leq g, \quad (32)$$

$$-3y_1 + 5y_2 - y_3 \leq g. \quad (33)$$

Перепишем и упростим уравнения, исключив игрока, у которого только два возможных плана (в этом случае игрока A). Имеем:

$$2x_1 - 3(1 - x_1) \geq g, \quad \text{откуда} \quad 5x_1 - 3 \geq g, \quad (34)$$

$$-2x_1 + 5(1 - x_1) \geq g, \quad \text{откуда} \quad -7x_1 + 5 \geq g, \quad (35)$$

$$3x_1 - (1 - x_1) \geq g, \quad \text{откуда} \quad 4x_1 - 1 \geq g. \quad (36)$$

Построим график прямой $g = 5x_1 - 3$, как показано на рис. 18.2. Эта прямая делит плоскость на две области. Для любой точки (x_0, g) , находящейся над прямой, будет выполняться соотношение

$$g > g_0 = 5x_0 - 3,$$

а для точки ниже прямой — соотношение

$$g < g_0 = 5x_0 - 3.$$

В данном случае нас интересует неравенство $5x_1 - 3 \geq g$. Поэтому искомое решение будет на этой прямой или под ней. Проведем все три прямые в интервале $0 \leq x_1 \leq 1$ и отметим все точки ниже всех трех прямых (см. рис. 18.3). Мы выбираем точку, для

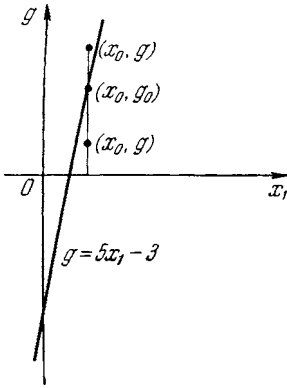


Рис. 18.2. График $g = 5x_1 - 3$.

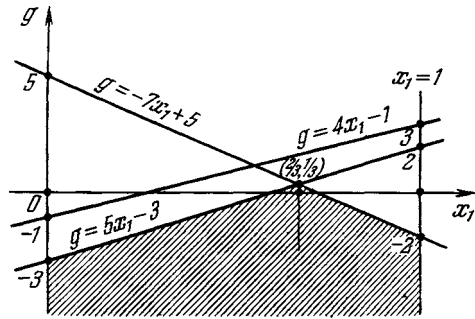


Рис. 18.3. График для g в интервале $0 \leq x_1 \leq 1$.

которой g наибольшее. В нашем примере эта точка находится на пересечении прямых $g = 5x_1 - 3$ и $g = -7x_1 + 5$, т. е. $x_1 = 2/3$, $x_2 = 1/3$. Отсюда $g = 5x_1 - 3 = 10/3 - 3 = 1/3$.

Так как уравнение (31) приводит к

$$3x_1 - x_2 = 3(2/3) - 1/3 = 5/3 > 1/3 = g$$

или

$$3x_1 - x_2 > g,$$

то $y_3 = 0$. Соотношения (28) и (32) дают

$$y_1 + y_2 = 1, \quad y_2 = 1 - y_1,$$

т. е.

$$2y_1 - 2(1 - y_1) = 1/3, \quad 4y_1 = 2 + 1/3 = 7/3,$$

или $y_1 = 7/12$ и $y_2 = 5/12$, и, подставляя эти величины в уравнение (33), получим

$$-3(7/12) + 5(5/12) + 0 = 1/3 = g.$$

Таким образом, решением является

$$x_1 = 2/3, \quad x_2 = 1/3, \quad y_1 = 7/12, \quad y_2 = 5/12, \quad y_3 = 0, \quad g = 1/3.$$

Таблица 18.11

			В	
Частота			y_1	y_2
			План	S
А	x_1	P	-2	-4
	x_2	Q	-1	3
	x_3	R	1	2

Применим графический метод к игре с седловой точкой, рассмотренной нами ранее и приведенной в таблице 18.11.

$$x_1 + x_2 + x_3 = 1, \quad (37)$$

$$y_1 + y_2 = 1, \quad (38)$$

$$-2x_1 - x_2 + x_3 \geq g, \quad (39)$$

$$-4x_1 + 3x_2 + 2x_3 \geq g, \quad (40)$$

$$-2y_1 - 4y_2 \leq g, \quad (41)$$

$$-y_1 + 3y_2 \leq g, \quad (42)$$

$$y_1 + 2y_2 \leq g. \quad (43)$$

Соотношения (38), (41), (42) и (43) приводят к

$$2y_1 - 4 \leq g, \quad -4y_1 + 3 \leq g, \quad -y_1 + 2 \leq g, \quad -y_1 + 2 \leq g.$$

Графики соответствующих прямых в интервале $0 \leq y_1 \leq 1$ приведены на рис. 18.4. В этом случае нас интересуют точки над прямыми, и мы берем наименьшую из этих точек. Этой точкой является точка с $y_1 = 1$, так что $g = 1$ и $y_2 = 0$.

Неравенство (41) дает $-2y_1 - 4y_2 = -2 < 1 = g$. Следовательно, $x_1 = 0$.

Неравенство (42) дает $-y_1 + 3y_2 = -1 < 1 = g$. Следовательно, $x_2 = 0$.

Неравенство (43) дает $y_1 + 2y_2 = 1 = 1 = g$.

Из $x_1 + x_2 + x_3 = 1$ и $x_1 = 0$, $x_2 = 0$ следует, что $x_3 = 1$. Таким образом, окончательным решением является

$$x_1 = 0, \quad x_2 = 0, \quad x_3 = 1, \quad y_1 = 1, \quad y_2 = 0, \\ g = 1.$$

Решение матричным методом*.

Обозначения:

$$A = \|a_{ij}\| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix};$$

$B = [b_{ij}]$ — любая квадратная подматрица матрицы A порядка $r > 1$;

$J_r = [1, 1, \dots, n]$ — матрица $(1 \times r)$;

C^T — матрица, транспонированная к C ;

$\text{adj } B$ — матрица, присоединенная к B ;

$\bar{X} = (1 \times r)$ — матрица, полученная из X вычеркиванием элементов, которые соответствуют строкам, вычеркнутым из A при получении B ;

$\bar{Y} = (1 \times r)$ — матрица, полученная из Y вычеркиванием элементов, которые соответствуют столбцам, вычеркнутым из A при получении B .

Решение.

1. Выберем квадратную подматрицу B матрицы A порядка r ($r \geq 2$) и вычислим

$$\bar{X} = \frac{J_r \text{adj } B}{J_r (\text{adj } B) J_r^T} = \|x_{r1}, x_{r2}, \dots, x_{rr}\|,$$

$$\bar{Y} = \frac{J_r (\text{adj } B)^T}{J_r (\text{adj } B) J_r^T} = \|y_{r1}, y_{r2}, \dots, y_{rr}\|.$$

2. Если некоторое $x_{ri} < 0$ или $y_{rj} < 0$, то отбрасываем найденную матрицу B и пробуем другую матрицу B .

3. Если $x_{ri} \geq 0$ ($i = 1, 2, \dots, r$) и $y_{rj} \geq 0$ ($j = 1, 2, \dots, r$), вычисляем

$$g = \frac{|B|}{J_r (\text{adj } B) J_r^T}$$

и строим X и Y из $\bar{X}$ и $\bar{Y}$, добавляя в соответствующих местах нули.

* Этот раздел предназначен для читателей, знакомых с понятиями теории матриц и с действиями над ними.

Проверяем, удовлетворяются ли неравенства

$$\sum_{i=1}^m x_i a_{ij} \geq g \text{ для каждого } j = 1, 2, \dots, n$$

и неравенства

$$\sum_{j=1}^n y_j a_{ij} \leq g \text{ для каждого } i = 1, 2, \dots, m.$$

Если одно из соотношений не выполнено, то пробуем другое B . Если все соотношения справедливы, то X , Y и g — искомые решения.

Пример. Пусть платежной матрицей будет матрица, приведенная в таблице 18.12. У нас есть только три подматрицы и $J_2 = \|1, 1\|$:

Таблица 18.12

	y_1	y_2	y_3
x_1	2	-2	3
x_2	-3	5	-1

$$B_1 = \left\| \begin{array}{cc} -2 & 3 \\ 5 & -1 \end{array} \right\|, \quad B_2 = \left\| \begin{array}{cc} 2 & 3 \\ -3 & -1 \end{array} \right\|, \quad B_3 = \left\| \begin{array}{cc} 2 & -2 \\ -3 & 5 \end{array} \right\|.$$

Рассмотрим B_1 :

$$\text{adj } B_1 = \left\| \begin{array}{cc} -1 & -3 \\ -5 & -2 \end{array} \right\|, \quad (\text{adj})^T = \left\| \begin{array}{cc} -1 & -5 \\ -3 & -2 \end{array} \right\|,$$

$$\bar{X} = \frac{\|1, 1\| \left\| \begin{array}{cc} -1 & -3 \\ -5 & -2 \end{array} \right\|}{\|1, 1\| \left\| \begin{array}{cc} -1 & -3 \\ -5 & -2 \end{array} \right\| \left\| \begin{array}{c} 1 \\ 1 \end{array} \right\|} = \frac{\| -6, -5 \|}{\| -6, -5 \| \left\| \begin{array}{c} 1 \\ 1 \end{array} \right\|} = \frac{\| -6, -5 \|}{-11} = \left\| \frac{6}{11}, \frac{5}{11} \right\|,$$

$$\bar{Y} = \frac{\|1, 1\| \left\| \begin{array}{cc} -1 & -5 \\ -3 & -2 \end{array} \right\|}{-11} = \frac{\| -4, -7 \|}{-11} = \left\| \frac{4}{11}, \frac{7}{11} \right\|.$$

Так как $6/11 > 0$, $5/11 > 0$, $4/11 > 0$ и $7/11 > 0$, то мы продолжаем дальше:

$$X = \|6/11, 5/11\|, \quad Y = \|0, 4/11, 7/11\|,$$

или

$$g = \frac{\left\| \begin{array}{cc} -2 & 3 \\ 5 & -1 \end{array} \right\|}{-11} = \frac{2-15}{-11} = \frac{13}{11}.$$

Теперь проверяем основные соотношения:

$$2x_1 - 3x_2 = 2 \cdot 6/11 - 3 \cdot 5/11 = 12/11 - 15/11 = -3/11 < 13/11.$$

Это соотношение противоречит требованию $2x_1 - 3x_2 \geq g$, поэтому B_1 не подходит.

Рассмотрим теперь B_2 :

$$\text{adj } B_2 = \left\| \begin{array}{cc} -1 & -3 \\ 3 & 2 \end{array} \right\|, \quad (\text{adj } B_2)^T = \left\| \begin{array}{cc} -1 & 3 \\ -3 & 2 \end{array} \right\|,$$

$$\bar{X} = \frac{\|1, 1\| \left\| \begin{array}{cc} -1 & -3 \\ 3 & 2 \end{array} \right\|}{\|1, 1\| \left\| \begin{array}{cc} -1 & -3 \\ 3 & 2 \end{array} \right\| \left\| \begin{array}{c} 1 \\ 1 \end{array} \right\|} = \frac{\|2, -1\|}{\|2, -1\| \left\| \begin{array}{c} 1 \\ 1 \end{array} \right\|} = \frac{\|2, -1\|}{1} = \|2, -1\|.$$

Так как $x_{r2} = -1 < 0$, мы отбрасываем B_2 .

B_3 должна дать нам ответ, так как это единственная оставшаяся подматрица, а мы знаем, что решение существует.

$$\text{adj } B_3 = \begin{vmatrix} 5 & 2 \\ 3 & 2 \end{vmatrix}, \quad (\text{adj } B_3)^T = \begin{vmatrix} 5 & 3 \\ 2 & 2 \end{vmatrix},$$

$$\bar{X} = \frac{\|1, 1\| \begin{vmatrix} 5 & 2 \\ 3 & 2 \end{vmatrix} \begin{vmatrix} 1 \\ 1 \end{vmatrix}}{\|1, 1\| \begin{vmatrix} 5 & 2 \\ 3 & 2 \end{vmatrix} \begin{vmatrix} 1 \\ 1 \end{vmatrix} \begin{vmatrix} 1 \\ 1 \end{vmatrix}} = \frac{\|8, 4\|}{\|8, 4\| \begin{vmatrix} 1 \\ 1 \end{vmatrix}} = \frac{\|8, 4\|}{12} = \left\| \frac{2}{3}, \frac{1}{3} \right\|,$$

$$\bar{Y} = \frac{\|1, 1\| \begin{vmatrix} 5 & 3 \\ 2 & 2 \end{vmatrix} \begin{vmatrix} 1 \\ 1 \end{vmatrix}}{\|1, 1\| \begin{vmatrix} 5 & 2 \\ 3 & 2 \end{vmatrix} \begin{vmatrix} 1 \\ 1 \end{vmatrix} \begin{vmatrix} 1 \\ 1 \end{vmatrix}} = \frac{\|7, 5\|}{12} = \left\| \frac{7}{12}, \frac{5}{12} \right\|.$$

Так как $2/3 > 0$, $1/3 > 0$, $7/12 > 0$, $5/12 > 0$, продолжаем дальше:

$$X = \left\| 2/3, 1/3 \right\|, \quad Y = \left\| 7/12, 5/12, 0 \right\|,$$

$$x_1 = 2/3, \quad x_2 = 1/3, \quad y_1 = 7/12, \quad y_2 = 5/12, \quad y_3 = 0,$$

$$g = \frac{\begin{vmatrix} 2 & -2 \\ -3 & 5 \end{vmatrix}}{12} = \frac{10-6}{12} = \frac{4}{12} = \frac{1}{3},$$

$$2x_1 - 3x_2 = 2 \cdot 2/3 - 3 \cdot 1/3 = 1/3 = g,$$

$$-2x_1 + 5x_2 = -2 \cdot 2/3 + 5 \cdot 1/3 = 1/3 = g,$$

$$2y_1 - 2y_2 + 3y_3 = 2 \cdot 7/12 - 2 \cdot 5/12 + 3 \cdot 0 = 14/12 - 10/12 = 1/3 = g,$$

$$-3y_1 + 5y_2 - y_3 = -3 \cdot 7/12 + 5 \cdot 5/12 - 0 = -21/12 + 25/12 = 1/3 = g.$$

Следовательно, решением будет

$$x_1 = 2/3, \quad x_2 = 1/3, \quad y_1 = 7/12, \quad y_2 = 5/12, \quad y_3 = 0, \quad g = 1/3.$$

Оно совпадает с решением той же игры, полученным на стр. 411.

ИТЕРАТИВНЫЙ МЕТОД

Итеративный метод применяется тогда, когда неприменим графический метод и когда практически неприменимы алгебраический и матричный методы. Этот метод, который дает приближенное значение цены игры, причем истинное значение можно получить с любой нужной степенью точности. Как будет объяснено ниже, этот метод недостаточен для нахождения оптимальных стратегий.

Мы рассмотрим этот метод на примере. Пусть платежной матрицей будет матрица, показанная в таблице 18.13. Эта матрица эквивалентна матрице, приведенной в таблице 18.14.

Мы составим таблицу 18.15 из 16 колонок (см. стр. 416). Колонки (1) и (16) — гипотетические номера ходов. Колонки (2), (3), (9) и (10) — действительные ходы. Заметим, что

колонка (2) — колонка (10) — ход, выбранный игроком A ,

колонка (3) — колонка (9) — ход, выбранный игроком B .

Колонки (4), (5) и (6) — выигрыши A , соответствующие ходам B (колонка (3)). Колонки (11), (12) и (13) — выигрыши B , соответствующие ходам A (колонка (10)). Колонка (7) — лучший следующий ход A , основан-

ный на всех выборах B (колонки (4), (5) и (6)). Колонка (14) — лучший следующий ход B , основанный на всех выборах A (колонки (11), (12) и (13)). Заметим, что

Таблица 18.13

		B			
		План	S	T	U
A	P	0	2	0	
	Q	0	0	2	
	R	3	0	0	

Таблица 18.14

		A			
		План	P	Q	R
B	S	0	0	3	
	T	2	0	0	
	U	0	2	0	

колонка (7) для любого хода = колонкам (2) и (10) для следующего хода;

колонка (14) для любого хода = колонкам (3) и (9) для следующего хода.

Колонка (8) — наибольший полный выигрыш игрока A , колонка (15) — наименьший полный проигрыш игрока B , определяемый числом сделанных ходов.

Колонки (4), (5) и (6) для любого хода = колонкам (4), (5) и (6) для предыдущего хода плюс результат B при настоящем ходе.

Колонки (11), (12) и (13) для любого хода = колонкам (11), (12) и (13) для предыдущего хода плюс результат A при настоящем ходе.

Например, для десятого хода

P Q R					S T U				
(1)	(3)	(4)	(5)	(6)	(10)	(11)	(12)	(13)	(16)
9	$T =$	6	8	6	$Q =$	12	2	8	9
		2	0	0		0	0	2	
10		8	8	6		12	2	10	10

Каждый раз, когда мы имеем два равных хода, выбираем первый по порядку PQR или STU .

Решение.

(а) Начинаем с ходов P для A и S для B и заполняем соответствующими результатами колонки (1), (2), (3), (4), (5), (6), (9), (10), (11), (12), (13) и (16).

(б) Для A R дает наибольшую цену. Поэтому ставим R в (7) и в колонки (2) и (10) хода 2. Для B лучшими ходами являются S или U ; выбираем первый, т. е. S . Поэтому ставим S в (14) хода 1 и в (3) и (9) хода 2.

(в) Наибольший итог для A теперь 3. Поэтому в (8) ставим $3/1$. Наименьший итог для B теперь 0. Поэтому в (15) ставим $0/1$.

Таблица 18.15

Планирование А и В

Планирование А								Планирование В								Номер хода n
Номер хода n	Ходы		Платежи			Наибольший выигрыш		Ходы		Платежи			Наибольший выигрыш			
	A	B	P	Q	R	План	g/n	B	A	S	T	U	План	g/n		
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	
1	P	S	0	0	3	R	3/1	S	P	0	2	0	S	0/1	1	
2	R	S	0	0	6	R	6/2	S	R	3	2	0	U	0/2	2	
3	R	U	0	2	6	R	6/3	U	R	6	2	0	U	0/3	3	
4	R	U	0	4	6	R	6/4	U	R	9	2	0	U	0/4	4	
5	R	U	0	6	6	Q	6/5	U	R	12	2	0	U	0/5	5	
6	Q	U	0	8	6	Q	8/6	U	Q	12	2	2	T	2/6	6	
7	Q	T	2	8	6	Q	8/7	T	Q	12	2	4	T	2/7	7	
8	Q	T	4	8	6	Q	8/8	T	Q	12	2	6	T	2/8	8	
9	Q	T	6	8	6	Q	8/9	T	Q	12	2	8	T	2/9	9	
10	Q	T	8	8	6	P	8/10	T	Q	12	2	10	T	2/10	10	
11	P	T	10	8	6	P	10/11	T	P	12	4	10	T	4/11	11	
12	P	T	12	8	6	P	12/12	T	P	12	6	10	T	6/12	12	
13	P	T	14	8	6	P	14/13	T	P	12	8	10	T	8/13	13	
14	P	T	16	8	6	P	16/14	T	P	12	10	10	T	10/14	14	
15	P	T	18	8	6	P	18/15	T	P	12	12	10	U	10/15	15	
16	P	U	18	10	6	P	18/16	U	P	12	14	10	U	10/16	16	
17	P	U	18	12	6	P	18/17	U	P	12	16	10	U	10/17	17	
18	P	U	18	14	6	P	18/18	U	P	12	18	10	U	10/18	18	
19	P	U	18	16	6	P	18/19	U	P	12	20	10	U	10/19	19	
20	P	U	18	18	6	P	18/20	U	P	12	22	10	U	10/20	20	
21	P	U	18	20	6	Q	20/21	U	P	12	24	10	U	10/21	21	
22	Q	U	18	22	6	Q	22/22	U	Q	12	24	12	S	12/22	22	
23	Q	S	18	22	9	Q	22/23	S	Q	12	24	14	S	12/23	23	
24	Q	S	18	22	12	Q	22/24	S	Q	12	24	16	S	12/24	24	
25	Q	S	18	22	15	Q	22/25	S	Q	12	24	18	S	12/25	25	
26	Q	S	18	22	18	Q	22/26	S	Q	12	24	20	S	12/26	26	
27	Q	S	18	22	21	Q	22/27	S	Q	12	24	22	S	12/27	27	

(г) Прибавляем результаты хода 2 к ходу 1.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
1	P	S	0	0	3	R	3/1	S	S	0	2	0	S	0/1	1
		S=0	0	0	3				R=3	0	0				
2	R	S	0	0	6			S	R	3	2	0			

(д) Повторяем шаг (б) для хода 2.

(е) Повторяем шаг (в) для хода 2.

(ж) Повторяем шаг (г) для хода 2 и получаем данные для хода 3 и т. д. Мы приходим к следующим результатам:
 Наименьший выигрыш A 0,80 (ход 10),
 Наибольший проигрыш B 0,71 (ход 14).
 Таким образом,

$$0,71 \leq g \leq 0,80.$$

A играл:	P 12 раз	B играл:	S 7 раз
	Q 11 раз		T 9 раз
	R 4 раза		U 11 раз
Итого 27 раз		Итого 27 раз	

Таким образом,

$$\begin{aligned} x_1 &\approx 12/27 = 0,44, & y_1 &\approx 7/27 = 0,26, \\ x_2 &\approx 11/27 = 0,41, & y_2 &\approx 9/27 = 0,33, \\ x_3 &\approx 4/27 = 0,15, & y_3 &\approx 11/27 = 0,41. \end{aligned}$$

Это приближенные результаты, основанные только на 27 ходах. Большая точность может быть получена при увеличении числа ходов.

Точные решения игры можно найти (с относительной легкостью):

(а) алгебраическим методом, заменив все соотношения равенствами;

(б) матричным методом, приняв $B = \begin{vmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 3 & 0 & 0 \end{vmatrix}$.

Результаты следующие:

$$\begin{aligned} x_1 &= 0,375, & y_1 &= 0,250, \\ x_2 &= 0,375, & y_2 &= 0,375, \\ x_3 &= 0,250, & y_3 &= 0,375, \\ g &= 0,75. \end{aligned}$$

Следует отметить, что хотя этот метод может быть использован для получения цены игры с любой необходимой степенью точности, он может оказаться непригодным при определении оптимальных стратегий. Более детальное рассмотрение этого вопроса см. в [22, часть 4].

ИГРЫ n ЛИЦ С НУЛЕВОЙ СУММОЙ

Коалиции. Характеристиками игры n лиц с нулевой суммой являются цены игры для различных коалиций, которые могут образовывать игроки. Если, например, в игре принимают участие игроки A , B , C и D , то могут быть образованы коалиции, показанные в таблице 18.16.

Таблица 18.16

Различные коалиции

	Группа I против группы II			Группа I против группы II	
1	A, B, C	D	5	A, B	C, D
2	A, B, D	C	6	A, C	B, D
3	A, C, D	B	7	A, D	B, C
4	B, C, D	A			

Если цена игры для A, B, C есть g , то цена игры для D будет $-g$, так как рассматривается игра с нулевой суммой. Итак, в игре четырех лиц должно быть семь цен, или характеристик, игры.

В самом деле, каждая цена получается из решения прямоугольной игры, в которой две коалиции являются двумя противниками. Найдем характеристики следующей игры трех лиц с нулевой суммой:

У игрока A есть два возможных плана: P_1, P_2 .

У игрока B есть два возможных плана: Q_1, Q_2 .

У игрока C есть два возможных плана: R_1, R_2 .

Платежи в зависимости от выбора планов показаны в таблице 18.17.

Таблица 18.17

Платежная матрица

Выбор			Платеж			Выбор			Платеж		
A	B	C	A	B	C	A	B	C	A	B	C
P_1	Q_1	R_1	2	1	-3	P_2	Q_1	R_1	3	-2	-1
P_1	Q_1	R_2	-1	1	0	P_2	Q_1	R_2	-2	0	2
P_1	Q_2	R_1	-1	-2	3	P_2	Q_2	R_1	0	-1	1
P_1	Q_2	R_2	0	2	-2	P_2	Q_2	R_2	-1	1	0

Таким образом, мы имеем три возможных коалиции:

1. A против B, C .
2. B против A, C .
3. C против A, B .

Решим каждую из соответствующих прямоугольных игр.

1. A против B, C . Платежная матрица в выражениях для A показана в таблице 18.18. Эта матрица имеет седловую точку. Таким образом, мы имеем решение для A против B, C :

Таблица 18.18

 A против B, C

		B, C			
План		Q_1, R_1	Q_1, R_2	Q_2, R_1	Q_2, R_2
A	P_1	2	-1	-1	0
	P_2	3	-2	0	-1

Лучшим планом игрока A является план P_1 .

Лучшая комбинация планов для игроков B и C — Q_1, R_2 .

Цена игры для A есть $g(A) = -1$.

Цена игры для B, C есть $g(B, C) = 1$.

2. B против A, C . Исследовав таблицу 18.19, мы находим, что A, C никогда не будут играть комбинацию P_1, R_2 или P_2, R_2 . Таким образом, матрицей игры будет матрица, показанная в таблице 18.20.

Таблица 18.19

В против A, C

		A, C			
		План	P_1, R_1	P_1, R_2	P_2, R_1
B	Q_1	1	1	-2	0
	Q_2	-2	2	-1	1

Таблица 18.20

		A, C		
		y	1-y	
		План	P_1, R_1	P_2, R_1
B	x	Q_1	1	-2
	1-x	Q_2	-2	-1

Решая алгебраически, получаем

$$x - 2(1 - x) = -2x - (1 - x).$$

Отсюда

$$x - 2 + 2x = -2x - 1 + x, \text{ т. е. } x = 1/4.$$

Следовательно,

$$g(B) = 1/4 - 2(1 - 1/4) = 1/4 - 6/4 = -5/4.$$

Точно так же

$$y - 2(1 - y) = -2y - (1 - y),$$

т. е.

$$y - 2 + 2y = -2y - 1 + y, \text{ т. е. } y = 1/4.$$

Решение для B против A, C следующее.

 Оптимальная стратегия игрока B — играть план Q_1 с частотой $1/4$ и план Q_2 с частотой $3/4$.

 Оптимальная стратегия игроков A и C — для C играть R_1 , а для A играть P_1 с частотой $1/4$ и P_2 с частотой $3/4$.

 Цена игры для B есть $g(B) = -1,25$.

 Цена игры для A, C есть $g(A, C) = 1,25$.

3. C против A, B. Исследовав таблицу 18.21, находим, что A, B никогда

Таблица 18.21

C против A, B

		A, B				
		y	1-y	0	0	
		План	P_1, Q_1	P_1, Q_2	P_2, Q_1	P_2, Q_2
C	x	R_1	-3	3	-1	1
	1-x	R_2	0	-2	2	0

не будут играть комбинации P_2, Q_1 или P_2, Q_2 . Поэтому

$$\begin{aligned} -3x + 0(1-x) &= 3x - 2(1-x), \\ -3x &= 3x - 2 + 2x, \text{ или } x = 1/4, \\ g(C) &= -3 \cdot 1/4 = -3/4, \\ -3y + 3(1-y) &= 0 - 2(1-y), \\ -3y + 3 - 3y &= -2 + 2y, \text{ т. е. } y = 5/8. \end{aligned}$$

Таким образом, решение C против A, B следующее.

Оптимальная стратегия для C — играть R_1 с частотой $1/4$ и R_2 с частотой $3/4$.

Оптимальная стратегия A и B : для A — играть P_1 , а для B — играть Q_1 с частотой $5/8$ и Q_2 с частотой $3/8$.

Цена игры для C есть $g(C) = -0,75$.

Цена игры для A, B есть $g(A, B) = 0,75$.

Характеристики игры следующие:

$$\begin{aligned} g(A) &= -1,0, & g(B, C) &= 1,0, \\ g(B) &= -1,25, & g(A, C) &= 1,25, \\ g(C) &= -0,75, & g(A, B) &= 0,75. \end{aligned}$$

Мы не будем заниматься построением оптимальных стратегий игроков, когда они не образуют коалиции. Наше изучение игр основывалось на предположении, что игрок планирует свою стратегию против наилучшей игры противника. В этом случае противником игрока A , например, является коалиция BC , и поэтому мы должны исследовать оптимальную стратегию коалиции, что мы и сделали.

Следующий пункт основывается на классическом подходе фон Неймана и Моргенштерна [18]. Ниже мы дадим их определение решения игры n лиц с нулевой суммой, где $n \geq 3$.

Обязательства. Рассмотрим платежную матрицу, показанную в таблице 18.17. Если A, B и C выбирают планы P_1, Q_1 и R_1 , то выигрыши игроков будут 2, 1 и -3 соответственно. C выигрывает -3 (т. е. проигрывает три единицы), и A и B выигрывают 2 и 1 соответственно.

Разумно ли ожидать, что C будет считать этот результат своим нормальным выигрышем? Конечно, нет. В действительности мы знаем, что даже когда он играет один против A и B в коалиции, он может выбрать свой план так, что ожидаемый проигрыш будет не больше 0,75 (см. стр. 417). Точно так же A и B не согласятся считать нормальным выигрышем любой выигрыш, который вынуждает их платить больше, чем 1,00 и 1,25 соответственно.

Рассуждая подобным образом, можно допустить такие выигрыши, как, скажем, $A = 0,50$, $B = -0,25$ и $C = -0,25$. В этом случае ни один из выигрышей игроков не меньше его цены игры. Так, $0,50 > -1,00$; $-0,25 > -1,25$; $-0,25 > -0,75$. Мы называем такой нормальный выигрыш *обязательством*.

Для игры n лиц в математических символах это можно описать так. Пусть $g_1, g_2, g_3, \dots, g_n$ — цены игры для игроков 1, 2, $\dots, n$ соответственно, когда каждый игрок играет против коалиции всех других игроков. Тогда $[x_1, x_2, \dots, x_n]$ — обязательство, если

$$x_i \geq g_i \quad \text{для} \quad i = 1, 2, \dots, n$$

и $\sum x_i = 0$ (так как игра с нулевой суммой). В примере, данном выше, следующие наборы выигрышей являются обязательствами:

- (1) $[-0,75, 0,25, 0,50]$, (3) $[0,30, -0,80, 0,50]$,
 (2) $[2,0, -1,25, -0,75]$, (4) $[-0,15, 0,85, -0,70]$.

В то же время наборы

- (5) $[0,50, 0,50, -1,00]$ (так как $-1,00 < -0,75$),
 (6) $[0,26, -1,26, 1,00]$ (так как $-1,26 < -1,25$)

обязательствами не являются. Игроки B и C (не в коалиции) предпочтут обязательство (1) обязательству (2) или обязательство (4) обязательству (2).

Игроки A и B (не в коалиции) предпочтут обязательство (4) обязательству (1).

Мы используем термин *доминирование* для обозначения этого предпочтения. Мы говорим: (1) *доминирует* над (2) по отношению к игрокам B и C , (4) доминирует над (1) по отношению к игрокам A и B .

Отношение доминирования используется только для двух игроков или больше и символически записывается следующим образом:

$$\begin{aligned} [0,75, 0,25, 0,50] &\succ_{B,C} [2,0, -1,25, -0,75], \\ [-0,15, 0,85, -0,75] &\succ_{A,B} [-0,75, 0,25, 0,50]. \end{aligned}$$

Вообще для обязательств

$$[y_1, y_2, \dots, y_n] \succ_{i=i_1, i_2, \dots, i_r} [x_1, x_2, \dots, x_n]$$

при условии

$$y_{ik} > x_{ik} \quad \text{при} \quad k = 1, 2, \dots, r.$$

Мы говорим также об одном обязательстве, доминирующем над другим. Это выражение означает, что существует группа игроков (два или более), для которых осуществляется доминирование. Так, (3) $\succ$ (4) (группа игроков A, C). Но между обязательствами (1) и (3) не существует отношения доминирования, так как нет такой группы игроков, для которых было бы справедливым доминирование.

Определение решения. Теперь мы в состоянии дать определение фон Неймана и Моргенштерна для игры n лиц с нулевой суммой. Решением данной игры n лиц является такое множество обязательств, что:

1. Ни одно обязательство этого множества не доминирует над другим обязательством того же множества.

2. Для любого обязательства, не принадлежащего этому множеству, найдется обязательство нашего множества, доминирующее над ним.

В примере, который мы здесь рассматривали, решение состоит из следующего множества трех обязательств:

$$\begin{aligned} K &= [0,50, 0,25, -0,75], \\ L &= [0,50, -1,25, 0,75], \\ M &= [-1,0, 0,25, 0,75], \end{aligned}$$

поскольку они удовлетворяют двум предыдущим требованиям. Сравним их, например, с обязательствами, данными выше: $L \succ (1)$ по отношению к A, B ; $M \succ (2)$ по отношению к B, C ; $L \succ (3)$ по отношению к A, C ; $L \succ (4)$ по отношению к A, C .

Может быть показана достоверность следующих результатов.

1. Каждая игра трех лиц имеет решение. Решение может состоять из (а) единственного конечного множества из трех обязательств (как было в приведенном примере), (б) бесконечного множества обязательств; таких множеств существует бесконечно много.

2. Каждая игра четырех лиц имеет решение.

ИГРЫ БЕЗ ОГРАНИЧЕНИЯ НА СУММУ ВЫИГРЫШЕЙ

Обязательства. Игра без ограничения «нулевой суммы» — это такая игра, в конце которой суммарные платежи не обязательно должны быть равны нулю. Такие игры могут представлять экономические ситуации, в которых конечный результат, например в случае двух конкурирующих компаний, не обязательно сводится к тому, что одна компания выигрывает точно столько, сколько теряет другая. Легче всего представить игру n лиц без ограничения нулевой суммы как игру $n + 1$ лица с нулевой суммой, в которой $(n + 1)$ -й игрок — фиктивный игрок, выигрыш которого дополняет суммарный выигрыш остальных игроков до нулевой суммы. В случае двух конкурирующих компаний можно допустить, что третьим игроком, который обеспечивает нулевую сумму игре, является природа. Необходимо, однако, отметить, что введение $(n + 1)$ -го игрока в данную игру n лиц не сводит решение к решению игры $n + 1$ лица с нулевой суммой, так как $(n + 1)$ -й игрок не волен вступать в любые возможные коалиции. Мы рассматриваем $(n + 1)$ -го игрока только с точки зрения его наибольшего возможного проигрыша.

Эти соображения приводят нас к следующему определению обязательства. Пусть $g_1, g_2, \dots, g_n$ будут цены игры для игроков $1, 2, \dots, n$ соответственно. Пусть g_{n+1} будет ценой игры для фиктивного игрока $n + 1$. Напомним, что цена игры g игроков $1, 2, \dots, n$, входящих в коалицию, есть

$$G = g(1, 2, \dots, n) = -g_{n+1}.$$

Мы называем $[x_1, x_2, \dots, x_n]$ обязательством, если

$$x_i \geq g_i \quad \text{для } i = 1, 2, \dots, n$$

и

$$\sum_{i=1}^n x_i = G.$$

Например, пусть в игре двух лиц $g_1 = -1$, $g_2 = 0$, $g_{n+1} = g_3 = -3$. Тогда $G = -g_3 = 3$ и $[1, 2]$, $[-1, 4]$, $[3, 0]$ являются обязательствами, а следующие пары обязательств не являются. $[-4, -1]$, так как $-1 < g_2 = 0$; $[1, 1]$, так как $1 + 1 = 2 < G = 3$.

Заметим, что это определение не противоречит определению, данному для игр с нулевой суммой. Для этих игр мы имеем $G = 0$. Понятие доминанты, что можно легко проверить, также не затрагивается этим определением.

Решение игры n лиц. Определение решения игры n лиц без ограничения нулевой суммы то же самое, что и для игр с нулевой суммой, а именно:

решением является множество обязательств (в смысле данного выше определения) такое, что:

Таблица 18.22

Игра двух лиц
с ненулевой суммой

План		Выигрыш	
А	В	А	В
P_1	Q_1	3	-1
P_1	Q_2	-1	1
P_2	Q_1	1	0
P_2	Q_2	-2	-1

1. Ни одно обязательство множества не доминирует над другим обязательством этого множества.

2. Для любого обязательства, не входящего в множество, найдется обязательство множества, доминирующее над ним.

Пример I. Рассмотрим игру двух лиц с ненулевой суммой, приведенную в таблице 18.22. Платежная матрица игрока A имеет вид

		Q_1	Q_2
A	P_1	3	-1
	P_2	1	-2

Эта матрица имеет седловую точку, так что

$$g(A) = -1.$$

Платежная матрица игрока B следующая:

		P_1	P_2
B	x	-1	0
	$1-x$	1	-1

Эта игра не имеет седловой точки. Решая эту игру, получаем

$$x(-1) + (1-x) \cdot 1 = x(0) + (1-x)(-1), \text{ т. е. } x = 2/3,$$

откуда

$$g(B) = 2/3(-1) + 1/3 \cdot 1 = -1/3.$$

Для $n+1$ игрока матрица будет иметь вид

	P_1, Q_1	P_1, Q_2	P_2, Q_1	P_2, Q_2
C	-2	0	-1	3

У этой матрицы вновь есть седловая точка. Поэтому

$$g(C) = -2, \quad G = 2.$$

Таким образом, решение игры составляют любые обязательства $[x_1, x_2]$, для которых

$$x \geq -1, \quad x_2 \geq -1/3, \quad x_1 + x_2 = 2.$$

Пример II. Пусть в игре трех лиц с ненулевой суммой будут следующие цены игры:

$$g(A) = 0, \quad g(B) = -1, \quad g(C) = -2, \quad G = 1.$$

Решение игры может быть получено из геометрического анализа рис. 18.5, который понятен без дополнительных пояснений.

Любая точка треугольника LMN является обязательством, например, $P = [1,5, 0,5, -1]$ есть обязательство, так как

$$x_1 = 1,5 > 0 = g(A), \quad x_2 = 0,5 > -1 = g(B), \quad x_3 = -1 > -2 = g(C),$$

$$\sum x_i = x_1 + x_2 + x_3 = 1,5 + 0,5 - 1 = G.$$

Заметим, что любая точка (обязательство) из области, покрытой точками, доминируется обязательством P , а любая точка из области, не покрытой точками, доминирует над P .

Решение игры задается точками D , E и F :

$$D = [2, 1, -2], \quad E = [0, 1, 0],$$

$$F = [2, -1, 0].$$

Можно легко удостовериться в том, что ни одно из этих обязательств не доминирует над другим и что одно из обязательств D , E или F доминирует над любым другим обязательством.

ЗАКЛЮЧЕНИЕ

За последние годы появилось несколько книг по теории игр.

Маккинси в «Введении в теорию игр» [13] дает прекрасный математический анализ общей теории с основным упором на игры двух лиц. Он также рассматривает связь теории игр с линейным программированием и теорией статистических решений. Эта работа содержит довольно полную, включающую последние работы, библиографию по теории игр. Дополнительно в последней части [13] Маккинси обсуждает ряд нерешенных задач в этой области.

Блэкуэлл и Гиршик [1] написали книгу по вопросам теории игр, «...предназначенную в первую очередь в качестве учебника по теории принятия решений для аспирантов первого года, специализирующихся по статистике», и требующую большей математической подготовки, чем упоминавшаяся ранее книга фон Неймана и Моргенштерна*. Хотя их исследования по теории игр — только средство получения результатов теории статистических решений, эта работа является важным шагом вперед в применении теории игр к промышленным проблемам. (Работа содержит также прекрасную библиографию.)

Некоторые из наиболее важных основных направлений исследования в области теории игр представлены в двух томах под редакцией Куна и Такера [7, 8]. Существенный вклад в эту область представляют многочисленные отчеты Беллмана, Боненбласта, Дрешера, Харриса, Хелмера, Карлина, Шэпли, Шоу и др., изданных РЭНД Корпорейшн. Важная фор-

* Ранняя элементарная нематематическая публикация теории фон Неймана дана Макдональдом [11]. Следующая по математической сложности — очень остроумная и ясная работа Вильямса [22].

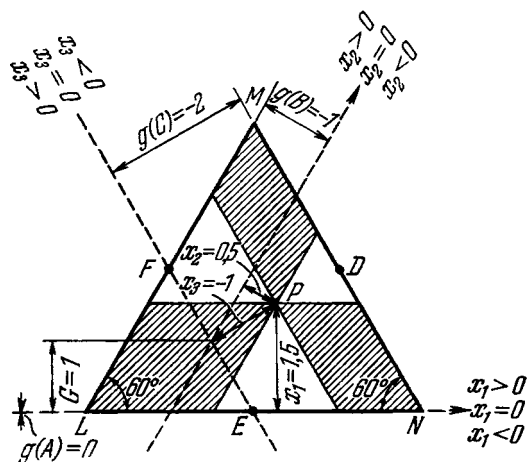


Рис. 18.5. Графическое решение игры трех лиц с ненулевой суммой.

мальная связь между линейным программированием и теорией игр показана Данцигом [5].

Очень мало было сделано в области приложения теории. Можно было бы сослаться на военные приложения, но они не были опубликованы. Такие статьи, как работы Макдональда [9], Шубика [20], Кэйвуда и Томаса [3], не рассматривают реальных промышленных приложений, но касаются возможности таких приложений.

В заключение стоит привести краткую оценку теории игр, данную Вильямсом [22, стр. 217]: «...хотя в настоящее время уже выяснены, несмотря на множество ограничений теории, многие ее специфические приложения, ее наибольший, пока неясный, вклад состоит, может быть, в том, что она дает людям, имеющим дело с сверхсложными проблемами, самую общую ориентацию. Даже если эти проблемы не поддаются строгому решению (в настоящее время это несомненно так, а возможно, так будет и в неопределенном будущем), она дает основу для работы над ними. Идея стратегии, различия между игроками, роль случайных событий, понятие матрицы выигрышей, идеи чистой и смешанной стратегии и т. д. дают драгоценную ориентацию лицам, которым необходимо обдумывать сложные конфликтные ситуации».

ЛИТЕРАТУРА

1. Blackwell D. and Girshick M. A., *Theory of Games and Statistical Decisions*, John Wiley & Sons, New York, 1954.
2. Brown G. W., *Iterative Solution of Games by Fictitious Play*, Chap. XXIV in T. C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, Cowles Commission Monograph No. 13, John Wiley & Sons, New York, 1951.
3. Caywood T. E. and Thomas C. J., *Applications of Game Theory in Fighter Versus Bomber Combat*, J. Opns. Res. Soc. Am., 3, no. 4, 402—411 (1955).
4. Cushen W. E., *Operational Gaming*, in J. F. McCloskey and J. M. Coppinger (eds.), *Operations Research for Management*, II, The Johns Hopkins Press, Baltimore, 1956.
5. Dantzig G. B., *A Proof of the Equivalence of the Programming Problem and the Game Problem*, Chap. XX in ref. 2.
6. Dresher M., *Games of Strategy*, Math. Mag., 25, 93—99 (1951).
7. Kuhn H. W. and Tucker A. W. (eds.), *Contributions to the Theory of Games*, I, Annals of Mathematics Study No. 24, Princeton University Press, Princeton, 1950.
8. Kuhn H. W. and Tucker A. W. (eds.), *Contributions to the Theory of Games*, II, Annals of Mathematics Study No. 28, Princeton University Press, Princeton, 1953.
9. McDonald John, *Applications of Game Theory in Business and Industry*, Games Symposium of the American Association for the Advancement of Science, St. Louis, Dec. 29, 1952.
10. McDonald John, *Poker: An American Game*, Fortune, 37, 128—131 and 181—187 (1948).
11. McDonald John, *Strategy in Poker, Business, and War*, W. W. Norton and Co., New York, 1950.
12. McDonald John, *The Theory of Strategy*, Fortune, 38, 100—110 (1949).
13. McKinsey J. C. C., *Introduction to the Theory of Games*, McGraw-Hill Book Co., New York, 1952.
14. Morgenstern O., *The Theory of Games*, Sci. Amer., 180, 22—25 (1949).
15. Nash J. F., *Equilibrium Points in n -Person Games*, Proc. Nat. Acad. Sci., 38, 48—49 (1950).
16. Nash J. F., *Non-Cooperative Games*, Ann. Math., 54, 286—295 (1951).
17. Neumann J. von, *Zur Theorie der Gesellschaftsspiele*, Math. Ann., 100, 295—320 (1928).
18. Neumann J. von and Morgenstern O., *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 3rd ed., 1953.

19. Paxson E. W., Recent Developments in the Mathematical Theory of Games. *Econometrica*, 17, 72—73 (1949).
20. Shubik Martin, The Uses of Game Theory in Management Science, *Mgmt. Sci.*, 2, no. 1, 40—54 (1955).
21. Stone R., The Theory of Games, *Econ. J.*, 58, 185—201 (1948).
22. Williams J. D., The Compleat Strategyst, McGraw-Hill Book Co., New York, 1954.

Глава 19

МОДЕЛИ ТОРГОВ

Торги за приобретение прав на собственность или за преимущества при предоставлении услуг являются типичным примером состязательной политики. Конкуренты соревнуются за обладание правами или преимуществами согласно правилам, устанавливаемым государственным органом и (или) лицами, выставившими их на аукцион. В той или иной форме в деловых отношениях торги существуют всегда: торги за получение контрактов, концессий, за лицензии на право использования патентов и т. д. В конечном счете назначение цен на товары или услуги можно также рассматривать как торги за деньги покупателей.

Этому аспекту состязательного поведения до сих пор не уделялось достаточного внимания с точки зрения исследования операций. В этой области потенциальные возможности и мощь операционного подхода вообще еще не были показаны. К сожалению, из соображений промышленной секретности подробности успешного применения методов исследования операций к разработке стратегии торгов не могли быть опубликованы. Поэтому любое обсуждение этих проблем должно быть либо совершенно абстрактным, либо основываться на переформулированном изложении реальной задачи.

Ряд простых задач, возникающих при исследовании проблемы торгов, может быть решен с использованием методов теории игр. Но в ситуации, когда число участников торгов становится большим или оно точно не известно, подобный подход уже неприменим. В этом случае можно успешно использовать аналитические методы, некоторые из которых будут изложены в этой главе. Заметим, что даже простейшие задачи торгов легче рассмотреть, применяя ниже описанные аналитические приемы, чем используя матричные методы, ставшие классическим аппаратом теории игр.

Существуют два вида торгов, и поэтому возникают две различные состязательные ситуации.

В первой, называемой *закрытыми торгами*, два или более конкурента независимо предлагают цены (ставки) за приобретение прав на собственность или предоставление услуг. В большинстве случаев каждый участник торгов имеет право только на одну ставку и ведущий торги принимает высшую или низшую цену согласно действующим правилам. Другой вид торгов — это *аукционы*, или *открытые торги*. При аукционных торгах два или более конкурента продолжают открыто, в присутствии друг друга, предлагать цены, пока никто больше не пожелает увеличить ставку. Участник торгов, внесший последнее предложение цены, считается выигравшим торги.

В качестве введения к построению моделей закрытых и аукционных (открытых) торгов рассмотрим сейчас два простых частных примера. Затем перейдем к исследованию одного случая закрытых торгов, более сложного, но зато и более близкого к реальной действительности, — торгов нескольких участников за получение контракта.

АУКЦИОННЫЕ ТОРГИ: ДВА ЛИЦА И ДВА ОБЪЕКТА

Рассмотрим сначала простой случай двух объектов известной стоимости V_1 и V_2 , предназначенных для продажи с аукциона один после другого. Предположим, что имеется только два покупателя A и B , располагающие для торгов S_A и S_B долларами соответственно. Предположим также, что каждая из величин S_A и S_B меньше суммы $V_1 + V_2$ и $1/2 < S_A/S_B < 2$. Мы предположим также, что A знает, какой суммой денег располагает B *.

Покупатель A желает знать, когда он должен прекратить поднимать цену, торгуясь за первый объект. Для решения этой задачи A должен в первую очередь определить, какую цель преследует он, участвуя в этих торгах. Наиболее очевидной целью является максимизация собственного дохода. Но возможно, что A имеет целью минимизацию дохода B или его цель состоит в минимизации чистой разности между доходами A и B . Интересно отметить, что последние две цели ведут к различным оптимальным политикам. В качестве первого примера рассмотрим случай, когда целью A является максимизация разности $R_A - R_B$, где R_A и R_B — доходы A и B соответственно.

Максимизация разности доходов. Предположим, что B предложил X долларов за первый из двух предназначенных для продажи объектов и минимальное допустимое увеличение цены есть Δ . A должен решить, предложить ли цену $X + \Delta$ или позволить B выиграть торги с ценой X . A может рассуждать следующим образом.

1. Если A позволит B получить первый объект за X долларов, то для второго объекта у B останется $S_B - X$ долларов. Если $S_A > S_B - X$ **, то A наверняка способен получить второй объект за $S_B - X + \Delta$. Тогда, если A позволяет B получить первый объект за X ,

$$R_B = V_1 - X, \quad (1)$$

$$R_A = V_2 - (S_B - X + \Delta), \quad (2)$$

и разность доходов A и B равна в этом случае

$$(V_2 - S_B + X - \Delta) - (V_1 - X).$$

2. Если A предлагает $X + \Delta$ и B позволяет ему получить первый объект по этой цене, то

$$R_A = V_1 - (X + \Delta), \quad (3)$$

$$R_B = V_2 - [S_A - (X + \Delta) + \Delta]. \quad (4)$$

Поскольку A будет иметь лишь $S_A - (X + \Delta)$ после первой продажи, B выигрывает вторые торги за $S_A - (X + \Delta) + \Delta$, и разность между доходами A и B будет равна

$$R_A - R_B = [V_1 - (X + \Delta)] - (V_2 - S_A + X). \quad (5)$$

Таким образом, A должен позволить B получить первый объект за X долларов, если разность между доходами A и B в этом случае больше, чем когда A идет на повышение и предлагает цену $X + \Delta$. Тогда условие, что

* Так как мы ищем оптимальную стратегию для A , то нет нужды знать, имеет ли B сведения об A .

** Это следует из предположений, что $1/2 < S_A/S_B < 2$ и лица A и B — разумные игроки.

A должен предложить $X + \Delta$, имеет вид

$$(V_2 - S_B + X - \Delta) - (V_1 - X) \leq (V_1 - X - \Delta) - (V_2 - S_A + X) \quad (6)$$

или

$$4X \leq 2V_1 - 2V_2 + S_A + S_B.$$

Отсюда A предлагает $X + \Delta$, если

$$X \leq \frac{2(V_1 - V_2) + (S_A + S_B)}{4}. \quad (7)$$

В этих условиях, если B действует разумным образом и торгуется корректно, A будет повышать цену до значения X , определяемого выражением

$$X = \frac{2(V_1 - V_2) + (S_A + S_B)}{4}. \quad (8)$$

Разность между доходами A и B находится подстановкой значения X из выражения (8) в выражение (5):

$$R_A - R_B = \frac{S_A - S_B}{2} - \Delta. \quad (9)$$

Эта разность не зависит от стоимости предложенных к продаже объектов и при $\Delta \rightarrow 0$ зависит лишь от разности между начальными капиталами двух покупателей.

Доход A равен

$$R_A = \frac{V_1 + V_2}{2} - \frac{(S_A + S_B)}{4} - \Delta. \quad (10)$$

Для иллюстрации этих результатов предположим, что A и B располагают для торгов 100 и 110 долларами соответственно и что каждый знает, какой суммой располагает другой. Пусть B предложил цену 60 долларов за первый из двух объектов. A должен тогда решить, предлагать ли 61 доллар или не предлагать. Из неравенства (7) следует, что A должен предложить 61 доллар, если

$$\$60 \leq \frac{2(V_1 - V_2) + (\$100 + \$110)}{4}.$$

Предположим, что A оценивает первый объект в 75 долларов, второй — в 100 долларов. Тогда

$$\frac{2(\$75 - \$100) + (\$100 - \$110)}{4} = \$40 < \$60.$$

Таким образом, A не следует предлагать за первый объект более 40 долларов. Следовательно, он должен позволить B получить первый объект за 60 долларов, и при этом он уверен, что получит второй объект не более чем за

$$(\$110 - \$60 + \$1) = \$51.$$

Исходя из оценки A стоимости объектов, чистый доход B будет равен $75 - 60 = 15$ долларов, а чистый доход A $100 - 51 = 49$ долларов. Тогда A закончит торги лучше B на $49 - 15 = 34$ доллара.

Если, однако, A предложил бы 61 доллар за первый объект и выиграл торги, то его чистый доход был бы $75 - 61 = 14$ долларов, а B смог бы получить второй объект за $100 - 61 + 1 = 40$ долларов. Чистый доход B был бы $100 - 40 = 60$ долларов, и разность доходов равнялась бы $60 - 14 = 46$ долларов в пользу B .

Необходимо заметить, что если оценки B стоимости продаваемых объектов отличны от оценок A , то относительная выгода различных возможностей будет для него иной.

Максимизация ожидаемого дохода. Рассмотрим теперь случай, когда целью A является максимизация собственного дохода. A теперь предложит цену $X + \Delta$ за первый объект, когда ожидаемый при этом им доход будет больше или равен доходу, который он получит, если позволит B получить первый объект за цену X , т. е. когда

$$V_1 - (X + \Delta) \geq V_2 - (S_B - X + \Delta) \quad (11)$$

или, что то же самое, когда

$$X \leq \frac{V_1 - V_2 + S_B}{2}. \quad (12)$$

Подобным же образом, если B максимизирует при торгах свой доход, он предложит цену $X + \Delta$, когда

$$X \leq \frac{V_1 - V_2 + S_A}{2}. \quad (13)$$

Если A и B торгуются корректно, последнее предложение цены будет Δ плюс наименьшая из величин

$$\frac{V_1 - V_2 + S_B}{2} \quad \text{и} \quad \frac{V_1 - V_2 + S_A}{2}.$$

Если $S_A > S_B$, доход A при рассматриваемой цели можно найти, положив $X = (V_1 + V_2 + S_B)/2$ в левой части неравенства (11). Это дает

$$R_A = \frac{V_1 + V_2 - S_B}{2} - \Delta \quad (S_A > S_B). \quad (14)$$

Если $S_A < S_B$, доход A находят, положив $X = (V_1 - V_2 + S_A)$ в левой части неравенства (11). В этом случае

$$R_A = \frac{V_1 + V_2 - S_A}{2} - \Delta \quad (S_A < S_B). \quad (15)$$

Заметим, что доход A (определяемый равенствами (14) или (15)), полученный при использовании в качестве цели максимума собственного дохода, всегда превосходит его доход (уравнение (9)), получаемый, когда целью является максимизация разности между доходами A и B .

Рассмотрим теперь ранее приведенный пример, но предполагая, что целью A является максимизация ожидаемого собственного дохода, т. е. поведение A никак не связано с поведением B . Мы показали, что если B получает первый объект за 60 долларов, а A — второй объект за 51 доллар, то доход A равен $100 - 51 = 49$ долларов.

Если же A получает первый объект за 61 доллар, то его доход будет только $75 - 61 = 14$ долларов.

В терминах максимального дохода наилучшее решение, которое следует принять, очевидно, и в данном случае совпадает с решением, основанным на принципе максимума разности.

Используя неравенство (12), можно установить, что A не должен предлагать за первый объект более чем 42,50 доллара, если он хочет максимизировать свой доход, так как

$$\frac{\$75 - \$100 + \$110}{2} = \$42,50.$$

С другой стороны, если его целью является максимизация разности доходов, A не должен предлагать больше чем 40 долларов.

ЗАКРЫТЫЕ ТОРГИ: ДВА ЛИЦА И ДВА ОБЪЕКТА

Рассматриваемая ситуация похожа на описанную ранее, с тем изменением, что оба объекта стоимостью V_1 и V_2 торгуются одновременно двумя конкурентами, каждый из которых участвует в двух закрытых торгах.

Чтобы сделать этот пример более простым, предположим, что A и B располагают каждый S долларами для покупки обоих объектов и $S < (V_1 + V_2)$. В случае равенства предложений мы будем предполагать, что победитель торгов определяется жребием.

Правильную стратегию можно определить, рассматривая рис. 19.1. Оптимальные цены, предлагаемые A за объекты 1 и 2, A_1 и A_2 соответственно, находятся путем назначения цен A_1 и A_2 таким образом, что доходы от успешных торгов каждого из объектов равны. Таким образом, если мы через d обозначим общий доход от успешного завершения торгов одного объекта, то

$$d = \frac{V_1 + V_2 - S}{2}. \quad (16)$$

Отсюда

$$A_1 = V_1 - d = \frac{V_1 - V_2 + S}{2} \quad (V_1 - V_2 + S = 0), \quad (17)$$

$$A_2 = V_2 - d = \frac{V_2 - V_1 + S}{2} \quad (V_2 - V_1 + S \geq 0). \quad (18)$$

Рис. 19.1. а) Объект 1.
б) Объект 2.

Если значение A_1 или A_2 в одном из уравнений (17) или (18) получается отрицательным, то решение состоит в изменении «отрицательной» цены на нуль и предложении всей располагаемой суммы денег S за другой объект.

Эта стратегия также оптимальна для B . Если и A и B придерживаются оптимальной стратегии, предлагаемые ими цены будут одинаковыми и результаты торгов определяются жребием. Ожидаемый доход как A , так и B равен, как следует из уравнения (16), d .

Если B уклонится при торгах от этой стратегии, например предложит $V_1 - d + \delta$ за объект 1 и $V_2 - d - \delta$ за объект 2, A получит объект 2 и B получит объект 1. Но доходы их будут различными.

$$\text{Доход } B = V_1 - (V_1 - d + \delta) = d - \delta, \quad (19)$$

$$\text{Доход } A = V_2 - (V_2 - d) = d. \quad (20)$$

Доход A превышает доход B на величину δ . Таким образом, используя оптимальные стратегии A_1 и A_2 , A получает доход, по крайней мере равный d , и препятствует получению B дохода, превышающего d .

Предположим, что в нашем примере A и B располагают каждый 100 долларами. Тогда, согласно выражению (17), A должен предложить за первый объект

$$A_1 = \frac{\$75 - \$100 + \$100}{2} = \$37,50$$

и за второй объект

$$A_2 = \frac{\$100 - \$25 + \$100}{2} = \$62,50.$$

Доход A в каждом из этих случаев (выиграй он торги) был бы равен 37,50 доллара (т. е. $\$75 - \$37,50$ и $\$100 - \$62,50$). Эта же стратегия является наилучшей и для B . Тогда, если это известно обоим, приобретение объектов будет зависеть от жребия и является делом случая.

Если начальные суммы различны и число объектов возрастает, то задача становится более сложной, но аналитические методы все еще позволяют найти наилучшую стратегию.

ЗАКРЫТЫЕ ТОРГИ: ЧИСЛО ЛИЦ ВЕЛИКО И ТОЧНО НЕ ИЗВЕСТНО

Исследуем теперь один пример, более близкий к действительности. Рассмотрим случай, когда правительственное учреждение приглашает большое число компаний из одной отрасли промышленности принять участие в торгах за контракты. Каждая компания, заинтересованная в получении контракта, должна принять участие в закрытых торгах, и компания, предложившая низшую цену, получает контракт. В этой ситуации число конкурентов, которые в действительности будут участвовать в торгах за каждый контракт, может быть неизвестным. Мы обсудим сначала пример торгов за один контракт и распространим затем полученные результаты на случай, когда торги ведутся за несколько контрактов одновременно.

Чтобы принять некоторую стратегию, компания должна совершенно точно определить цель ее участия в торгах. В рассматриваемом примере имеются много возможностей для выбора цели и принятие различных целей приведет к различным стратегиям. Перечислим ряд возможных целей.

1. Наиболее вероятной целью является максимизация общего ожидаемого дохода.

2. Следующей целью может быть возврат по крайней мере некоторой части капиталовложений.

3. Возможной целью является минимизация ожидаемых потерь.

4. Целью, которую обычно обнаруживают в настоящих состязательных ситуациях, является минимизация дохода конкурентов. Конкурент, заработавший много денег, становится более опасным конкурентом и при длительном соревновании повредит компании.

5. Может быть важным получить контракт даже в убыток, чтобы поддерживать определенный уровень производства.

Эти и другие цели, так же как и различные комбинации целей, можно обнаружить в задачах торгов. Каждая задача будет иметь различное решение в зависимости от целей, преследуемых компанией, участвующей в торгах. В рассматриваемых далее примерах мы будем предполагать, что единственной целью компании является максимизация полного ожидаемого дохода. Это, несомненно, одна из наиболее часто встречаемых целей и одна из простейших для рассмотрения в моделях торгов исследуемого типа.

Общая модель. Пусть C — оценка стоимости выполнения контракта. Истинная стоимость, определяемая после выполнения всех работ, будет, конечно, отличаться от оценочной стоимости. Поэтому важно определить величину смещения оценки и ее колебания. Это можно сделать, изучив данные за прошлое об оценках и истинной стоимости контрактов, что позволит получить распределение величины, характеризующей, какую часть от оценки составляет истинная стоимость (см. рис. 19.2).

Пусть $h(S) dS$ — вероятность того, что частное от деления истинной стоимости на ее оценку лежит между S и $S + dS$. Пусть x — предложенная

цена выполнения контракта. Тогда, если контракт получен по цене x , окончательный доход будет равен $x - SC$.

Обозначим теперь через $P(x)$ вероятность того, что предложение цены x будет минимальным и контракт будет получен. Тогда, если предложение цены x принято, средний доход $E(x)$, на который можно надеяться, равен

$$E(x) = \int_0^{\infty} P(x)(x - SC)h(S) dS. \quad (21)$$

Значение x , при котором это среднее значение дохода достигает максимума, должно быть предложено в качестве цены. Так как $P(x)$ не зависит от S и

$$\int_0^{\infty} h(S) dS = 1,$$

формула (21) принимает вид

$$E(x) = P(x)(x - C'), \quad (22)$$

где

$$C' = C \int Sh(S) dS$$

— несмещенная оценка стоимости выполнения контракта. В общем случае $E(x)$ имеет вид, сходный с кривой, изображенной на рис. 19.3. Предложение цены, приносящее максимальный средний доход, легко находится как абсцисса точки максимума кривой среднего дохода.

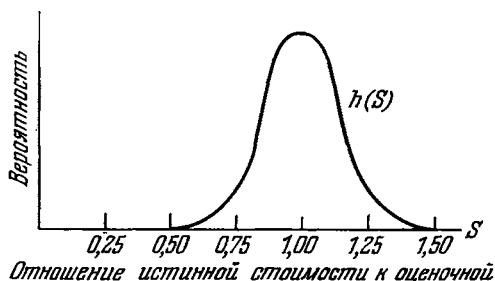


Рис. 19.2. Надежность оценки стоимости.

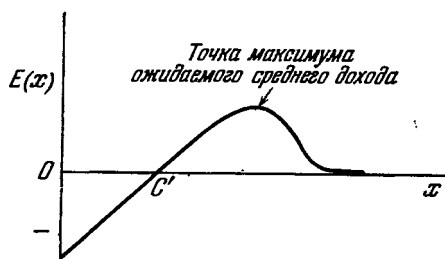


Рис. 19.3. Ожидаемый средний доход как функция предлагаемых цен.

Если кривая среднего дохода получена, то действительно легко определить цену, приносящую максимальный доход. Но основная трудность состоит в определении $P(x)$, вероятности получения контракта, как функции предложенной его цены.

Вероятность получения контракта. Один из возможных путей определения вероятности получения контракта при заданной предложенной цене состоит в изучении данных о прошлых торгах. Предполагается, что результаты прошедших торгов за контракты всегда объявляются. Они позволяют изучить поведение возможных конкурентов.

Допустим, мы изучаем конкурента A . Для каждого контракта, относительно которого A торговался и для которого наша компания производила оценку стоимости, мы образуем отношение цены, предложенной A , к нашей оценке. Если в прошлом было достаточно контрактов, за которые A торго-

вался, можно построить кривую, характеризующую поведение A относительно наших оценок стоимости контракта. Такие характеристики можно построить для всех возможных конкурентов. Несколько примеров таких кривых приведено на рис. 19.4.

Теперь, если нам точно известно, какие конкуренты будут участвовать вместе с нами в торгах, относительно легко вычислить вероятность получения контракта при заданном предложении цены.

Предположим, что каждый конкурент будет вести себя на предстоящих торгах так же, как он вел себя в прошлом, что является наилучшим предположением при отсутствии дополнительной информации. Вероятность оказаться ниже (предложить дешевле) конкурента A при предложении цены x равна площади области вправо от точки x/C под кривой распределения, характеризующей поведение на торгах A . Аналогично вероятность оказаться ниже B равна площади области вправо от точки x/C на кривой распределения B . И вероятность, что предложение x окажется самой дешевой ценой выполнения

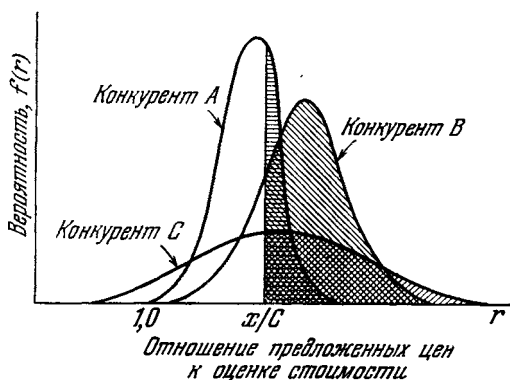


Рис. 19.4. Кривые поведения конкурентов.

контракта, когда конкуренты известны, равна просто произведению вероятностей поражения всех известных конкурентов.

Если точно не известно, сколько конкурентов будет участвовать в торгах, задача становится существенно более трудной. В этом случае необходимо ввести понятие среднего участника торгов. Характеристику поведения среднего конкурента можно получить, используя все предыдущие частные x/C для построения частотной кривой (см. рис. 19.5).

Рис. 19.5. Кривая поведения среднего конкурента.

Пусть $f(r)$ — плотность вероятности распределения отношения цен, предложенных средним конкурентом, к нашей оценке стоимости контракта. Тогда вероятность того, что предложение цены x будет меньше предложения среднего конкурента, равна $\int_{x/C}^{\infty} f(r) dr$. И вероятность предложить меньше k средних конкурентов равна тогда $\left[\int_{x/C}^{\infty} f(r) dr \right]^k$.

Предположим теперь, что мы можем определить вероятность того, что в торгах будут участвовать k конкурентов. Тогда, если эта вероятность равна $g(k)$, вероятность $P(x)$ того, что предложение цены x окажется наименьшим из всех предложенных цен, равна

$$P(x) = \sum_{k=0}^{\infty} \left\{ g(k) \left[\int_{x/C}^{\infty} f(r) dr \right]^k \right\}. \quad (23)$$

Кривая $f(r)$ может быть получена путем подбора параметров некоторой известной кривой в соответствии с располагаемыми данными.

Плотность γ -распределения часто дает хорошее совпадение с данными такого типа. Она имеет вид

$$f(r) = \frac{a^{(b+1)}}{b!} r^b e^{-ar}, \quad (24)$$

где a и b — постоянные параметры, которые получают путем сравнения частотной кривой с кривой плотности γ -распределения. Разумно также предположить, что число конкурентов распределено по закону Пуассона. Отсюда, если λ — оценка числа конкурентов, получаем

$$g(k) = \frac{\lambda^k e^{-\lambda}}{k!}. \quad (25)$$

Существуют критерии, позволяющие определить, согласуются ли оба принятых распределения с имеющимися данными.

Если оценка λ произведена корректным образом, то мы найдем

$$P(x) = e^{-\lambda} \sum_{k=0}^{\infty} \left\{ \frac{1}{k!} \left[\lambda \left(\int_{x/C}^{\infty} \frac{a^{b+1}}{b!} r^b e^{-ar} dr \right) \right]^k \right\} = \quad (26)$$

$$= e^{-\lambda} \exp \left[\lambda \left(\int_{x/C}^{\infty} \frac{a^{b+1}}{b!} r^b e^{-ar} dr \right) \right] = \quad (27)$$

$$= \exp \left\{ -\lambda \left[1 - \sum_{i=0}^b \frac{1}{i!} \left(\frac{ax}{C} \right)^i e^{-\frac{ax}{C}} \right] \right\}. \quad (28)$$

Сумма в скобках является значением функции распределения Пуассона, легко получаемой из таблиц.

Определение оптимального предложения цен. Математическое ожидание (среднее значение) дохода

$$E.P. = (x - C') P(x) = (x - C') \exp \left\{ -\lambda \left[1 - \sum_{i=0}^b \frac{1}{i!} \left(\frac{ax}{C} \right)^i e^{-\frac{ax}{C}} \right] \right\}, \quad (29)$$

где C' определяется согласно уравнению (22). Теперь можно начертить кривую зависимости средней прибыли от предлагаемых цен и легко определить оптимальное значение x (рис. 19.3).

Это решение можно получить и аналитическим путем, но получаемые формулы малоприспособны для использования.

В описанном методе определения оптимального предложения цен существенными параметрами являются оценка стоимости выполнения контракта C и оценки среднего числа ожидаемых конкурентов λ . Способы получения оценки стоимости выше уже обсуждались. Возможны несколько различных подходов к получению оценки числа конкурентов.

Во многих случаях компания располагает сведениями о намерениях конкурентов. В сочетании с опытом руководителей компании эти данные позволяют получить правильную оценку возможного числа конкурентов.

Величина и значимость контракта также могут влиять на число участников торгов. Эту связь можно исследовать, строя по прошлым данным линию регрессии числа участников торгов на оценки стоимости выполнения контрактов.

Если она значима, то оценка стоимости контракта позволяет, используя уравнение регрессии, получить оценку числа конкурентов.

Например, пусть зависимость между числом участников торгов и оценками стоимости выполнения контрактов имеет вид, представленный на рис. 19.6. Регрессия, с учетом имеющихся опытных данных, была выбрана линейной, и проведенная на чертеже прямая представляется вполне удовлетворительной. Ее мы впоследствии используем для получения оценки числа конкурентов, исходя из известной оценки стоимости выполнения контракта.

Моделирование торгов за несколько контрактов. Особенно широкий интерес представляет случай, когда государственная администрация объявляет торги за несколько контрактов одновременно. Каждая компания должна решить, за какие контракты она намерена торговаться и каковы будут ее предложения цен по каждому из торгуемых контрактов. Задача еще более усложняется при наличии для компаний-участников ряда ограничений. Это могут быть самоограничения, устанавливаемые каждой компанией для себя, или они могут накладываться администрацией, объявляющей торги.

Вот несколько примеров возможных ограничений.

1. Компания не может участвовать в торгах за контракты, общая оценочная стоимость выполнения которых превосходит заданную величину.
2. Компания не может получить более установленного числа контрактов.
3. Компании должно быть гарантировано получение по крайней мере некоторого числа контрактов.
4. Общая сумма предложения цен компании не должна превосходить заданной величины.

Ясно, что наличие одного из этих ограничений или некоторой их комбинации приводит к различной тактике предложения цен. Методы решения задачи будут связаны поэтому с природой существующих ограничений.

Чтобы показать общую идею решения подобных задач, мы предположим наличие ограничения 4 и решим задачу в случае торгов за n контрактов одновременно. Ограничение подобного типа может появиться различными путями. Например, администрация, предложившая торги, не желает, чтобы какая-нибудь компания получила слишком большой объем работ, и поэтому устанавливает верхнюю границу L для общей суммы предложений каждой компании. Или, может быть, ввиду финансовых затруднений или нехватки производственных мощностей, такое ограничение устанавливается руководством компании. Таким образом, если x_i — предложенная компанией цена за i -й контракт, то мы ищем максимум общего ожидаемого среднего дохода при дополнительном условии

$$x_1 + x_2 + \dots + x_i + \dots + x_n \leq L. \quad (30)$$

Для решения этой задачи мы поступим относительно каждого контракта так же, как и в предыдущем параграфе: построим кривые средних прибылей. Графическая иллюстрация для трех контрактов показана на рис. 19.7.

Построим для каждого контракта таблицу из трех столбцов (см. табл. 19.1). Числа в первом столбце, где указываются предложения цен,

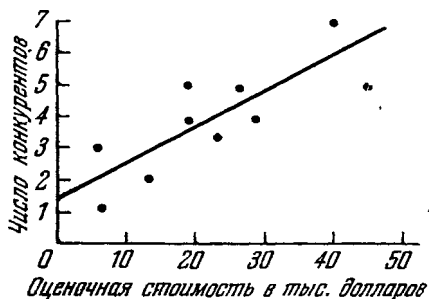


Рис. 19.6. Число конкурентов как функция оценочной стоимости.

Таблица 19.1

Определение оптимального предложения цен при наличии ограничений

Первый контракт			Второй контракт			Третий контракт		
Предложения цен, долл	Ожидаемый средний доход, долл	Изменение среднего дохода, долл	Предложения цен, долл	Ожидаемый средний доход, долл	Изменение среднего дохода, долл	Предложения цен, долл	Ожидаемый средний доход, долл	Изменение среднего дохода, долл
100 000	0		100 000			100 000		
105 000	3 000	+3000	105 000			105 000		
110 000	5 500	+2500	110 000			110 000		
115 000	7 500	+2000	115 000			115 000		
120 000	9 000	+1500	120 000	0		120 000		
125 000	9 900	+900	125 000	2000	+2000	125 000		
130 000	10 200	+300	130 000	3700	+1700	130 000		
135 000	9 900	-300	135 000	5100	+1400	135 000		
140 000	9 000	-900	140 000	6100	+1000	140 000		
145 000	7 500	-1500	145 000	6700	+600	145 000		
150 000	5 500	-2000	150 000	6800	+100	150 000	0	
155 000	3 100	-2400	155 000	6400	-400	155 000	3 000	+3000
160 000	300	-2800	160 000	5600	-800	160 000	5 500	+2500
165 000	0	-300	165 000	4400	-1200	165 000	7 500	+2000
170 000	0		170 000	2900	-1500	170 000	9 000	+1500
175 000	0		175 000	1100	-1800	175 000	10 000	+1000
180 000	0		180 000	0	-1100	180 000	10 400	+400
185 000	0		185 000	0		185 000	10 200	-200
190 000	0		190 000	0		190 000	9 400	-800
195 000	0		195 000	0		195 000	8 000	-1400
200 000	0		200 000	0		200 000	6 000	-2000
205 000	0		205 000	0		205 000	3 500	-2500
210 000	0		210 000	0		210 000	500	-3000
215 000	0		215 000	0		215 000	0	-500

должны возрастать, переходя даже разумные пределы, начиная от оценочной стоимости выполнения контракта до значений, когда ожидаемый средний доход становится равным нулю.



Рис. 19.7. Кривые ожидаемых средних доходов.

Во втором столбце таблицы помещают ожидаемый средний доход от контракта при данном предложении цены за его выполнение. Третий столбец указывает на изменения в среднем доходе, связанные с увеличением предлагаемой цены. Полученная таблица будет похожа на таблицу 19.1.

Рассмотрим частную ситуацию, описанную в таблице 19.1, где мы заинтересованы в получении трех контрактов. Предположим, что верхний предел общего предложения цен за три контракта установлен в 440 000 долларов.

Первый этап состоит в определении оптимальных предложений цен за каждый контракт. В рассматриваемом случае они равны 130 000 долларов за первый контракт, 150 000 долларов за второй контракт и 180 000 долларов за третий контракт. Общая сумма оптимальных предложений равна 460 000 долларов. Следовательно, необходимо сократить сумму предлагае-

мых цен на 20 000 долларов. Просмотрим третий столбец таблицы, где указаны изменения ожидаемого дохода, для определения, какие цены должны быть уменьшены, чтобы уложиться в заданную сумму. Легко видеть, что первое уменьшение следует сделать на втором контракте, так как это сократит средний доход только на 100 долларов, в то время как уменьшение на 5000 долларов цен первого и второго контрактов приводит к сокращению ожидаемого дохода соответственно на 300 и 400 долларов.

Продолжая аналогичным образом, мы в итоге получим совокупность предложений цен, удовлетворяющую заданным ограничениям. Оптимальные цены для таблицы 19.1 равны 125 000 долларов для первого контракта, 140 000 долларов для второго контракта и 175 000 долларов для третьего контракта.

В общем случае рассмотренная здесь задача является задачей нелинейного программирования, так как средний доход каждого контракта есть нелинейная функция предложения цен. Предложенный метод решения отличается относительной простотой и позволяет решить задачу довольно быстро, даже когда анализируются торги за большое число контрактов.

В предыдущих разделах книги (части IV—VIII) рассматривались модели, построенные для различных задач исследования операций. В данном разделе рассматриваются: (1) проверка адекватности модели и полученного с помощью модели решения, (2) разработка процедуры корректировки решения при его использовании и (3) использование полученного решения (т. е. проведение его в жизнь). Такое расположение материала (сначала построение модели, а затем испытание, корректировка и использование) не означает, что исследователи должны придерживаться такого порядка в своей работе. Очень часто построение модели и ее испытание проводятся одновременно. Различные соображения об использовании решения могут повлиять на вид модели, а следовательно, и на полученное решение.

Даже если решение правильно и уже опробовано, тем не менее оно остается оптимальным только до тех пор, пока система не претерпевает значительных изменений. Однако среди систем, изучаемых в исследовании операций, очень мало достаточно стабильных систем. Поэтому необходима система контроля для обнаружения значительных отклонений в системе и введение соответствующих корректировок в решение.

И наконец, решение надо провести в жизнь. Эта стадия является очень ответственной, от нее во многом зависит успех всего исследования. Поэтому надо уделить особое внимание работам по внедрению и контролю.

Глава 20

ДАННЫЕ ДЛЯ ИСПЫТАНИЯ МОДЕЛИ

ВВЕДЕНИЕ

В этой главе рассматриваются задачи, связанные со сбором данных для испытания моделей, рассмотренных в предыдущих частях книги. Эта тема настолько обширна, что было бы нелепо пытаться дать обзор всех относящихся к ней работ в одной главе, так что мы приведем лишь ссылки на соответствующую литературу.

Данные, необходимые для испытания модели, в значительной степени определяются видом испытаний, которые намерен провести исследователь. При испытании модели ставится вопрос: почему модель перестает правильно отображать реальный процесс и, следовательно, перестает быть полезной? Это сам по себе сложный вопрос, ответ на который зависит от понимания цели испытания. С точки зрения целей, рассматриваемых в данной главе, наиболее важными можно считать четыре недостатка, приводящих к нарушению адекватности отображения моделью реального процесса.

1. В модели отображается зависимость эффективности системы (зависимой переменной) от одной или нескольких (независимых) переменных, которые в действительности не влияют на эффективность системы. Таким образом, модель становится неадекватной, когда в нее включены переменные, отсутствующие в реальной системе. Например, известные модели, в которых среднее число заказов, необходимых для поступления определенной суммы, дано на счет фирмы, является функцией этой суммы. Но по крайней мере в одном случае установлено, что эти две величины не зависели друг от друга.

2. Модель может оказаться неверной, если в ней не учитываются переменные, которые в действительности влияют на эффективность системы. Например, модель может не учитывать зависимости общей стоимости производства от затрат, связанных с колебаниями количества рабочей силы, хотя для исследуемого случая эти затраты являются весьма существенными.

3. Модель может неправильно отображать действительную зависимость, которая существует между эффективностью (E) и одной или несколькими независимыми переменными (x_i, y_i). Например, согласно модели общая стоимость промышленной операции возрастает линейно с увеличением задержки в отгрузке товаров. Стоимость выражается через nc_d , где n — число дней задержки и c_d — штраф за один день задержки. Однако на самом деле общая сумма штрафа за задержку поставки может быть равной $(c_d)^n$.

4. И наконец, даже если все перечисленные выше условия не нарушены, но значения входящих в модель параметров оценены неправильно, модель может давать неверные результаты.

Фактически условие 4 является обобщением условий 1 и 2, так как включение в модель или исключение из нее различных переменных зависит от значений входящих в нее параметров.

При испытании модели можно начать с анализа ее в целом, т. е. с определения той точности, которую обеспечивает перспективное или ретроспективное предсказание на модели эффективности системы. Если же при этом выясняется, что модель не соответствует реальной системе, то нужно продолжить испытания с целью определения, какой из четырех указанных « пороков » присущ данной модели. Конечно, можно было бы сначала испытывать модель по каждому из рассмотренных недостатков, а потом уже испытывать ее в целом. Обычно на практике порядок испытания зависит от того

насколько сильна уверенность в правильности модели. Если считать, что модель достаточно надежна, то начинают сразу испытывать ее в целом. Если же нет достаточной уверенности в адекватности модели, то начинают испытывать (проверять) те аспекты модели, в которых больше всего сомневаются. Если модель испытывается в целом, то проверить ее правильность можно путем сравнения предсказанных моделью результатов с результатами реального процесса.

Рассмотрим теперь, какие данные необходимы для испытаний и каким образом их получать.

«НЕПРЕДУСМОТРЕННЫЕ» ДАННЫЕ И РЕЗУЛЬТАТЫ

Следует отметить важную особенность всех испытаний, которую нельзя предусмотреть заранее: исследователь сталкивается со многими неожиданными явлениями, которые нельзя предвидеть даже при самом тщательном планировании. Иногда при испытании оказывается, что модель, которая на первый взгляд правильно описывает процесс, не отображает очень важной стороны процесса. Например, в модели производства и хранения, правильно отображающей затраты на хранение и подготовку производства, может быть пропущена очень важная статья расходов, связанная с колебаниями количества рабочей силы. Или в модели предполагается, что поставки сырья могут быть произведены в любое время, хотя на самом деле поставщик предполагает осуществлять поставки дважды в месяц.

Такие примеры свидетельствуют о плохой связи между заказчиком и исследовательской группой. Однако это лишь указывает на недостатки исследования, но не дает ответа, как их исправить. В данной книге приводится ряд примеров успешного решения вопроса правильной связи между исследовательской группой и заказчиком.

На первый взгляд некоторые доказательства правильности модели часто кажутся настолько простыми, что как будто не требуется разработки специальных процедур сбора соответствующих данных. Однако история науки изобилует очень большим числом примеров, когда для обоснования простых утверждений приходится строить очень сложные процедуры сбора фактических данных. Опыт подсказывает, что каждое эмпирическое доказательство является сложным и необходимо разрабатывать специальные методы сбора экспериментальных данных, если только не удастся обосновать теоретическое положение, не прибегая к сложным опытам. Простоту доказательств нельзя признавать, руководствуясь интуицией.

Процесс сбора доказательств в общих чертах состоит из следующих элементов: наблюдение значений переменных (включая измерение), выборка (включая планирование эксперимента), преобразование данных, использование данных для испытания, анализ полученных результатов, а иногда и разработка нового процесса сбора доказательств.

НАУЧНЫЕ НАБЛЮДЕНИЯ

Научное наблюдение заключается в нахождении наилучших возможных (но не обязательно физически реализуемых) условий и процедур, с помощью которых можно установить значения переменных. Под «наилучшими возможными» здесь понимаются наилучшие из тех, которые *мы можем* себе представить, а не наилучшие в абсолютном смысле.

На первый взгляд этот этап кажется совершенно ненужным. Зачем разрабатывать процедуры, которые мы не сможем осуществить? На самом

деле такие идеальные (или оптимальные) условия и процедуры наблюдения не являются продуктом праздного воображения; всегда очень важно знать, насколько достоверны результаты, которые мы в конечном счете получаем. Кроме того, что еще более важно, идеальные условия и процедуры являются своего рода эталоном, с помощью которого мы можем оценить реально достижимые условия и операции наблюдения, выявить их недостатки и внести необходимые коррективы в полученные данные. Использование идеализированных операционных (т. е. научных) наблюдений * для корректировки полученных данных принято во всех науках.

Рассмотрим известный физический пример: определение ускорения свободно падающего тела. Идеализированные условия для измерения такого ускорения (наряду с другими условиями) предполагают, например, создание абсолютного вакуума, в котором падает данное тело. На самом деле физик никогда не может создать абсолютного вакуума, но он может провести эксперимент таким образом, чтобы оценить, как падало бы тело, если бы оно находилось в абсолютном вакууме. Он определяет, как воздействуют на ускорение колебания атмосферного давления, используя математические зависимости, чтобы связать колебания атмосферного давления с изменениями ускорения. Затем путем экстраполяции определяется, что произошло бы в абсолютном вакууме, и таким образом выводится идеальное ускорение свободно падающего тела.

Двумя главными типами количественных переменных являются счетные и метрические. Для оценки счетных переменных требуется их подсчет, в то время как для оценки метрических переменных требуется их измерение.

Научные наблюдения счетных переменных. Часто ошибочно считают, что счет является очень простым (если не простейшим) способом определения количества. Если бы кому-нибудь понадобилось сосчитать несколько раз, сколько букв «е» встречается на этой странице, то он бы получил различные результаты при каждом подсчете. Работник склада или сотрудник, проводящий перепись, могут подтвердить, насколько трудна эта операция. Считать легко, но трудно считать правильно.

Существует два типа ошибок при счете: превышение и преуменьшение истинного количества. Большее количество можно получить, если сосчитать один и тот же предмет несколько раз или сосчитать предметы, которые вообще не должны входить в счет. Меньшее количество получается, если пропустить те предметы, которые надо сосчитать. Кроме того, эти ошибки могут произойти вследствие ошибок при сопоставлении предметов с числами натурального ряда (например, большее количество получается, если пропустить одно или несколько чисел, а меньшее, если одно и то же число повторить дважды).

Процесс счета состоит из двух стадий: (1) определения предметов, которые надо сосчитать, и (2) сопоставления количеству предметов последовательных чисел натурального ряда. Ошибки при счете возникают на любой из стадий.

Желательно разработать наилучшую возможную процедуру счета, даже если мы не сможем на практике воспользоваться этой процедурой. Такая процедура состоит в определении стандартных условий и стандартных операций, с помощью которых можно провести подсчет идеальным образом. Кроме того, надо дать точное определение элементов, которые необходимо подсчитывать. После разработки такой процедуры можно

* Подробно о научном определении см. Акоф [1, стр. 305] и Хемпел [20].

оценить различные практически реализуемые процедуры счета и выбрать наилучшую из них. Кроме того, с помощью идеальной процедуры можно оценить ошибки, которые могут встречаться в выбранной нами реальной процедуре.

Научное наблюдение свойств. Разработка идеализированной процедуры для измерения свойств зависит прежде всего от самого свойства. В процедуру счета, например, необходимо включать предметы, обладающие определенным свойством, и исключать остальные. Однако в данном случае нас интересует, в какой степени предмет обладает определенным свойством. Другими словами, мы переходим от решения типа «да — нет» к решению типа «сколько».

Научное наблюдение свойств означает определение следующих характеристик идеализированного процесса измерения.

1. Определение предметов, явлений или классов предметов или явлений, которые следует подвергнуть наблюдению.
2. Определение условий, в которых должны проводиться наблюдения.
3. Определение изменений в условиях, которые, быть может, необходимо провести в процессе наблюдения.
4. Определение операций, которые необходимо выполнить, а также приборов и единиц измерения, которыми должен пользоваться наблюдатель.
5. Определение отсчетов (замеров), которые необходимо получить.
6. Схема анализа полученных данных.

В формальном описании процесса измерения* даются логические и математические операции, которые следует выполнить с полученными данными, чтобы оценить переменные. В научном определении наблюдения указывается, как в идеальном случае мы собирали бы нужные данные.

В операционном определении процесса сбора данных показывается, как в действительности мы намерены собирать данные. При испытании модели ошибки могут возникать на каждой из этих трех стадий планирования процесса испытания модели. Несмотря на то, что потребители результатов операционных исследований обычно более склонны сомневаться в данных, используемых при испытании модели, чем в самой модели, многие ученые все же имеют тенденцию исключать вопросы сбора данных из проведения исследований. Профессионалы же более склонны сомневаться в самой модели. Но если нужно «продать» и практически использовать результаты исследования, то операционист должен быть готов к тому, чтобы обосновать экспериментальные данные, так же как он обычно может обосновать свою модель.

Обосновать данные — значит не только показать их соответствие наблюдаемым фактам, но и продемонстрировать их представительность с точки зрения используемой процедуры выборки.

ВЫБОРКА

Оценка переменных состоит в измерении свойств отдельных элементов, или в подсчете числа элементов группы, или в измерении свойств этой группы. Определение свойств отдельного элемента означает определение тех условий, в которых необходимо проводить наблюдения. Если эти условия выполняются и можно провести наблюдение без ошибок,

* Более подробно о теории измерения см. [3, 4, 31, 32, 37, 38 и 39].

то требуется всего одно наблюдение. Если же нужные условия не соблюдаются, то наблюдения будут ошибочны и ошибки эти можно оценить только после двух или более опытов. Сколько проделать опытов и каких именно — все это вопросы, относящиеся к теории выборки. Поскольку на практике трудно выдержать идеальные условия наблюдения, нужно разработать методику проведения эксперимента, чтобы с помощью различных методов, например дисперсионного и корреляционного анализа, оценить величины отклонений от истинных значений и объяснить их причины. С другой стороны, можно провести наблюдения над определенным подмножеством элементов (выборкой) и сделать выводы, относящиеся ко всем изучаемым элементам данного множества.

Если мы оцениваем свойство множества элементов или производим их подсчет, то часто практически невозможно провести наблюдения над каждым элементом (например, над каждой гайкой и болтом, изготовляемыми на предприятии). Поэтому также необходимо получить выборку из всего множества.

В общем случае цель получения выборки очень проста и понятна. Просто мы хотим получить нужную информацию с минимальными затратами и за минимальное время. Как и другие задачи этой книги, задачу определения наилучшей выборки можно рассматривать как задачу уравнивания экономических факторов.

Если мы потратим много денег и времени на сбор данных, то получим более точные сведения, но за эту точность мы будем платить деньгами и временем. Каково же правильное решение? Некоторые научные организации сначала назначают комиссию для тщательного исследования вопроса, а затем уже делают выводы. Некоторые деловые люди, наоборот, склонны делать быстрые и решительные выводы. Между двумя этими крайностями находится оптимальное решение, уравнивающее риск от ошибочного решения и расходы по сбору данных.

Прежде чем перейти к обсуждению теории выборки, приведем некоторые примеры. Когда женщина покупает кукурузу, она может выбрать несколько початков, чтобы выяснить, свежая ли кукуруза. Она выбирает початки, которые кажутся ей наиболее типичными, из всего «множества», не пользуясь никакими операционными правилами, а просто руководствуясь здравым смыслом. Выборка на основе здравого смысла является наиболее распространенной, поскольку она самая дешевая, простая и быстрая. Мы постоянно применяем такой принцип выборки: выбирая место для еды, галстук или туфли. Но хотя эту выборку просто применять, ее очень трудно оценить. Обычно мы не знаем, как будет действовать человек, руководствуясь здравым смыслом, при выборе предметов, и поэтому отсутствует адекватный метод определения характеристик произведенной таким способом выборки. Некоторые стремятся выбирать предметы, которые обладают, по их мнению, наиболее типичными «усредненными» свойствами из всей данной совокупности, считая, что таким способом можно получить наиболее «представительные выборки». Например, на железной дороге выбирают «типичную неделю» в октябре, как наиболее «представительную» для транспортных условий. Другая компания выбирает «типичную область» для изучения запросов потребителя. Такой тип выборки, также основанный на здравом смысле, называют «целенаправленным». В результате этой процедуры можно получить очень плохую выборку в смысле достоверности информации обо всем множестве, особенно если нас интересует определение средних значений и дисперсий различных характеристик всех элементов данного множества. В общем случае ни из той, ни из другой выборки мы не

можем определить вероятность выбора элемента из данного множества, а именно это и является основным требованием для определения надежности оценок, которые мы получаем с помощью данных из выборки.

Выборка представляет собой группу предметов из определенного множества. Определение «множества» предметов, явлений, условий, из которого нужно сделать выборку, дается в определении той переменной, которая оценивается. Это множество содержит все возможные данные определенного типа, которые можно получить.

Оценка выборок и оценки по выборкам. Определяя процедуру выборки, мы должны решить, каким образом будем ее проводить, а также определить метод получения оценок для множества на основе выборки. При этом следует учитывать два показателя: (1) вероятность того, что полученная с помощью выборки оценка будет отличаться от истинного значения интересующего нас параметра всего множества предметов на величину, большую некоторой величины x ; (2) стоимость процедуры выборки. Обычно вероятность отклонения от истинного значения с увеличением объема выборки уменьшается, но стоимость при этом увеличивается.

Вероятность таких отклонений отчасти определяется разбросом значений оценок, полученных при обработке результатов выборки. Этот разброс, как правило, измеряется дисперсией оценки, а сама дисперсия представляет собой квадрат среднего квадратического отклонения полученных по выборке оценок от истинных значений измеряемой величины. Среднее квадратическое отклонение и размах являются другими мерами разброса.

Другой важной характеристикой процесса выборки (которая связана с вероятностью отклонения оценки от истинного значения) является «смещение». *Смещением* называется отклонение математического ожидания значения оценки, полученной из выборки, от истинного значения. Ожидаемое значение в свою очередь является средней оценкой, полученной при многократном повторении процесса выборки. Название «смещение» для этого свойства оценок является очень неудачным. Возможно, что определенный метод оценки дает некоторое смещение, однако этот метод может быть очень хорошим с точки зрения стоимости и общей ошибки. Например, относительные оценки являются смещенными, но в некоторых случаях они имеют гораздо меньшую дисперсию, чем несмещенные линейные оценки, так что вероятность отклонения на заданную величину от истинного значения оказывается меньше при использовании относительных оценок.

Если известны смещение и дисперсия процессов выборки и оценки, то обычно можно вычислить вероятность получения определенных отклонений, если, конечно, процессы выборки и оценки проводятся правильно. Это условие является очень существенным. Во многих случаях на практике на него не обращают внимания, и в результате процесс выборки — оценки оказывается неоптимальным. Очень часто ошибки, обусловленные неправильным проведением процесса выборки, и ошибки вычисления оценок увеличиваются с увеличением объема выборки n . Таким образом, обобщенная модель выборки состоит из следующих составляющих:

Стоимость выборки = (1) стоимость смещения и дисперсии +
(2) + стоимость исследования n элементов + (3) потери из-за ошибок
в проведении процесса выборки.

Если потери (3) возрастают с увеличением n , то может оказаться, что даже 100%-ная выборка дает очень большие ошибки, так что такой процесс выборки оказывается очень плохим.

Типы выборочных планов. При неограниченной случайной выборке каждая выборка может быть получена с равной вероятностью. Ограниченная случайная выборка означает неравную вероятность получения каждой выборки. Но всегда, когда процесс случайной выборки организован правильно, известна вероятность получения любой выборки.

В данной книге нет необходимости подробно освещать все методы выборки. Единственное, что можно сделать,— кратко описать некоторые процессы выборки, которые широко используются в операционных исследованиях. Однако следует подчеркнуть, что это изложение никак не претендует на широту охвата методов и теории процесса выборки *. Эти методы должны быть на вооружении не только у каждого операциониста, но и у каждого ученого вообще.

Все схемы процессов выборки основаны на очень простых практических соображениях.

1. Элементы множества могут распадаться на отличные друг от друга по значимости группы (например, по расположению, по денежным суммам в накладных и т. п.). В этом случае есть смысл производить выборку из каждой группы, так как при этом уменьшается дисперсия оценок и, что более важно, можно изменять объем выборки, необходимый по каждой группе. Накладные с более крупными суммами более важны, чем накладные с меньшими суммами; поэтому нам нужно брать большие выборки из более важных групп.

2. Элементы множества часто распадаются на классы (например, груз, указанный в накладной; жильцы дома, квартала или города; товары на складе и т. д.). Если мы хотим исследовать некоторый элемент определенного класса, то мы можем с таким же успехом исследовать остальные элементы этого класса. Следовательно, теперь выборка производится не по исходным элементам множества, а по классам. При этом может увеличиваться дисперсия оценок, но зато значительно сокращаются затраты на получение самой выборки, т. е. возникает обычная экономическая задача уравнивания различных факторов.

3. Не обязательно заранее составлять полный выборочный план. Можно определять следующие шаги, основываясь на информации, полученной из предыдущей выборки.

Приведем общую классификацию основных типов выборочных планов.

I. Фиксированная выборка: план выборки фиксирован и не подвергается изменениям в отношении правил отбора данных.

A. Неограниченная случайная выборка: случайная выборка ** производится из всей генеральной совокупности и может быть двух видов: (1) простая случайная выборка: каждому элементу множества присваивается определенный номер, и с помощью таблицы случайных чисел производится выборка; (2) систематическая случайная выборка: в упорядоченном множестве выбирается случайным образом начальный элемент, и затем выбираются последующие элементы на фиксированном расстоянии от начального и каждого следующего.

Б. Ограниченная случайная выборка: множество разбивается на подгруппы (которые в свою очередь могут разбиваться на подгруппы и т. д.), и в качестве выборки берется несколько подгрупп и/или берется случайная выборка из некоторых или всех подгрупп.

* Обзор по современной теории выборки см. в [1, 7, 11, 19, 35 и 50].

** Таблицы случайных чисел можно взять из [22, 26, 40 и 41]. Подробности получения таких чисел можно найти в [25].

1. Многоступенчатая случайная выборка: производится случайная выборка из подгрупп, которые в свою очередь выбраны (а) с равной вероятностью, или (б) с вероятностью, пропорциональной относительному размеру подгруппы, или по некоторому другому критерию.

2. Послойная случайная выборка: случайная выборка производится из каждой подгруппы множества. Объем выборки из подгрупп может быть следующим: (а) не зависит от размера подгруппы (т. е. выборки равного объема); (б) пропорционален относительному размеру подгруппы или (в) пропорционален относительному размеру подгруппы и дисперсии элементов внутри подгруппы (оптимальное распределение).

3. Выборка подгрупп: производится случайная выборка подгрупп, все элементы которых включаются в общую выборку.

4. Послойная выборка подгрупп: комбинация Б2 и Б3, включающая более двух стадий выборки.

II. Последовательная выборка *: берется небольшая случайная выборка и анализируется; на основе этого анализа принимается решение, продолжать ли процесс выборки, и если да, то каким образом. Эти выборки могут состоять из:

А. Групп, как в двойной или составной выборке, или

Б. Отдельных элементов, взятых по одному за раз.

Существует самостоятельный раздел теории выборки, который называется «планированием эксперимента» и занимается разработкой планов проведения эксперимента над переменными исследуемой модели. Вместо того, чтобы фиксировать значения всех переменных, кроме одной, исследуемой, можно построить систему сбора данных при «оптимальных» значениях некоторых переменных модели. Этот метод выборки предполагает, что можно свободно менять переменные реального процесса или по крайней мере переменные реалистической модели. Так, компания, сдающая напрокат своим клиентам бытовые приборы, должна обладать возможностью менять их свойства определенным образом, чтобы испытывать модель для получения оптимального варианта обслуживания при минимальных затратах. Информацию по вопросам планирования эксперимента можно найти в [8, 13, 17 и 27]. Особо следует упомянуть о методах экспериментального поиска, разработанных Боксом [2], которые предназначены для отыскания экстремума функции многих переменных с помощью оптимального плана поиска. Очевидно, что эти методы необходимы операционистам, поскольку во многих моделях требуется находить оптимальную (минимальную или максимальную) точку некоторой функции.

Обработка данных. Наблюдения над выбранными элементами или над выборкой ситуаций дают исходные данные, с помощью которых оцениваются различные переменные и, следовательно, создаются условия для испытания модели в целом или части модели. Во многих случаях бывает необходимо сравнение полученных данных, подготовка их к печати, кодировка, перфорирование и т. д. Здесь мы не собираемся рассматривать обработку данных (эти вопросы можно найти в X главе книги Акофа [11]). Нас интересует только форма, в которую следует превратить данные, чтобы их можно было использовать для испытания.

В общем случае такой формой является или оценка значения параметра, или выведенная «статистика» (функция результатов наблюдения), которая

* О последовательной выборке подробнее см. [5, 34, 43 и 45].

показывает зависимость между двумя или несколькими переменными. Например, при испытании модели оптимального размера партии необходимо сначала оценить переменные, определяющие затраты, чтобы вычислить общие затраты, «предсказываемые» моделью. После определения прогнозов на модели они сравниваются с экспериментальными данными, чтобы определить «статистику», которую можно использовать для оценки адекватной модели.

Рассмотрим сначала задачу оценки. Для упрощения изложения выборка и оценка рассматриваются отдельно, хотя при планировании исследования их необходимо рассматривать совместно. Затраты на реализацию плана выборки включают стоимость анализа данных, которую нельзя определить, не зная процедуры оценки. Процедура оценки, «наилучшая» для одной выборки, может оказаться плохой для другой. Следовательно, выражение «наилучшая оценка» относится только к определенному типу выборки. «Наилучшей» оценкой для данного типа выборки считается оценка, которая минимизирует ожидаемые стоимости ошибок (переоценки и недооценки). Однако в настоящее время мы не можем получать таких оценок. Большинство оценок, которыми мы пользуемся, выбираются в предположении, что стоимости ошибок пере- и недооценки (приблизительно) пропорциональны квадрату величины ошибки. Это предположение во многих случаях оправдывается, т. е. в четыре раза опасней ошибиться на три единицы, чем на одну, в девять раз опасней ошибиться на две единицы, чем на одну, и т. д. Одним из «преимуществ» такого предположения является сравнительная математическая простота минимизации квадрата ошибок (принцип «наименьших квадратов» *).

Кроме того, минимизируя квадраты, можно одновременно минимизировать другие функции ошибок. Тем не менее следует принять во внимание спорность такого предположения **.

В исследовании операций очень важно определить дисперсию оценок. Нельзя полностью определить ценность модели, если мы не знаем, насколько постоянны оценки ее эффективности. Больше того, если неизвестна надежность оценки эффективности модели, то ее нельзя использовать. Необходимо знать, будет ли действительно повышаться эффективность, как показывает оценка, или, наоборот, уменьшаться. (Одним из серьезных недостатков линейного программирования является трудность оценки надежности оптимального значения.) Определить надежность оценки эффективности можно двумя способами.

1. Если имеется модель вида $E = f(x, y)$, известны средние квадратические отклонения оценок x и y , то для некоторых функций можно аналитическим путем определить среднее квадратическое отклонение оценки для E .

2. Сравнивая теоретически рассчитанные оценки эффективности с действительной эффективностью, можно стандартными методами определить среднеквадратическое отклонение для оценок эффективности E .

* О методе наименьших квадратов, а также о других оценках характеристик множества, применяемых при различных планах процессов выборки, см. [12, главы 8 и 9].

** Р. А. Фишер [16] предложил оценку, обеспечивающую «максимум информации». Такая оценка должна максимизировать вероятность совпадения наблюдаемых значений с истинными (принцип «наибольшего правдоподобия»). Однако на практике очень часто эти оценки мало отличаются от оценок, полученных методом наименьших квадратов, а также от третьего типа оценок, называемых «минимаксными оценками» (минимизируется максимальный риск ошибки).

Вид представления данных и используемые оценки зависят от того, какие сведения о модели нас интересуют (см. стр. 439 этой главы). Статистические испытания, определяющие, насколько важны или незначительны те или иные переменные, описаны в [12, 16, 18, 21, 23, 28, 29, 33, 48 и 49]. Эти испытания дают возможность ответить на следующие вопросы:

(а) Следует или не следует включать в модель определенную переменную.

(б) Является ли аналитическая функция линейной или имеет другой вид.

(в) Является ли закон распределения нормальным или нет.

(г) Не пропущена ли в модели какая-то переменная, которая должна быть включена в модель.

В следующей главе рассмотрены некоторые методологические проблемы, связанные с такими испытаниями.

ПРОВЕРКА ДОКАЗАТЕЛЬСТВА — ИСПЫТАНИЕ РЕШЕНИЯ

Мы рассмотрели, какие стороны модели можно проверить путем изучения ее свойств. В заключение следует упомянуть об испытании модели в целом путем исследования результатов работы системы (ретроспективные испытания). В 21-й главе будут описаны испытания модели в течение повторяющегося процесса принятия решений.

Обеспечивает ли решение лучшую работу? Полученное операционистом решение задачи должно быть лучше обычного практического решения или должно содержать доказательство, что лучшее решение невозможно. Утверждение решения руководителем в основном определяется практической эффективностью этого решения, а не заложенной в нем логикой или правильностью соответствующих характеристик модели.

При ретроспективном испытании нужно восстановить прошлые результаты. Это означает, что нужно определить значения неуправляемых независимых переменных (например, спроса и затрат на подготовку производства в производственном планировании). Значения этих переменных редко можно достаточно просто определить, а если их и можно вообще выявить, то только косвенным путем — из документов. Обработка таких документов приводит к трезвым скептическим выводам, поскольку все знают, что числа редко означают то, что они должны были бы означать. Поэтому во многих случаях необходимо восстановить процедуру, с помощью которой были получены эти данные, т. е. фактически восстановить операционную модель (обычно неявно выраженную), с помощью которой были получены эти данные. Лишь после этого указанные данные можно сравнивать с нормативными и вносить необходимые коррективы. Например, отдел сбыта дает данные по заказам прошлых лет, а после проверки может оказаться, что это данные по отгруженным товарам. Это объясняется не только неравномерным распределением отгрузки товаров во времени. От некоторых заказов могли отказаться из-за невозможности быстро осуществить отгрузку товаров. Поэтому может потребоваться серьезная и длительная работа, чтобы получить данные по заказам на основе имеющихся данных по отгрузке.

Во многих случаях имеющиеся данные относятся к такому короткому промежутку времени, что есть смысл использовать все эти данные без всякой выборки во времени. В этих случаях обычно исследователи обнаруживают, что руководители не считают короткий промежуток времени достаточно

представительным для определения нормальных или типичных показателей. И действительно, легко убедиться, что нет ничего более нетипичного, чем типичный год.

Если имеется большой объем данных о прошлой работе, важно провести правильную послойную выборку с учетом разделения данных по характеристикам, используемым руководством для суждения о том, насколько типичным был исследуемый год. Например, руководители отличают хороший год от плохого в зависимости от обеспечения предприятия сырьем. Испытание решения будет тем убедительнее, чем более полно будут представлены различные периоды работы.

Само испытание обычно состоит из сравнения действительной эффективности с рассчитанной эффективностью, определяемой предложенным решением. «Прямого» сравнения обычно недостаточно. Например, при сравнении производственных затрат в примере главы 2 следует учитывать возможные ошибки в определении значений переменных модели. Это можно выполнить несколькими методами.

1. Предположим, что модель выражена в форме $E = f(x, y)$. Если известны дисперсии оценок x и y , то во многих случаях можно рассчитать дисперсию оценки E . Если известна дисперсия оценки E , то можно определить, является ли разница случайной величиной или нет.

2. При испытании модели в целом можно эмпирически получить значение дисперсии оценки E . Затем эту величину можно использовать, как указано в пункте 1.

3. Если же нельзя получить опытным путем значение дисперсии оценки E (как в примере главы 2), то можно попробовать поставить следующий вопрос: при каких величинах ошибок переменных преимущества предложенного решения сводятся на нет? Например, в задаче, приведенной в главе 2, такая ошибка в оценке обоих стоимостных показателей (фиксированной и переменной стоимости на партию) выражается в занижении действительного значения на 10%. В данной задаче есть основания полагать, что ошибки в оценках затрат не превысили предельного значения. Это можно доказать следующим образом. Если бы показатели действительно были занижены до предельных значений, то компания не получила бы никакой прибыли, используя предложенное решение. Однако в действительности была получена определенная прибыль. Следовательно, ошибки в определении оценок показателей затрат не превысили предельных значений.

При ретроспективных испытаниях следует осторожно подходить не только к временным выборкам (как было показано выше), но и к другим типам выборок. Например, нет надобности подсчитывать снижение производственных затрат по каждой из нескольких тысяч деталей, а следует взять выборку из этих деталей. Если учитывается много предприятий, потребителей и т. д., то по ним также следует провести выборку. Если план выборки составлен недостаточно правильно, то нельзя делать обоснованных выводов по выборке для всего множества.

Таким образом, при ретроспективном испытании решения необходимо не только сравнивать действительные и гипотетические оптимальные показатели по прошлым результатам, но также определять одним из описанных методов, насколько существенны наблюдавшиеся расхождения.

Иногда можно провести непосредственное испытание решения, собирая для этого необходимые данные. Однако, так же как и при ретроспективном анализе, возникают те же задачи корректировки данных и выборки (поскольку редко можно провести испытание всей системы). При таких испытаниях обычно можно более строго контролировать сбор данных, но

нельзя провести испытания в различных условиях. Таким образом, наблюдаемые отклонения от оптимума определяются более точно, но условия испытаний оказываются более ограниченными, чем при ретроспективных испытаниях.

Всегда, когда это возможно, желательно проводить испытания обоих видов. Это дает большую уверенность в правильности полученного решения, в то время как экстраполяция будущего по прошлому часто представляется слишком «неубедительной», что затрудняет «сбыт» полученного решения.

ЗАКЛЮЧЕНИЕ

В данной главе были рассмотрены процедуры проверки адекватности модели и полученного с ее помощью решения. Тот факт, что сначала рассматривалось построение моделей, а затем испытание моделей, совсем не означает, что исследователи должны следовать этой методике. Некоторые ученые полагают, что начинать исследования лучше всего с выявления имеющихся данных, поскольку они убеждены, что только таким путем можно «создавать теории», отображающие реальную действительность. Другие, напротив, считают, что нет смысла собирать данные до тех пор, пока нет модели, так как только после построения модели можно установить, какие данные в самом деле нужны.

Однако развитие науки показывает, что эти расхождения во взглядах объясняются только разной манерой работы отдельных ученых и не имеют отношения к вопросу о выборе правильного метода исследования. Одна из причин создания исследовательских коллективов в отличие от индивидуальной научной работы состоит в том, что при групповой работе контролируется излишнее увлечение одной части ученых экспериментом (работой над данными), а другой части — теорией (моделями). Исследовательский процесс представляет собой динамическое взаимодействие эксперимента (данных) и теории (моделей), причем не столь существенно, что предшествует чему. Но «динамические взаимодействия» трудно описывать в столь систематической книге, как эта. Гораздо проще показать их на каком-нибудь конкретном примере. Чтобы дать систематическое описание процедур исследования операций, необходимо было придерживаться определенной последовательности в изложении материала; поэтому сначала были даны сведения по несколько произвольно выбранным принципам построения модели, а затем — по испытанию моделей.

Многие утверждают, что «достаточно хорошо построенная» модель не нуждается в испытании. Или встречается несколько иное утверждение, что коль скоро ученый в явном виде принимает определенные допущения при построении модели, то с точки зрения самого исследования нет надобности в каких-либо дополнительных обоснованиях. Оба эти утверждения, по существу, направлены против сбора соответствующих доказательств адекватности модели, поскольку под «доказательством» мы понимаем информацию, которая потенциально может заставить нас изменить наше мнение. Если исследователь удовлетворен своей моделью, то для него доказательства могут быть излишними, хотя они могут быть очень важны для тех, кто хочет воспользоваться этой моделью.

Такие представления, характеризующие классическую научную философию рационализма, являются неприемлемыми, когда исследование охватывает широкий круг прикладных вопросов. Разделение теоретической и экспериментальной сторон исследования может быть необходимо время от времени. Но постоянное разделение этих аспектов научного исследования

может привести к трагическому разладу; ученый может строить такие модели, которые невозможно практически испытать. Как бы то ни было, поскольку исследование операций является исследованием организационных решений, которые проводятся в жизнь, необходимо уделять вопросам испытания моделей столь же серьезное внимание, как и вопросам их построения. Нет никакого смысла отыскивать теоретически оптимальные решения, которые оказываются неприменимыми на практике.

ЛИТЕРАТУРА

1. A c k o f f R. L., The Design of Social Research, University of Chicago Press, Chicago, 1953.
2. B o x G. E. P., The Exploration and Exploitation of Response Surfaces, Technical Report No. 4, Philadelphia Ordnance District, Dept. of the Army with Institute of Statistics, North Carolina State College, Raleigh, 1954.
3. C a m p b e l l N. R., An Account of the Principles of Measurements and Calculations, Longmans, Green & Co., New York, 1928.
4. C h u r c h m a n C. W., A Materialist Theory of Measurement, in R. W. Sellars, V. J. McGill, and Marvin Farber (eds.), Philosophy for the Future, The Macmillan Co., New York, 1949.
5. C h u r c h m a n C. W., Statistical Manual: Methods of Making Experimental Inferences, Pittman-Dunn Laboratory, Frankford Arsenal, Philadelphia, 1951.
6. C h u r c h m a n C. W., Theory of Experimental Inference, The Macmillan Co., New York, 1948.
7. C o c h r a n W. G., Sampling Techniques, John Wiley & Sons, New York, 1953.
8. C o c h r a n W. G. and C o x G. M., Experimental Designs, John Wiley & Sons, New York, 1950.
9. C o o m b s C. H., R a i f f a H. and T h r a l l R. M., Some Views on Mathematical Models and Measurement Theory, in R. M. Thrall, C. H. Coombs, and R. L. Davis (eds.), Decision Processes, John Wiley & Sons, New York, 1954.
10. C r a m e r H a r a l d, Mathematical Methods of Statistics, Princeton University Press, Princeton, 1945.
11. D e m i n g W. E., Some Theory of Sampling, John Wiley & Sons, New York, 1950.
12. D i x o n W. J. and M a s s e y F. J., Jr., Introduction to Statistical Analysis, McGraw-Hill Book Co., New York, 1951.
13. F e d e r e r W. T., Experimental Design, The Macmillan Co., New York, 1955.
14. F e l l e r W., An Introduction to Probability Theory and Its Applications, John Wiley & Sons, New York, 1950.
15. F e r r i s C. D., G r u b b s F. E. and W e a v e r C. L., Operating Characteristics for the Common Tests of Significance, Ann. math. Statist., 17, 178—197 (1946).
16. F i s h e r R. A., Statistical Methods for Research Workers, Oliver & Boyd, London, 1948.
17. F i s h e r R. A., The Design of Experiments, Oliver & Boyd, London, 1949.
18. H a l d A., Statistical Theory with Engineering Applications, John Wiley & Sons, New York, 1952.
19. H a n s e n M. H., H u r w i t z W. N. and M a d o w W. G., Sampling Survey Methods and Theory, John Wiley & Sons, New York, 1953.
20. H e m p e l C. G., Fundamentals of Concept Formation in Empirical Science, in «International Encyclopedia of Unified Science» (Part III), University of Chicago Press, Chicago, 1952.
21. H o e l P. G., Introduction to Mathematical Statistics, John Wiley & Sons, New York, 2nd ed., 1954.
22. H o r t o n H. B., Random Decimal Digits, Interstate Commerce Commission, Washington, 1949.
23. J o h n s o n P. O., Statistical Methods in Research, Prentice-Hall, New York, 1949.
24. K e n d a l l M. G., Advanced Theory of Statistics, Chas. Griffin & Co., London, 2 vols., 1947.
25. K e n d a l l M. G. and S m i t h B. B., Randomness and Random Sampling of Numbers, J. R. statist. Soc., 101, 147—166 (1938).
26. K e n d a l l M. G. and S m i t h B. B., Tables of Random Sampling Numbers, in «Tracts for Computers», no. 24, Cambridge University Press, Cambridge, 1940.
27. M a n n H. B., Analysis and Design of Experiments, Dover Publications, New York, 1949.

28. Massey F. J., Jr., The Kolmogorov—Smirnov Test for Goodness of Fit, J. Amer. statist. Ass., 46, 68—78 (1951).
29. Mode E. B., The Elements of Statistics, Prentice-Hall, New York, 1941.
30. Mood A. M., Introduction to the Theory of Statistics, McGraw-Hill Book Co., New York, 1950.
31. Nagel Ernest, Measurement, Erkenntnis, 2, 313—333 (1931).
32. Nagel Ernest, On the Logic of Measurement, Thesis, Columbia University, New York, 1930.
33. Snedecor G. W., Statistical Methods, Iowa State College Press, Ames, 4th ed., 1946.
34. Statistical Research Group, Sequential Analysis of Statistical Data: Application, Columbia University Press, New York, 1946.
35. Stephan F. F., History of the Uses of Modern Sampling, J. Amer. statist. Ass., 43, 12—39 (1948).
36. Stephan F. F., Sampling, Amer. J. Sociology, 55, 371—375 (1950).
37. Stephan F. F., Mathematics, Measurement, and Psychophysics, in S. S. Stevens (ed.), Handbook of Experimental Psychology, John Wiley & Sons, New York, 1951.
38. Stevens S. S., On the Problem of Scales for the Measurement of Psychological Magnitudes, J. univ. Sci., 9, 94—99 (1939).
39. Stevens S. S., On the Theory of Scales of Measurement, Science, 103, 677—680 (1946).
40. The RAND Corporation, A Million Random Digits, The Free Press, Glencoe, 1955.
41. Tippett L. H. C., Tables of Random Sampling Numbers, in «Tracts for Computers», no. 15, Cambridge University Press, Cambridge, 1927.
42. Wald Abraham, Contributions to the Theory of Statistical Estimation and Testing Hypotheses, Ann. math. Statist., 10, 299—326 (1939).
43. Wald Abraham, Foundations of a General Theory of Sequential Decision Functions, Econometrica, 15, 279—313 (1947).
44. Wald Abraham, On the Principles of Statistical Inference, in «Notre Dame Mathematical Lectures», no. 1, Notre Dame University, Notre Dame, 1942.
45. Wald Abraham, Sequential Analysis, John Wiley & Sons, New York, 1947.
46. Wald Abraham, Statistical Decision Functions, Ann. Math., 20, 165—205 (1949).
47. Wald Abraham, Statistical Decision Functions which Minimize the Maximum Risk, Ann. Math., 46, 265—280 (1945).
48. Walker H. M., Elementary Statistical Methods, Henry Holt & Co., New York, 1943.
49. Wilks S. S., Elementary Statistical Analysis, Princeton University Press, Princeton, 1949.
50. Yates Frank, Sampling Methods for Censuses and Surveys, Chas. Griffin & Co., London, 1949.

Глава 21

КОРРЕКТИРОВКА И ПРИМЕНЕНИЕ РЕШЕНИЯ

ВВЕДЕНИЕ

Во многих, если не в большинстве операционных проектов рассматриваются многократно решаемые руководителями задачи. Например, в задаче управления запасами разрабатывается «правило принятия решения», которое можно применять неоднократно по мере необходимости и при изменении ситуации. Каждую неделю или месяц можно определять количество производимой продукции, а для этого необходимо применять полученное решение. Но сама изучаемая система редко остается неизменной. Изменяется структура системы. Изменяются соотношения между переменными системы (ее параметры), да и сами параметры не остаются неизменными.

Поэтому в предлагаемом решении необходимо изменять соответствующие соотношения и параметры по мере их изменения в самой системе. Со временем могут меняться затраты, распределение спроса по одному или всем показателям и соотношения между переменными. Следовательно, следует периодически переоценивать соотношения и параметры и проверять выполнимость различных допущений модели, на основании которой разрабатывается правило принятия решений. Таким образом, решение необходимо корректировать, иначе оно перестанет быть эффективным из-за изменений в системе.

После того как решение было испытано и была разработана процедура корректировки повторяющихся решений, необходимо применить решение на практике, т. е. для полного завершения проекта необходимо внедрение полученного решения. Целью исследования операций является улучшение операций, а не простая информация о положении дел. Поэтому необходимо разработать и провести в жизнь план внедрения.

Правда, четкая универсальная методология построения оптимальных процедур корректировки и внедрения еще не разработана. Однако достаточно хорошо известно, как разрабатывать процедуры, которые дают больший эффект на этапе внедрения, чем при отсутствии какого-либо плана или передаче этого этапа полностью в руки «заказчика».

Поскольку мы пока не умеем строить оптимальную процедуру корректировки решений, на этой заключительной стадии исследования открываются богатые возможности для проявления изобретательности, которая порождается прежде всего на основе глубоких знаний статистики, ее принципов и средств, а особенно методов статистического контроля качества. Знакомство с общими идеями и логикой статистического контроля качества*, а также различными методами и приемами этой дисциплины** играет существенную роль в разработке эффективных систем корректировки решения.

КОРРЕКТИРОВКА РЕШЕНИЯ (УПРАВЛЕНИЕ РЕШЕНИЕМ)

Эффективность решения операционной задачи может снизиться при изменении значений параметров системы или зависимостей между ними (см. стр. 439, гл. 20). Параметр, не игравший вначале существенной роли, может стать важным, и наоборот. На эффективность решения могут также влиять изменения величин и зависимостей параметров, которые не теряют своего значения для системы. Здесь важно напомнить, что если определенные параметры имеют значения, близкие к нулю, то соответствующие переменные системы уже не играют большой роли.

Рассмотрим несколько примеров. Во-первых, несущественные переменные могут стать важными. Например, в модели управления производством и запасами могут не учитываться затраты на хранение полуфабрикатов (как в примере главы 2) из-за малой длительности производственного цикла. Но при изменении номенклатуры выпускаемой продукции или объема производства производственный цикл может значительно измениться. А потому затраты на хранение полуфабрикатов могут стать значительными. И наоборот, сокращение производственного цикла в системе со значительными затратами на хранение полуфабрикатов может уменьшить значение этих переменных. В задаче, связанной с закупкой товаров, возможен вариант, когда все цены закупаемых товаров меняются достаточно плавно в ограниченном

* См. [3 и 24].

** См. [5, 6, 9, 13, 15, 22, 23 и 25].

диапазоне и, следовательно, можно не учитывать скачков цен. Если же изменяется объем закупок или политика цен поставщика, то скачки цен могут оказывать значительное влияние.

Во-вторых, значение параметра может существенно измениться. Например, вполне вероятно, что могут измениться спрос, скорости поставок, мощности наличного оборудования, затраты и т. д. Эти изменения должны найти отражение в решении, т. е. значения параметров нужно периодически переоценивать.

В-третьих, могут измениться функциональные зависимости в решении. Например, потери из-за нехватки товаров могут быть линейной функцией времени, когда спрос превышает предложение, или эти потери могут быть даже постоянными (не зависят от времени нехватки). Когда же предложение превышает спрос, то потери из-за нехватки могут расти со все возрастающей скоростью в зависимости от времени нехватки.

Не всякое изменение параметра или зависимости играет существенную роль.

В общем случае изменение существенно, если (1) корректировка решения в соответствии с этими изменениями дает повышение эффективности и (2) стоимость такой корректировки и проведение ее в жизнь не превышают повышения эффективности.

Таким образом, разработка системы «управления» решением состоит из трех этапов.

1. Составление перечня переменных, параметров и зависимостей, которые или включены в решение, или должны быть включены при определенных изменениях.

2. Разработка процедуры обнаружения значительных изменений каждого из включенных в перечень параметров и каждой зависимости.

3. Определение действий или корректировок, которые необходимо внести в решение в случае обнаружения значительных изменений.

Последние два этапа взаимосвязаны, потому что определение значимости изменения (этап 2) зависит от затрат на проведение корректировок, определенных на этапе 3.

Целесообразно рассматривать коррекцию параметров и коррекцию зависимостей отдельно, поскольку требования к соответствующим процедурам различны. Рассмотрим сначала коррекцию параметров.

Коррекция параметров.

Первый шаг в разработке процедуры коррекции (управления) состоит в составлении перечня всех переменных и зависимостей, изменения которых влияют на эффективность решения. Очевидно, что этот перечень должен включать все переменные и зависимости, которые есть в решении. Но далеко не столь очевидно, что в этот перечень должны входить также переменные, которые не включены в решение, но при определенных (вполне вероятных) условиях должны быть включены. В приведенном примере такими переменными являются затраты на хранение полуфабрикатов. Составление перечня должно быть основано на анализе системы, который предшествует построению модели. Однако требуется вновь рассмотреть каждую переменную, которая была отброшена при анализе, с тем, чтобы выяснить, не будет ли эта переменная существенной в определенных условиях и не изменится ли ее значение настолько, что она начнет оказывать заметное влияние на решение.

Обычно удобно «разбивать» некоторые параметры на их «наблюдаемые» составляющие. Например, затраты на подготовку производства можно раз-

бить на трудоемкость подготовки (человеко-часы) и стоимость человеко-часа, а затраты на создание запаса можно разбить на затраты хранения, капиталовложения, налоги и т. д. При таком разбиении обычно гораздо проще определить причину изменения, если оно произошло, а это является, как будет показано ниже, наиболее существенным элементом в «управлении» решением.

Включенные в перечень переменные рационально разделить на два класса.

1. Переменные, значения которых на интервале действия решения можно определить заранее, например: число модификаций производимых изделий, число рабочих дней за отчетный период и продажная цена изделия. Корректировка значений таких переменных осуществляется (а) путем обмена информацией между теми, кто знает их истинные значения, и теми, кому они нужны при принятии решения, или (б) путем обеспечения последних соответствующими справочными материалами (например, календарь при определении числа рабочих дней за отчетный период).

2. Переменные, значения которых нельзя определить заранее, например: число проданных изделий, число отработанных часов или количество автомашин, прибывающих в пункт назначения в единицу времени. Значения таких переменных нужно заранее оценить.

Обычно достаточно просто разработать процедуру коррекции переменных первого класса, хотя не всегда просто провести эту процедуру в жизнь. Для переменных же второго класса обычно довольно трудно разработать процедуру коррекции. Поэтому остановимся подробнее на коррекции (подстройке) переменных, значения которых необходимо оценить.

В этом случае задача сводится к тому, чтобы определить (на основе накопленных данных или оценок), следует ли изменить в будущем полученную в прошлом оценку или следует изменить саму процедуру определения оценки. Например, при принятии решения на рассматриваемом интервале можно использовать средние затраты на подготовку производства, вычисленные по результатам прошлой работы. При этом необходимо контролировать, произошли ли значительные изменения истинного значения этой средней величины, и если произошли, то следует определить, какое значение использовать на будущее. Если же оценивается «месячная потребность в деталях» и значение этой величины прогнозируется каждый месяц, то необходимо проверить предположения, связанные с получением этих оценок. Например, возможно предположение, что оценка является несмещенной, т. е. считается, что математическое ожидание отклонений предсказанной оценки от действительной величины равно нулю. Значение смещения следует проверить, чтобы определить, произошло ли его изменение, и если произошло, то какое, что в свою очередь может потребовать внесения изменений в предсказываемые оценки. Следует отметить, что в этом случае корректируются оценки, а не истинные значения самих переменных.

Для контроля любой переменной важно определить, изменилось ли ее истинное значение или одна (несколько) из характеристик оценки этой переменной. Такой контроль заключается в проверке гипотезы об отсутствии изменений в самой переменной или в ее оценке (характеристики оценки сами являются переменными).

Ошибки при обнаружении изменений. Ошибки при обнаружении изменений могут быть двух типов:

Тип I — утверждение, что изменение произошло, хотя на самом деле его не было.

Тип II — отрицание изменения, которое произошло.

Чтобы уяснить сущность коррекции значений переменных, важно понять эти два типа ошибок.

Первый тип ошибки характеризуется уровнем значимости α , при котором испытывается гипотеза H_0 (null hypothesis) об отсутствии изменения. Если, например, при испытании гипотезы используется уровень значимости 5%, то это означает, что предполагается наличие изменения, хотя в действительности в 5% всех случаев изменения не произошло. Одна из задач разработки процедуры коррекции заключается в определении уровня значимости (т. е. приемлемой ошибки I типа), при котором следует испытывать гипотезу об отсутствии изменения.

Ошибка II типа является более сложной, поскольку значение предположения об отсутствии изменения, которое в действительности произошло,

определяется величиной изменения. Следовательно, ошибка II типа выражается не одной величиной, а различными вероятностями, соответствующими каждой возможной величине изменения. Эти вероятности зависят от вида статистического испытания, проводимого для определения изменения. Вероятности, связанные с различными ошибками II типа, удобно показывать на так называемой рабочей характеристике (см. рис. 21.1). На рис. 21.1 показаны характеристики для одного вида статистических испытаний, для «двустороннего z-испытания». По горизонтальной оси отложены различные возможные изменения (в единицах среднего квадратического отклонения),

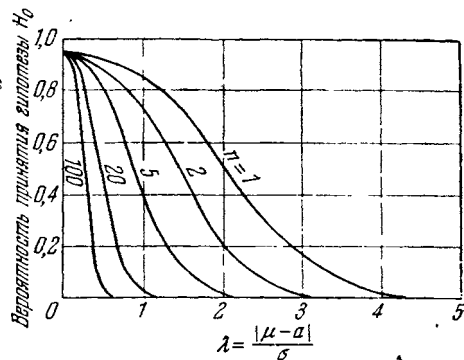


Рис. 21.1. Рабочие характеристики для нормального двустороннего z-испытания.

которые следует проверить на гипотезу об отсутствии изменения. При использовании кривой гипотезу следует представить в такой форме, при которой истинная средняя величина (например, спроса) равна a (т. е. $H_0/\mu = a$), где a представляет собой величину, используемую в решении. Значения λ , отложенные по горизонтальной оси, соответствуют различным величинам μ (в единицах среднего квадратического отклонения σ). Отклонения от a могут быть положительными и отрицательными.

Единица по оси λ означает, что истинное среднее значение μ находится на расстоянии σ от среднего значения a , используемого в решении. Например, предположим, что испытывается гипотеза о том, что истинное среднее значение спроса за месяц составляет 125 000 долларов. Тогда при значении среднего квадратического отклонения 11 200 долларов одна единица по горизонтальной оси означает, что истинное среднее значение спроса за определенный месяц заключено в диапазоне 113 800—136 200 долларов. Две единицы означают диапазон 102 600—147 400 долларов и т. д.

Кривые на рис. 21.1 представляют собой вероятности гипотез об отсутствии изменений H_0 для каждого возможного значения μ . Когда гипотеза H_0 справедлива (и, следовательно, $\mu = a$), то при уровне значимости 0,05 кривая проходит через точку с вероятностью 0,95 при $\lambda = 0$. Это означает, что с вероятностью 0,95 гипотеза H_0 принимается и с вероятностью $1 - 0,95 = 0,05$ она отвергается. При движении вправо переходим к рассмотрению возможных ошибок II типа, и вероятность начинает падать. Характер кривой зависит от объема выборки. Например, если рассматри-

ваются результаты только двух опытов *, то гипотеза H_0 принимается с вероятностью 0,72 при $\lambda = \sigma$; если же используются результаты 20 опытов, то вероятность падает до 0,04. Чем больше выборка, тем менее вероятно, что мы примем H_0 , когда произошло значительное изменение. Кривые на рис. 21.1 падают гораздо круче для больших выборок. Для маленьких выборок (1 или 2 опыта) вероятность принятия ложной гипотезы H_0 остается большой даже при значительном изменении в 2σ **.

Обнаружение значительных изменений и их корректировка.

В идеальной процедуре коррекции переменной должны быть приняты шесть взаимозависимых решений. (В некоторых случаях нет выбора между одним или несколькими решениями.) Перечислим эти решения.

1. Частота (т. е. период) контрольных замеров.
2. Число наблюдений за один контрольный замер, если возможно более одного наблюдения.
3. Способ выбора опытов для наблюдения (т. е. разработка плана выборки).
4. Процедура статистического испытания для определения изменения переменной.

5. Правило принятия решения, основанное на испытании.

6. Действия, которые необходимо предпринять, если испытание покажет, что величина переменной изменилась.

Кроме того, эти решения должны быть приняты таким образом, чтобы минимизировать сумму следующих затрат:

- 1) стоимости проведения наблюдений;
- 2) стоимости проведения испытания;
- 3) средней стоимости ошибки I типа (например, стоимости изменения величины переменной, когда изменение не оправдано);
- 4) средней стоимости ошибки II типа (например стоимости отказа от изменения, которое необходимо произвести).

В некоторых случаях следует минимизировать максимум затрат или некоторые другие стоимостные характеристики, а не среднюю стоимость. Однако в этой книге такие вопросы не исследуются, а рассматривается только одна стоимостная функция — математическое ожидание стоимости, поскольку она чаще других используется в качестве критерия для принятия решения при разработке процедуры коррекции ***.

К сожалению, в настоящее время не разработано процедуры принятия перечисленных шести решений при условии минимизации суммы указанных четырех статей затрат. Разработка оптимальной процедуры может проводиться обычным путем, т. е. можно построить модель, которая выражает математическое ожидание общих затрат как абстрактную (но не конкретную) функцию шести решений. Но некоторые выражения модели нельзя оценить. Например, в большинстве случаев нельзя определить средние затраты, связанные с ошибками I и II типов. Кроме того, неизвестны нужные рабочие характеристики или выражения, из которых они получены.

Можно сделать ряд замечаний по затронутым вопросам (подробности см. в [3, 21, 27, 28]). Теория вывода стала частью общей теории принятия решений, в которой значения величин в смысле желательных результатов

* Рабочие характеристики строятся в предположении, что проведенные наблюдения являются случайной выборкой из множества возможных опытов, причем распределение является нормальным и проводится определенное статистическое испытание, упомянутое выше.

** Подробнее о двух типах ошибок см. в [3, 26, 27 и 28].

*** Относительно критерия принятия решения при статистическом испытании см. [3, 26 и 27].

рассматриваются как основные факторы, определяющие поведение. Современная наука отходит от «интуитивного» обоснования принятого решения и все больше исходит из экономических предпосылок. И действительно, можно использовать «смещенный» метод испытания только из-за того, что он дает наиболее желательные общие результаты.

Кроме рабочих характеристик следует упомянуть о «стоимостных» кривых, отображающих стоимость принятия гипотезы в различных условиях. Оценку стоимости можно получить, только если каждой гипотезе поставить в соответствие определенную стратегию, поскольку словесная или символическая гипотеза является только удобным промежуточным звеном между доказательством (опытами) и действием.

Для иллюстрации рассмотрим две стоимостные кривые для двух различных стратегий. Первая стратегия состоит в том, чтобы принять партию материала, вторая — отказаться от приемки. Известно, что в партии есть определенный, но неизвестный процент p бракованных материалов. Таким образом, для первой стратегии стоимость принятия гипотезы практически равна нулю, например, до 1% брака, а затем начинает расти и при 3% брака возрастает довольно быстро из-за реакции покупателей. Стоимость второй стратегии означает потерю материала и зависит линейно от p . Следует отметить, что при первой стратегии существует зона нечувствительности от 0% до 1%, поэтому корректная постановка гипотезы будет следующей:

$$H_0/p \leq 0,01, \quad H_1/p > 0,01.$$

Предположим, есть два различных метода проверки партии. Какой лучше? Очевидно, что надо умножить вероятность принятия H_0 на стоимость принятия H_0 и прибавить вероятность принятия H_1 , умноженную на стоимость принятия H_1 для каждого p . Таким образом, получим функцию «риска» для каждого метода. Иными словами, риск равен стоимости для данного значения p , умноженной на вероятность принятия гипотезы.

Чтобы оценить методы испытания гипотез через функции риска, необходимо знать, как часто встречаются различные значения p . Например, возможно, что нет партий с $p > 0,03$. Тогда вид функции риска для $p > 0,03$ не влияет на оценку метода.

Таким образом, оценка методов испытания гипотез зависит от некоторых априорных значений различных параметров, подобных параметру p в приведенном примере. Это одна из основных проблем современной теории статистических решений. Одно из положений, основанных на классическом «принципе недостаточности оснований», заключается в том, что если нет никаких данных относительно p , то следует принять, что все значения p равновероятны. Это означает, что при оценке методов испытания гипотез следует рассматривать площадь под кривой риска как меру неэффективности метода (метод с наименьшим общим риском принимается лучшим). Другое положение состоит в том, что при отсутствии предварительной информации нужно рассматривать ситуацию как игровую. Метод экспериментатора при испытании рассматривается как стратегия в игре. Противником является «природа», стратегия которой состоит из некоторых определенных значений параметров. Если природа является сознательным противником, то она будет «выбирать» такие значения p , которые дадут максимальные средние потери для экспериментатора. Поэтому экспериментатор должен играть таким образом, чтобы минимизировать потери от стратегии природы, т. е. минимизировать максимальный риск.

Чтобы наглядно проиллюстрировать это положение, приведем матрицу игры. Предположим, что стратегиями экспериментатора являются два

метода — метод I и метод II. Предположим также, что p может принимать значения 1, 2, 3 и 4%. Это возможные стратегии природы. Таким образом, известна матрица игры:

Экспериментатор	Природа			
	$p=1\%$	$p=2\%$	$p=3\%$	$p=4\%$
Метод I	100	250	255	225
Метод II	50	175	325	175

Поставленные в матрице числа представляют собой «проигрыш» экспериментатора, который всегда положителен (это стоимости «штрафа», а не действительные общие потери). Экспериментатора устраивает метод II, поскольку он обеспечивает минимальные потери (50). Но если он действительно выберет этот метод, то природа (считаем ее мстительным противником) выберет $p = 3\%$, что вызовет значительные потери экспериментатора. Если же он выберет метод I, то природа по-прежнему выберет $p = 3\%$, и оба игрока будут придерживаться своих оптимальных стратегий (точка $p = 3\%$ (метод I) является седловой точкой). Следовательно, экспериментатор должен выбрать метод I.

С точки зрения методологии интересно, что для определения наилучшего метода необходимо сделать определенные предположения об исследуемых параметрах. Но если известно распределение вероятностей для p , как показано выше, то зачем испытывать гипотезы относительно p ? Чтобы ответить на этот вопрос, нужно понять «цикличность» всякого научного исследования, в том числе и процесса испытания гипотез. Делаются определенные предположения для оценки методов испытания гипотез, после испытания гипотез получаются определенные данные, на основе которых делаются новые предположения об оценке методов (например, испытания дают информацию о распределении действительного процента брака в партии). Такая цикличность исследования может завести в порочный круг, если не делать каждый раз соответствующие выводы.

Несмотря на указанные трудности возможны различные подходы к разработке системы «управления» решением, близкой к оптимальной. Каждый подход определяется конкретными условиями проведения исследования, так что нельзя указать единой процедуры, которая годилась бы для разработки системы, близкой к оптимальной, при любых условиях. Вместе с тем можно привести несколько общих соображений относительно требований к такой системе и проиллюстрировать их на конкретных примерах. Более подробно о методах коррекции решения см. [6, 9, 13, 15, 17, 18, 22, 23, 24].

1. Когда в результате испытания, проводимого с целью проверки решения, выявляется значительное изменение значения переменной, то это не означает, что нужно автоматически подставлять в решение новое значение. Нужно попытаться определить причину этого изменения. Если причина выяснена, то может оказаться, что она не останется в силе в течение продолжительного периода, и поэтому изменять значение переменной не следует. Например, математическое ожидание «спроса за месяц» может значительно измениться в один из месяцев, когда наблюдается «невероятно» высокий спрос. Причина может заключаться в том, что на следующий месяц

объявлено повышение цен и покупатели стараются запастись товарами перед повышением. В этом случае можно предсказать возврат к «нормальным» колебаниям спроса и не вносить никаких изменений в решение. Кроме того, обнаружение причины изменения может указывать на необходимость включения новой переменной в модель и в решение. Например, если в приведенном примере плановики заранее получают информацию о предстоящем повышении цен, то они могут соответственно изменить оценку спроса. Установленная причина изменений может отражать не временные, а достаточно устойчивые изменения системы. Например, введение автоматического оборудования может значительно повлиять на затраты по подготовке производства и на затраты прямого труда.

2. Когда не удастся найти какой-нибудь причины для объяснения происшедшего изменения, то следует заменить старое значение переменной новым. Однако в последующие периоды необходимо обратить особое внимание на эту величину, проверить, как она соответствует решению, и постараться вновь найти причину (или причины) изменения.

3. Частота проверки определенной величины (т. е. интервал контроля) может определяться временем применения правила принятия решения. Если, например, составляется производственный план на каждый месяц, то проверку надо выполнять также каждый месяц.

4. В некоторых случаях в течение всего периода контроля только один раз можно провести необходимые наблюдения. Например, месячные потребности обеспечения производства можно определить только раз в месяц; если при этом период контроля равен одному месяцу, то не возникает вопроса об объеме выборки.

5. Если же можно проводить наблюдения чаще, то при использовании процедуры статистического контроля предполагается, что наблюдения, проведенные на любом интервале времени, являются случайной выборкой из всех наблюдений, которые можно провести в течение всего периода контроля*. Важно проверить это предположение о случайности временного интервала. При этом мы можем также улучшить точность оценок, используемых в решении.

6. До сих пор речь шла о контроле отдельных величин. Однако на практике могут происходить одновременные изменения нескольких величин. При этом два изменения, каждое из которых само по себе незначительно, могут оказаться значительными, если они происходят одновременно. Но существует так много различных возможных комбинаций изменений, что операционисты не могут планировать испытаний для каждой комбинации. Это означает, что такие ситуации нужно анализировать по мере их возникновения. Ряд описанных методов контроля необходим, чтобы обнаружить, какие изменения произошли, но важность одновременных изменений обычно необходимо оценить, когда они возникают. Ответственность за такой контроль нужно поручить персоналу (группе исследования операций или другому подразделению), который хорошо понимает, как было получено решение. Это необходимо для того, чтобы понять важность одновременных изменений и найти средства их корректировки.

Прежде чем перейти к примерам, рассмотрим другой аспект системы коррекции решений.

* Возможно другое предположение, согласно которому наблюдения, проведенные в любой интервал времени, являются случайной выборкой из всех возможных наблюдений, которые можно провести в том же интервале времени. Если берется выборка наблюдений в течение периода контроля, то делаются оба предположения, но справедливость второго предположения проверяется процедурой случайной выборки.

Коррекция зависимостей.

Любое распределение вероятностей предполагает определенную зависимость между вероятностью события и значением переменной. Если же в модели и в решении есть функции распределения, например распределение спроса, то нужно контролировать параметры распределения (например, математическое ожидание и среднее квадратическое отклонение), а также форму закона распределения (например, нормальный закон или закон Пуассона).

Контроль параметров распределения выполняется так же, как для других величин, и потому не нуждается в дополнительном обсуждении. Нет никакой «стандартной» процедуры и для контроля закона распределения. Такой контроль можно выполнять путем периодической проверки по критерию «согласия» с помощью обычных статистических испытаний. Частота, с которой нужно проводить такие испытания, зависит от поступления соответствующих данных. Если эти данные наносить на график, то часто по характеру получающейся кривой можно определить, когда следует проводить испытания.

Например, в случае нормального распределения можно периодически наносить получаемые данные на вероятностную сетку, на которой построена также постулируемая кривая распределения. Исследование таких кривых может дать ключ к объяснению изменения параметров и формы закона распределения*.

Контроль других зависимостей, не имеющих формы распределения вероятностей, также включает контроль вида функции, определяющей соотношение между переменными, и контроль значений самих переменных. Предположим, в модели есть выражение nc , где n есть число изделий в производственной партии и c — зарплата, отнесенная к одному изделию. Выражение nc представляет собой выплату зарплаты рабочим на всю партию. В этом выражении устанавливается прямая зависимость между n и c . Контроль этого выражения состоит из двух этапов.

Во-первых, нужно периодически проверять соответствие линейного соотношения nc последним полученным данным по размеру зарплаты на партию.

Например, при увеличении размера партии может повыситься производительность труда, в связи с чем удельная зарплата падает по мере увеличения n . Эту закономерность можно усмотреть из графика общих затрат как функции от n .

Во-вторых, надо построить график отклонений реальной зарплаты на партию от теоретического графика nc в функции времени и проверить, имеют ли эти отклонения случайный или систематический характер. Если окажется, что полученный временной ряд не является случайным, то это означает, что изучаемая величина зависит от времени и, следовательно, время нужно ввести в рассматриваемое соотношение в качестве переменной. Необходимо определить причину этой временной зависимости (например, повышения опыта операторов), чтобы контролировать эту причинно-следственную связь.

Каждая задача исследования операций имеет свои отличительные особенности, в связи с чем постоянно возникают новые задачи контроля, для которых необходимо разрабатывать оригинальные процедуры контроля. На этой стадии исследования открываются широкие возможности для про-

* Подробности об испытаниях распределения на соответствие нормальному закону см. в [6 и 8].

явления творческих способностей операциониста. Это наглядно подтверждается нижеследующими примерами, в каждом из которых рассматривается разработка процедуры контроля, соответствующей особенностям конкретной задачи.

Примеры систем контроля (коррекции решения).

Пример 1 *. В проекте системы управления производством и запасами, разработанном для фирмы «Муллинс менюфэчуринг компани», прогноз месячного спроса продукции фирмы (кухонного оборудования) определялся

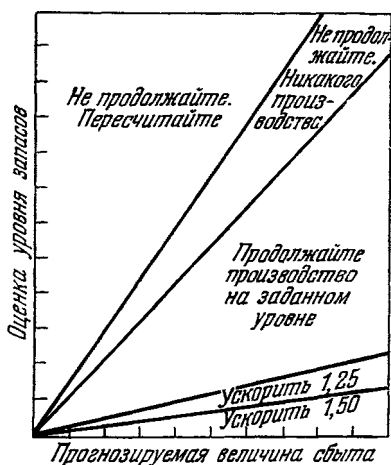


Рис. 21.2. Графики управления производством на данный месяц.

как треть среднего квартального сбыта по данным последних четырех кварталов. Хотя эта оценка была более достоверной, чем ряд других опробованных оценок, все же она была недостаточно надежной (большая дисперсия) и потому считалась сравнительно нестабильной. С помощью ретроспективных испытаний можно было показать, что если бы в решение входила эта оценка, то при составлении каждого календарного плана производства наблюдалась бы каждый раз нехватка товаров. Для этой величины был установлен следующий контроль.

Изучение данных по результатам нескольких прошлых лет показало, что нехватки товаров не бывает в те месяцы, когда отношение начальных запасов к объему сбыта в этом месяце больше 0,14. Другими словами, это отношение указывало на время опережения производства для пополнения запасов.

Поскольку производство на данный месяц планируется до того, как станет известной величина сбыта товаров за этот месяц, то в состав этого отношения должна входить сама оценка сбыта. Необходимо также оценить уровень начальных запасов на планируемый месяц. Таким образом, неудовлетворительная величина отношения указывает на плохое качество прогнозирования, так что следует провести необходимую корректировку оценок. Иными словами, если оценка отношения меньше 0,14, то это означает, что требуется выпустить большее количество изделий, чем рассчитано по уравнению «запуск — выпуск». Если же оценка отношения больше чем 0,14, то на планируемый месяц не требуется никакого выпуска данных изделий. В помощь плановику были построены простые графики (см. рис. 21.2).

На каждый месяц каждому изделию соответствует точка, которая представляет собой прогнозируемое отношение начальных запасов к объему сбыта. Эта точка может попасть в одну из пяти показанных областей. Область «Не продолжайте. Пересчитайте» означает, что требуется изменить метод прогнозирования. Изучение экономики изменения операций показало, что более подробных указаний, чем дает такой график, не требуется. Использование этого контрольного графика, а также полученного из модели процесса правила принятия решения относительно объема производимой продукции дало возможность значительно сократить запасы и в то же время фактически устранить нехватку товаров.

* Подробное описание этого проекта можно найти в [11, 12 и 16].

Пример 2 *. В другом исследовании, проведенном для отделения ламп компании «Дженерал электрик», была установлена возможность значительного сокращения числа размещаемых заказов, необходимых для выполнения годового плана сбыта продукции. Ретроспективные испытания показали, что такое сокращение в течение относительно короткого периода (свыше 2 лет) не оказывает влияния на объем сбыта. Однако для исследования влияния на более продолжительный период времени не было достаточных данных. Поэтому необходимо было разработать процедуру, чтобы определить, будет ли такое сокращение сказываться на объеме сбыта и когда.

Для этого была разработана процедура периодического проведения множественного регрессионного анализа зависимости между объемом текущего сбыта по балансу и изменением годового числа заказов за прошлые годы, полученным также по балансовому отчету. Например, в конце 1955 года был определен объем сбыта за 1955 год (по отчету) и сопоставлен с изменениями числа заказов за следующие годы: а) с 1952 по 1953, б) с 1953 по 1954 и в) с 1954 по 1955 год. В конце 1956 года добавляется еще один член: изменение заказов с 1955 по 1956 год. Такой расчет будет выполняться в течение 10 лет. Если обнаружится, что уменьшение объема сбыта за какой-нибудь год сильно зависит от сокращения заказов за какой-то год или комбинацию лет, то тогда можно изменить политику управления торговым оборотом.

Чтобы проводить эти периодические контрольные проверки наилучшим образом, понадобилось получать более точные сведения по торговым заказам, для чего были введены новые формы отчетности.

Пример 3 **. Заказчиком этого проекта является химическая компания, ежегодный оборот которой составляет приблизительно 55 миллионов долларов. Производство осуществляется на семи предприятиях, расположенных в разных частях страны. Штат рабочих и служащих компании составляет приблизительно 2500 человек. Компания производит большое количество разнообразных продуктов, требующих бактериологической и химической обработки.

Разработанный проект предусматривает управление производством и запасами трех продуктов, которые мы обозначим через A , B и C . Эти продукты выпускаются одним предприятием, и их производство начинается с одного и того же процесса ферментации. Различия между продуктами возникают на последующих стадиях производственного процесса, включающего приблизительно 50 этапов.

После тщательного изучения системы группа исследования операций разработала процедуру планирования производства. В этой процедуре необходимо было учесть интервал времени между окончанием планирования продукта (t_0) и окончанием его производства (t_1). Интервал (t_0 , t_1) был принят равным 10 дням. Процедура планирования заключалась в следующем.

1. В момент времени t_0 определяется количество каждого продукта на складе (S_{0A} , S_{0B} , S_{0C}).

2. Определяется наиболее высокий объем сбыта каждого продукта на один период за предшествующие девять десятидневных периодов (X_A , X_B , X_C). Затем определяются оценки объема сбыта в интервалах (t_0 , t_1) и (t_1 , t_2) как fX_A , fX_B и fX_C , где f есть коэффициент запаса, вычисленный по минимальной величине, которая в данной процедуре планирования

* Подробное описание этого проекта можно найти в [1 и 4].

** Подробное описание этого проекта см. в [2].

гарантирует от нехватки продукта. Анализ накопленных за прошлые годы данных показал, что величина f равна 2,0.

3. Тогда величина запаса S_1 , нужная в момент t_1 , определяется как fX_A , fX_B и fX_C .

4. Количество Q_A продукта A , которое нужно произвести за время (t_0, t_1) , определяется из выражения

$$Q_A = fX_A - (S_{0A} + a_{0A} - fX_A) = 2fX_A - S_{0A} - a_{0A},$$

где a_{0A} — количество продукта A , которое должно поступить на склад за время (t_0, t_1) . Q_B и Q_C определяется аналогично.

Предположим, например, что 25 единиц продукта находятся на складе в момент t_0 (т. е. $S_0 = 25$), 75 единиц были запланированы для производства на предшествующий период и будут получены за время (t_0, t_1) (т. е. $a_0 = 75$). Кроме того, максимальный объем сбыта за предшествующие девять десятидневных периодов равен 35 единицам (т. е. $X = 35$). Предположим, используется коэффициент запаса, равный 2 (т. е. $f = 2$). Тогда количество продукта, которое надо произвести (Q), определяется следующим образом:

$$Q = 2 \cdot 2 \cdot 35 - 25 - 75 = 40.$$

Для удобства была составлена таблица, показывающая выходы трех продуктов на различных стадиях ферментации в десятидневный период. Процесс изготовления носит такой характер, что, когда продукт A планируется, определенные количества продуктов B и C уже получаются. Используя полученную таблицу, можно определить ожидаемый выход каждого продукта, зная только расход материалов для каждого ферментатора. Был также разработан рабочий график для облегчения планирования на каждый десятидневный период.

Во время первого ретроспективного испытания этой простой процедуры планирования не делалось никакой попытки ограничить изменение числа ферментаторов, переходящих из одного периода планирования на следующий период. Результаты ретроспективного применения этой процедуры к периоду от января 1950 года до декабря 1952 года были следующими: не наблюдалось нехватки продуктов; средний уровень запаса продукта A уменьшился на 58%, продукта B — на 63%, продукта C — на 56% по сравнению с действительным средним уровнем запаса, полученным за трехлетний период. Эти сокращения сказались на уменьшении общего потребного числа ферментаторов. За трехлетний период было получено сокращение числа ферментаторов на 17%. Таким образом, введение этого метода дало бы возможность сократить и затраты, связанные с запасами, и издержки производства.

Проведенные исследования показали, что перенесенная на конечный продукт стоимость вложенных капиталов, стоимость хранения, наблюдения, износа, порчи и т. д. доходит до 2% от стоимости конечного продукта в расчете на месяц. На основе этой цифры была вычислена приблизительная стоимость запасов за трехлетний период. При ретроспективном испытании предложенной процедуры планирования средняя месячная стоимость запасов была уменьшена на 42%.

Полученные результаты были обнадеживающими. Но при этих ретроспективных испытаниях уровень производства был подвержен частым и значительным колебаниям. Это влекло частые изменения уровня рабочей силы, которые были неприемлемы для компании.

На предприятиях компании действует система старшинства, согласно которой любое изменение служебного положения одного работника на данном предприятии, производящем данный продукт, может повлечь за собой пять различных изменений положений работников на других предприятиях. По этой системе, если на определенном предприятии нужен какой-то работник, то его можно перевести с другой должности, на которую в свою очередь можно взять работника с должности более низкого уровня по системе старшинства. Поскольку в системе предусмотрено пять различных уровней старшинства, то требуется пять различных перемещений работников. Это связано с увеличением затрат на обучение и сокращением выпуска продукции. Поэтому процедура планирования производства должна учитывать стабилизацию рабочей силы, т. е. эта система должна обеспечивать минимальное число перемещений работников по каждому предприятию. Были получены данные относительно требуемого уровня рабочей силы для различных уровней производства.

Для обеспечения средней скорости ферментации, определяемой представленным производственным планом, были необходимы три бригады. При этом возникает вопрос: если бы в течение трех лет (1950—1952) работали три бригады и объем производства колебался бы от минимальной величины до максимальной, то каков был бы уровень запасов и насколько велик был бы дефицит продуктов на складе? Для ответа на этот вопрос был вновь использован тот же основной метод планирования с учетом ограничений по рабочей силе. При этом подсчете оказалось, что все заявки за трехлетний период были удовлетворены и уровень запасов сократился: для продукта *A* на 39%, *B* на 57%, *C* на 57%, что дало сокращение затрат по запасам примерно на 40%. Кроме того, количество потребных ферментаторов сократилось на 18%.

Предложенная процедура планирования была использована для расчета будущей работы предприятия в течение первых восьми месяцев 1953 года. В результате оказалось, что все запросы были удовлетворены и не было нехватки продуктов на складе.

Как уже отмечалось, процедура планирования испытывалась ретроспективно по результатам последних трех лет работы. Эти испытания оказались очень удачными, но при этом возникли два вопроса:

1. Почему (в каких условиях) эта процедура планирования эффективна?
2. Где гарантия того, что она будет сохранять свою эффективность? Или конкретнее, как и где можно внести изменения в процедуру, чтобы она продолжала действовать достаточно эффективно?

Процедура планирования разработана на основе двух главных характеристик сбыта.

1. В общем случае количество проданного продукта за один плановый период не превосходит удвоенной величины максимального количества продукта, проданного за предшествующие 9 плановых периодов.

2. Средние требования за десятидневный период могут быть удовлетворены при условии, что работают три бригады.

Чтобы гарантировать надежную процедуру планирования, необходимо проверять эти две характеристики, чтобы при их нарушении можно было обнаружить происшедшее отклонение. Кроме того, необходимо разработать методы изменения процедуры при нарушении характеристик. Сначала рассмотрим коэффициент планирования.

На рис. 21.3 приведен график коэффициента планирования для продукта *A* за истекший 41 месяц (коэффициент планирования рассматривается как отношение количества действительно проданного за месяц продукта

к максимальному месячному количеству продукта, проданного за предшествующие 3 месяца). Результаты по другим продуктам совершенно аналогичны.

Ниже показывается, что указанный коэффициент меняется в диапазоне от 0,5 до 1,5 со средним значением, равным 1. Исследования показали, что необходим коэффициент запаса, равный 2, при изменении коэффициента планирования в указанном диапазоне, чтобы избежать нехватки продукта. Например, при коэффициенте 1,75 наблюдается нехватка продукта на

складе *. Для контроля можно использовать рис. 21.3. До тех пор, пока отношение меняется в указанных пределах, можно использовать коэффициент запаса, равный 2.

Если же точка попадает за контрольный диапазон, то предлагается сначала попытаться найти причину этого явления. Например, отношение, приблизительно равное 2,5, объяснялось объявленным компанией на следующий месяц повышением цен. Для данного графика удалось объяснить каждую точку, выходящую за контрольный диа-

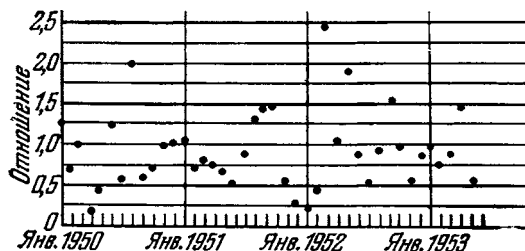


Рис. 21.3. Контрольный график коэффициента планирования для продукта А (отношение количества действительно проданного за месяц продукта к максимальному месячному количеству продукта, проданного за предшествующие 3 месяца).

зон. Каждая причина дала возможность выявить соответствующую переменную, пропущенную при построении модели. Если не учитывать трудностей внедрения, то следовало бы включить в модель все эти переменные. Однако в данном случае было сравнительно просто выявить отклонения и определить необходимые корректирующие воздействия. Поскольку было установлено, что процедура планирования гарантирует от нехватки продукта даже в периоды значительных отклонений, что подтвердилось в ряде случаев, то была введена коррекция — увеличение коэффициентов запаса f от 2,0 до 2,5, если две соседние точки выходят за контрольную зону и нет оснований полагать, что они вернуться в нее. Поэтому было предложено строить новый контрольный график в диапазоне от 1,0 до 2,0.

Были найдены причины для двух случаев отклонений точек ниже предельного уровня 0,5. Анализ показал, что при применении описанной процедуры планирования можно устранить возникновение слишком больших запасов. Для этого надо сделать коэффициент запаса равным 1,5 и передвинуть пределы на 0 и 1,0, если окажется, что три точки попадают ниже границы 0,5 и нет причин ожидать, что они вернуться в указанный диапазон.

Теперь рассмотрим вопрос, связанный с ограничением рабочей силы. Допустим, что O_h представляет собой максимальный месячный объем сбыта за предшествующие 3 месяца. Согласно процедуре планирования наличие продукта на складе в первый десятидневный период определяется как $\frac{2}{3}O_h$.

В конце каждого месяца можно рассчитать отношение среднего месячного запаса к общему месячному объему сбыта, причем средний месячный запас получается как сумма запасов в начале каждого планового периода, деленная на три. Рассчитанную величину отношения можно сравнить с верх-

* С помощью таких расчетов стало возможным показать компании, во сколько обходится гарантия от нехватки продукта в различных количествах.

ним и нижним пределами этого отношения, полученными из предположения, что для контроля остается коэффициент запаса, равный 2. Эти пределы получаются следующим образом.

В соответствии с используемым коэффициентом запаса 2 действительный объем сбыта за месяц не может быть больше $1,5O_h$ и не может быть меньше $0,5O_h$, т. е. в пересчете на десятидневный период получим соответствующие значения $1/2O_h$ и $1/6O_h$. Планирование рассчитано на десятидневный период, и плановики учитывают проданные продукты, когда становится известно о продаже. Исходя из пределов, определяемых коэффициентом запаса 2 (т. е. 0,5 и 1,5), вычислены значения верхнего и нижнего пределов отношения среднего месячного запаса к действительному месячному объему сбыта (2,33 и 0,63). Вычисления проводились следующим образом. Если O_h представляет собой максимальный объем сбыта за предшествующие 3 месяца, то, считая объем сбыта одинаковым за каждый десятидневный период, получим значение $1/3O_h$ в качестве максимального объема сбыта за предшествующие девять десятидневных периодов. Планирование предусматривает объем производства $2/3O_h$ на первый период нового месяца.

В таблице 21.1 приведен расчет указанного отношения для нижнего предела ожидаемого сбыта.

Таблица 21.1

Расчет верхнего предела (отношения)

10-дневный период	Запас (производство + «излишек»)	Ожидаемый сбыт
1	$2/3O_h$	$1/3(0,5O_h)$
2	$2/3O_h + 2/3O_h - 1/6O_h = 7/6O_h$	$1/3(0,5O_h)$
3	$2/3O_h + 7/6O_h - 1/6O_h = 5/3O_h$	$1/3(0,5O_h)$
Всего	$7/2O_h$	$1/2O_h$

Средний месячный запас $7/2O_h : 3 = 7/6O_h$,

Отношение среднего запаса

к общему месячному объему сбыта $= 7/6O_h : 1/2O_h = 2,33$.

Аналогично в таблице 21.2 приведен расчет отношения для верхнего предела ожидаемого сбыта.

Средний месячный запас $17O_h : (6 \times 3) = 17/18O_h$,

Отношение среднего запаса

к общему месячному объему сбыта $= 17/18O_h : 3/2O_h = 17/27 = 0,63$.

На рис. 21.4 показан график отношения среднего месячного запаса к общему месячному объему сбыта. Отмечены верхний и нижний пределы (2,33 и 0,63). Когда точка попадает выше верхнего предела, то это означает, что запасы могут стать слишком большими, и наоборот, когда точка попадает ниже нижнего предела, то это показывает, что запасы могут стать слишком малыми.

Соответствие этих данных теоретическим показывает, почему ретроспективные испытания, проведенные из расчета работы трех бригад, были

Т а б л и ц а 21.2

Расчет нижнего предела (отношения)

10-дневный период	Запас (производство + «излишек»)	Ожидаемый сбыт
1	$\frac{2}{3}O_h$	$\frac{1}{3}(1,5O_h)$
2	$\frac{2}{3}O_h + \frac{2}{3}O_h - \frac{1}{2}O_h = \frac{5}{6}O_h$	$\frac{1}{3}(1,5O_h)$
3 *	$O_h + \frac{5}{6}O_h - \frac{1}{2}O_h = \frac{4}{3}O_h$	$\frac{1}{3}(1,5O_h)$
Всего	$\frac{17}{6}O_h$	

* В начале второго периода (когда планируется производство продуктов, которые будут готовы в третьем периоде) известен сбыт за первый период. В данном случае он принимается равным $\frac{1}{3} \cdot 1,5O_h$. Тогда прогнозируемый спрос на третий период равен $2 \cdot \frac{1}{3} \cdot 1,5O_h = O_h$.

столь успешными. При ретроспективных испытаниях новой процедуры планирования ни одна точка не выпала за контрольные пределы, в то время как на практике наблюдаются отклонения от этих пределов (см. рис. 21.4).

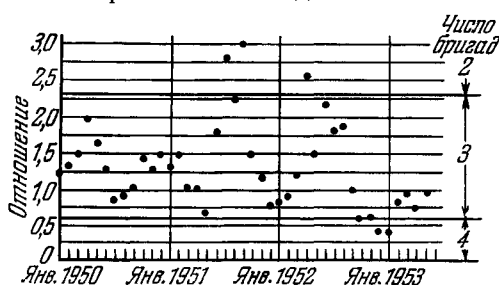


Рис. 21.4. Контрольный график для определения уровня рабочей силы (рассматривается отношение среднего месячного запаса к общему месячному объему сбыта).

Кроме того, анализ действительных запасов в точках, находящихся вне контрольных пределов, показал очень низкие или высокие уровни запасов или даже недостаток продуктов. Поскольку затраты, связанные с изменением количества бригад, очень высоки, то необходима высокая уверенность в том, что такое изменение действительно обязательно. Поэтому рекомендуется изменять количество бригад только после того, как будут получены последовательно три точки за одной и той же контрольной границей.

Когда решается вопрос об изменении количества бригад, то кроме графика рис. 21.4 следует учесть прогнозы по сбыту. Однако этот график наглядно показывает, когда работа трех бригад дает слишком большие или слишком малые запасы. Важно отметить, что этот график следует использовать при условии, что коэффициент планирования находится в контрольном диапазоне. Если же коэффициент планирования выходит из контрольного диапазона, то отношение запасы — сбыт также может выйти за заданные границы, но причина может заключаться не в бригадах, а в коэффициенте планирования.

Следует подчеркнуть, что оба вида контроля нельзя применять механически, требуется хорошее знание общей обстановки и конкретных условий.

ВНЕДРЕНИЕ В ПРОИЗВОДСТВО

После того как решение получено и испытано, его можно применить на практике. Может оказаться, что группа исследования операций не связана непосредственно с работой по внедрению, но такая практика порочна по двум причинам.

1. Независимо от того, сколько усилий было затрачено на получение решения и насколько тщательно оно было испытано, недостатки могут обнаружиться на этапе внедрения. Кроме того, могут потребоваться меры по улучшению решения. Если корректировка решения по результатам его практического использования находится в ведении лиц, не понимающих, как было получено решение, то эффективность такой корректировки будет незначительна. Например, производственники могут не усмотреть особого вреда в небольшом изменении, которое на самом деле иногда оказывается весьма существенным.

2. Проведение решения в жизнь может быть связано далеко не со столь очевидной процедурой, как она представлялась первоначально исследователям. Чтобы потенциальные возможности предложенного решения были полностью реализованы, необходимо перевести решение на язык рабочей процедуры. Такой перевод должен быть по возможности более точным, и только исследователи могут минимизировать вызванную им потерю эффективности решения.

Характер возникающих при внедрении трудностей зависит от того, является ли решение одноразовым или повторяющимся. В случае одноразового решения трудностей меньше, но они не устраняются полностью. В дальнейшем рассматриваются в основном более сложные повторяющиеся решения.

Задачи внедрения нельзя отделить от задач организации и административного управления исследованием, но, хотя эти вопросы и затронуты в рассматриваемом ниже примере, они детально разбираются только в следующей главе. В этой главе говорится только о преобразовании решения в рабочую процедуру.

Следует признать, что в исследовании операций пока что отсутствует четко разработанная методология внедрения. Основное внимание в исследовании операций уделялось разработке методов и средств принятия решений, но уже сейчас осознана важная роль задачи внедрения. Методология внедрения, которая сейчас принята, является довольно очевидной, но это не означает, что ею можно пользоваться механически. Однако гораздо правильней сознательно применить даже очевидные процедуры, чем полностью передать внедрение в руки производственников или даже руководства.

Чтобы преобразовать решение в рабочую процедуру, следует ответить на три вопроса:

1. Кто и что должен делать?
2. Когда?
3. Какая информация и какие средства нужны для этого?

Руководствуясь ответами на эти вопросы, можно разработать рабочую процедуру, провести необходимую подготовку персонала и практически внедрить решение.

Внедрение решения нельзя провести без участия людей, которые непосредственно проводят его в жизнь. Нужно выявить этих людей и определить их обязанности, что невозможно сделать, не изучив в деталях операции, выполняемые организацией, и принятое в ней разделение обязанностей. Анализ организации (см. главу 4) дает большую часть необходимой инфор-

мации, а дополнительная информация поступает от руководителей и исполнителей, работающих в контакте с группой исследования операций. Таким образом, на этой и всех остальных стадиях внедрения решения необходимо постоянное сотрудничество между руководством, исполнителями и операционистами. В следующей главе показано, как организовать такое сотрудничество.

Каждого работника, отвечающего в той или иной степени за внедрение, следует проинструктировать относительно тех обязанностей, которые ему надлежит выполнять. Следует предоставить в распоряжение этих работников необходимые средства и обучить их пользоваться ими, причем указанные средства не должны быть слишком сложны. Так, например, может оказаться целесообразным представить даже относительно простое уравнение в виде номограммы (как в примере, рассмотренном в главе 2) или таблицы. В некоторых случаях может потребоваться упрощение решения, даже если это повлечет за собой некоторую потерю его эффективности. Обычно решение реализуется кругом работников, математическая подготовка которых недостаточно высока. Поэтому, если мы хотим быть уверенными в том, что рекомендуемые правила принятия решений будут правильно применены, необходимо упростить их, прежде чем передать в руки исполнителей или руководителей. Это часто означает, что следует дать приближенный простой вариант изящного решения или полностью отказаться от тонких математических соображений и предложить «быстрое и грубое» правило принятия решения.

Следует, однако, отметить, что в некотором смысле почти каждое операционное решение является приближенным и в какой-то мере «быстрым и грубым». Это объясняется тем, что при построении каждой модели принимаются некоторые упрощающие допущения, которые уменьшают общность модели и полученных из нее решений, что, собственно, и означает признание определенной «быстроты и грубости». Для операциониста очень важно понять, что используемое приближенное решение может быть гораздо лучше более точного решения, которое нельзя использовать.

Приведем пример, показывающий важность и сложность этапа внедрения в исследовании операций.

Пример внедрения. В главе 4 приведен пример анализа работы компании, выпускающей двигатели для автомашин, небольших судов и строительного оборудования. Следует напомнить, что в результате анализа работы организации было обнаружено, что разработанные правила принятия решения по управлению производством и снабжением не могут разрешить всех трудностей. Это объясняется невозможностью внедрения многих рекомендаций. Трудности возникли в связи с длительностью процесса обработки заказа потребителя и недостатком информации, необходимой для применения разработанных правил принятия решений. В результате анализа системы были разработаны следующие рекомендации по внедрению предложенных правил принятия решений.

1. Уменьшить число различных подразделений, обрабатывающих заказ, с целью устранения дублирования в работе.

2. Обеспечить достаточный справочный материал для подразделений, обрабатывающих заказ, с целью упрощения и стандартизации процедур управления по многим заказам.

3. Правильно разместить подразделения, обрабатывающие заказ, с целью замены письменной связи устной.

4. Изменить формы документации по обработке заказов с целью включения в них только необходимой и достаточной информации.

5. Установить эффективный контроль за внесением технических изменений в изделия и за распространением соответствующей информации.

Предложенные изменения процедуры обработки заказа могли бы сократить время на обработку заказа с трех недель до одной недели, а также использовать оптимальные размеры партий выпускаемых и закупаемых предприятием изделий.

Предложенная система управления производством и запасами была ориентирована на использование вычислительной машины IBM 650. Система должна была функционировать на основе следующей общей процедуры.

1. Отдел контроля за материалами передает информацию о состоянии дел в отдел механизированного учета.

2. Стандартные решения можно обрабатывать на вычислительной машине IBM 650.

3. На основе информации, выданной отделом механизированного учета, отдел контроля за материалами формирует решения по производству и сбыту продукции.

При этом исходим ли мы из того, что разработанная система должна обеспечивать эффективное управление запасами, устранение нехватки продукции и необходимости интенсификации производства.

В результате введения новой процедуры обработки заказа и системы управления производством и запасами обеспечивалась возможность определения потребностей в деталях в среднем через неделю после получения заказа от потребителя. Используя вычислительную машину, можно непрерывно определять текущий уровень запасов. Кроме того, новая система позволяет определять будущие потребности в запасах.

Группа исследования операций, проработав в течение полугода, представила руководству компании разработанные правила принятия решений и рекомендации по внедрению. Руководство полностью одобрило предложенные рекомендации. Затем 12 месяцев было затрачено непосредственно на внедрение. Значительное оживление дел компании в этот период усложнило задачу внедрения, поскольку появились новые задачи. Если бы эти задачи не были учтены группой исследования операций, то эффективность разработанных правил принятия решений была бы в значительной мере снижена.

При разработке программы внедрения особое внимание уделялось ряду факторов, среди которых наиболее важными были:

Связь с руководством высшего уровня. В процессе внедрения проводились периодические устные отчеты руководству вместе с демонстрацией соответствующих материалов. Эти отчеты давали информацию о внедрении, причем подчеркивались не только достигнутые успехи, но также трудности, на которые руководству следовало обратить больше внимания. Отчеты были достаточно краткими и давали информацию об общем состоянии дел, но не содержали излишних подробностей, не требующихся для руководства высшего уровня.

Связь с ведущим составом (линейными руководителями). Вскоре после того, как руководство высшего уровня одобрило разработанные операционной группой мероприятия, было организовано совещание с ведущим составом компании. На этом совещании обсуждались различные нововведения, которые намечалось провести, а также объяснялись причины, побудившие провести исследования, и ожидаемые выгоды от внедрения разработанных рекомендаций. В начале совещания выступил один из руководителей компании, после которого выступали сотрудники группы исследования операций. В дальнейшем по мере необходимости устраивались встречи с ведущим

составом для обсуждения конкретных вопросов, связанных с внедрением.

Контакты с персоналом компании (исполнителями). Аналогичные совещания проводились и с остальным персоналом. Рассматривались конкретные детали работы отдельных подразделений. Проводились совещания в подразделениях, разъяснялись обязанности, давались ответы на вопросы, серьезно обсуждались возникающие трудности.

Особое внимание было уделено вопросу своевременного, тщательного сбора данных, которые поступают в информационную систему. Как бы хорошо ни была разработана система, ее работа будет определяться точностью поступающей информации. Например, в документах ОТК нужно правильно указывать наименование деталей, а также точное число проверенных, годных и забракованных деталей. Если бы эта информация была неправильной, то нельзя было бы эффективно управлять запасами, так как невозможны были бы заказы партий ненужных деталей или нехватка деталей, которую можно было бы предотвратить.

Системный подход. В процессе внедрения, так же как и на других стадиях исследования, операционная группа придерживалась системной точки зрения. Одним из средств, которое группа нашла эффективным в этом смысле, было использование временного графика внедрения. Работы по внедрению были разделены на несколько логических этапов. Для каждого этапа был разработан календарный план, на котором отображалось фактическое состояние проекта, периодически корректировавшееся. Благодаря этому группа исследования операций имела возможность оценивать ход выполнения проекта внедрения, не теряя из поля зрения общих целей.

В процессе внедрения группа исследования операций продолжала улучшать и дорабатывать новую систему управления производством и запасами. Часто первоначальные рекомендации видоизменялись в свете новых данных. Зная конечную цель, исследователи могли оценить, как влияет каждое изменение на систему в целом.

Внедрение такой системы, охватывающей почти все операции компании, было длительным процессом. Имели место случаи, когда возникало некоторое разочарование вследствие кажущегося отсутствия немедленных результатов. Одной из важных задач операционистов было выявление таких настроений и принятие неформальных мер для их ликвидации. Умение установить личные контакты является одной из важнейших и наиболее сложных задач на этапе внедрения. Успешное решение этой задачи дает ключ к успеху внедрения в целом.

ЗАКЛЮЧЕНИЕ

Часто ученые теряют интерес к задаче, когда она решена на бумаге. Они думают о следующей задаче и оставляют «грязную» работу по контролю и внедрению другим. В результате многие решения так и остаются только на бумаге или при внедрении дают неутешительные результаты. Исследователи должны ясно представлять себе, что от правильной постановки контроля и внедрения зависит успех проекта. Кроме того, при этом происходит подлинная проверка научных методов. При внедрении мы имеем дело с наиболее трудной задачей, возникающей перед наукой, отдельной личностью и социальными группами.

Как уже указывалось, разработка системы «управления» решением и программы внедрения есть только часть работы, связанной с проведением решения в жизнь. На этих, как и на других этапах исследования, очень

важное значение имеют организационные мероприятия, обучение и административное управление. Руководство, те, кто проводит в жизнь его решения, и исследователи образуют коллектив, успех работы которого зависит от того, насколько рационально были организованы индивидуальные и совместные усилия его членов, как они были подготовлены для совместной работы и насколько хорошо ведется административное управление. Эти вопросы рассматриваются в следующей главе.

ЛИТЕРАТУРА

1. A c k o f f R. L., Allocation of Sales Effort in Proceedings of the Conference on «What Is Operations Research Accomplishing in Industry?», Case Inst. Techn., 1955.
2. A c k o f f R. L., Production and Inventory Control in a Chemical Process, J. Opns. Res. Soc. Am., 3, 319—338 (1955).
3. C h u r c h m a n C. W., Theory of Experimental Inference, New York, 1948.
4. D a v i d s o n W. E., Operations Research in the Lamp Division of the General Electric Company, in Proceedings of the Conference on «What Is Operations Research Accomplishing in Industry?», Case Institute of Technology, Cleveland, 1955.
5. D i x o n W. J. and M a s s e y F. J., Jr., Introduction to Statistical Analysis, McGraw-Hill Book Co., New York, 1951.
6. D u n c a n A. J., Quality Control and Industrial Statistics, Chicago, 1952.
7. F e r r i s C. D., G r u b b s F. E. and W e a v e r C. L., Operating Characteristics for the Common Tests of Significance, Ann. math. Statist., 17, 178—197 (1946).
8. G e a r y R. C. and P e a r s o n E. S., Tests of Normality, London, 1938.
9. G r a n t E. L., Statistical Quality Control, New York, 2nd ed., 1952.
10. H a l d A., Statistical Theory with Engineering Applications, New York, 1952.
11. H a r e V. C., Jr., and H u g l i W. C., Applications of Operations Research to Production Scheduling and Inventory Control, II, Proceedings of the Conference on «What Is Operations Research Accomplishing in Industry?», Case Inst. Techn., 1955.
12. H i c k o x C. H., The Role of the Company Technician in Operations Research, in Proceedings of the Conference on «What Is Operations Research Accomplishing in Industry?», Case Institute of Technology, Cleveland, 1955.
13. J u r a n J. M. (ed.), Quality Control Handbook, New York, 1946.
14. K e n d a l l M. G. and S m i t h B. B., Randomness and Random Sampling of Numbers, J. R. statist. Soc., 101, 147—166 (1938).
15. K e n n e d y C. W., Quality Control Methods, Prentice-Hall, New York, 1948.
16. K n e c h t F. W., Jr., Applications of Operations Research to Production Scheduling and Inventory Control, I, in Proceedings of the Conference on «What Is Operations Research Accomplishing in Industry?», Case Institute of Technology, Cleveland, 1955.
17. L i t t a u e r S. B., Social Aspects of Scientific Method in Industrial Production, Phil. Sci., 21, 93—100 (1954).
18. L i t t a u e r S. B., Technological Stability in Industrial Operations, Trans. N. Y. Acad. Sci., Ser. II, 13, no. 2, 66—72 (1950).
19. L i t t a u e r S. B., The Development of Statistical Quality Control in the United States, Amer. Statistician, 4, 14—20 (1950).
20. M o s t e l l e r F r e d e r i c k, On Some Useful «Inefficient» Statistics, Ann. math. Statist., 17, 377—408 (1946).
21. N e y m a n J e r z y and P e a r s o n E. S., On the Problem of the Most Efficient Tests of Statistical Hypotheses, Phil. Trans., Ser. A, 231, 289—337 (1933).
22. P e a c h P a u l, An Introduction to Industrial Statistics and Quality Control, Edwards & Broughton Co., Raleigh, 2nd ed., 1947.
23. R u t h e r f o r d J. G., Quality Control in Industry — Methods and Systems, Pitman Publishing Corp., New York, 1948.
24. S h e w h a r t W. A., Statistical Methods from the Viewpoint of Quality Control, U. S. Department of Agriculture, Washington, 1939.
25. T i p p e t t L. H. C., Technological Applications of Statistics, New York, 1950.
26. W a l d A b r a h a m, Contributions to the Theory of Statistical Estimation and Testing Hypotheses, Ann. math. Statist., 10, 299—326 (1939).
27. W a l d A b r a h a m, On the Principles of Statistical Inference, in «Notre Dame Mathematical Lectures», no. 1, Notre Dame University, Notre Dame, 1942.
28. W a l d A b r a h a m, Statistical Decision Functions, Ann. math. Statist., 20, 165—205 (1949).
31. У. Черчмен и др.

Часть X

АДМИНИСТРАТИВНОЕ УПРАВЛЕНИЕ ОПЕРАЦИОННЫМИ ИССЛЕДОВАНИЯМИ

Перед организацией, стремящейся создать свое собственное операционное подразделение, возникают три важные административные задачи.

1. Отбор персонала для операционной группы из числа сотрудников организации или вне ее.

2. Обучение (подготовка) отобранных специалистов (кроме того редкого случая, когда представляется возможность нанять на работу целую подготовленную группу).

3. Организация операционной группы и ее работы.

Эти три задачи рассматриваются в главе 22, хотя третья задача была поставлена уже в главе 21, когда обсуждались проблемы практического внедрения результатов операционных исследований.

Глава 22

ОТБОР И ПОДГОТОВКА ПЕРСОНАЛА И ОРГАНИЗАЦИЯ ОПЕРАЦИОННЫХ ИССЛЕДОВАНИЙ

ВВЕДЕНИЕ

По-видимому, наступит день, когда утверждение, что данный специалист является операционистом, будет иметь столь же определенное значение, какое имеет, например, сегодня утверждение, что этот ученый — физик-теоретик. Титул «операционист» будет в общих чертах определять также то, какое образование получил его владелец и какую работу он способен выполнять. Но пока эта пора не наступила, отбор и подготовка персонала будут оставаться важными проблемами для любой организации, которая намерена создать собственное подразделение исследования операций.

В результате создания Американского общества исследования операций и Института наук организационного управления, а также бурного роста числа публикаций в этой области знаний (особенно благодаря изданию регулярно выходящих журналов «Operations Research»* и «Management Sciences») за последние несколько лет одна из указанных проблем несколько упростилась. Научно-исследовательские работники многих отраслей знаний получили представление об исследовании операций и заинтересовались этим предметом. В итоге частично облегчилась задача тех, кто при отборе студентов и сотрудников был вынужден прежде объяснять, что такое исследование операций, как говорится, «от печки». И теперь появилась возможность проводить более тщательный выбор, чем несколько лет назад. Однако проблема, кого именно отбирать и как обеспечить подготовку будущих операционистов, все еще не снята с повестки дня.

Даже в давно сложившихся науках чрезвычайно трудно сформулировать эффективные принципы отбора, подготовки и организации работы специалистов. Лучшее, что возможно сделать в новой области, — это предложить интуитивные правила, основанные на ограниченном опыте специалистов, выполнявших практические исследования. Поскольку этот опыт практиков весьма разнообразен, предлагаемые правила также существенно отличны друг от друга. Поэтому данная глава не содержит никаких претензий в смысле неоспоримости высказываемых в ней взглядов.

Значительно больше внимания уделялось организации крупных самостоятельных операционных групп, работающих над военными задачами, чем небольшим группам, занимающимся решением производственных задач и работающим непосредственно в промышленности. Отличное изложение принципов работы крупных операционных коллективов содержится в статье Макклоски [6]. Однако методика организации таких коллективов лишь в незначительной мере применима к промышленным операционным группам. Тем не менее проведению научно-исследовательских работ в промышленности и общим вопросам организации научно-исследовательских коллективов уделялось большое внимание [1, 2, 4, 5], и некоторые из полученных результатов можно использовать применительно к рассматриваемым здесь задачам. Основная проблема, возникающая в тех отраслях промышленности, которые заинтересованы в проведении операционных исследований, сводится к вопросу: с чего нам начинать? Вообще говоря, существует три пути решения этой проблемы.

* Прежнее название «Journal of the Operation Research Society of America».

1. Организовать группу из числа сотрудников фирмы, и если никто из них не имеет достаточной подготовки по исследованию операций, то нанять на работу в штат фирмы квалифицированного операциониста.

2. Сформировать группу, включающую как сотрудников фирмы, так и специалистов консультационной операционной фирмы, с которой заключается контракт на период, пока сотрудники данной фирмы не смогут выполнять исследования самостоятельно.

3. Заключить договор с консультационной операционной фирмой на проведение всех операционных исследований в данной компании без привлечения ее собственных сотрудников.

Последний путь выгоден тогда, когда требуется срочно решить некоторую конкретную задачу. Однако если пойти по этому пути, то компания не получает возможности проводить операционные исследования собственными силами. Поэтому в дальнейшем рассматриваются только первые два из указанных путей.

ОТБОР ПЕРСОНАЛА В ОПЕРАЦИОННУЮ ГРУППУ

Нас интересуют главным образом деловые качества, а не индивидуальные наклонности лиц, включаемых в операционную группу. Проблема отбора, естественно, связана с особенностями характера, однако она не содержит в себе ничего уникального по сравнению с проблемой отбора на любую научно-исследовательскую работу, если не считать одного фактора: весьма существенной является способность общения с производственниками, не имеющими специальной научной подготовки. Работающий в промышленности операционист должен уметь доступно излагать специальные технические вопросы. Он должен обладать способностью усваивать принятый в фирме или отрасли промышленности жаргон в такой степени, чтобы не только изъясняться на нем, но и понимать, когда им пользуются другие. Но самое важное заключается в том, что операционист должен с уважением относиться к хорошему администратору, мастеру и рабочему и никогда не проявлять высокомерия при общении с работниками, не имеющими научной подготовки.

В операционной группе должен быть по крайней мере один сотрудник, хорошо осведомленный по всем вопросам деятельности данной промышленной организации, и один ученый или инженер, обладающий богатым опытом и хорошей научной подготовкой. Один сотрудник может сочетать в себе оба качества, и в целом ряде случаев операционные исследования начинались успешно проводиться именно такой «группой», представленной одним специалистом. Однако организация группы в составе не менее двух человек обладает существенными преимуществами, ибо они могут не только обдумывать совместно стоящие перед ними проблемы, но и плодотворно «возражать друг другу», используя различные научные методы. Получаемые в итоге результаты могут оказаться больше простой суммы частей. Кроме того, опыт свидетельствует, что группа в составе двух сотрудников более быстро расширяет сферу своей деятельности и повышает эффективность, чем один индивидуально работающий операционист.

Ценность операциониста, знающего деятельность фирмы, прямо пропорциональна широте его знаний. Глубина знаний любой конкретной области этой деятельности не столь важна, как широта, ибо в исследовании операций фундаментальную роль играют взаимодействия. (Специализированные, детальные сведения можно получить, привлекая в случае необходимости тех работников фирмы, которые являются узкими специалистами

по соответствующим вопросам.) Этот операционист должен иметь также склонность к количественному мышлению, но вовсе не обязательно обладать высокой подготовкой по математике или ее приложениям. Важно только, чтобы его не смущали числа и символы и чтобы он стремился расширить свои математические познания. Ему не обязательно быть ученым или инженером. Эффективные операционные группы включали бухгалтерских работников, сотрудников отдела сбыта, снабжения, юристов и администраторов.

Если же этот операционист инженер по образованию, то второй сотрудник, если это представляется возможным, должен иметь чисто научную подготовку. Наоборот, если он сотрудник теоретического плана, то второй сотрудник, опять-таки если это возможно, должен быть инженером. Такая комбинация обеспечивает желательное равновесие между теорией и приложениями.

Ученый, стремящийся к установлению истины ради истины, мало подходящая кандидатура на роль операциониста. То же самое относится и к ученому, привязанному к лабораторным экспериментам, ибо сложные связи, характеризующие реальный мир, слишком резко отличаются от регулируемых лабораторных условий. Ученый, стремящийся к совершенству, которого больше интересует «полнота, завершенность» конечного результата, чем решение поставленной задачи в намеченный срок, также едва ли подходит на роль операциониста. А вот инженер или ученый, работающий в двух или трех близких областях специализации только потому, что не удовлетворялся «решениями», содержащимися в каждой отдельной области, является удачным кандидатом в операционисты.

Сотрудники операционной группы должны принадлежать к разряду «методологов», т. е. к числу исследователей, интересующихся прежде всего тем, каким образом проводится исследование, а не конкретным, специфическим содержанием изучаемой проблемы. Они должны представлять собой ученых, заинтересованных в постановке задач, построении моделей и разработке методов оценки адекватности моделей и измерения так называемых «неуловимых» переменных. Эти специалисты не должны быть «привязаны» к единственному набору научных или технических средств либо справочников. Они должны обладать гибкостью и проявлять интерес к возможности проведения исследований в тех областях, которые считались недоступными для точного изучения.

По меньшей мере один из сотрудников операционной группы должен иметь возможность читать и понимать растущую литературу по исследованию операций, особенно источники, указанные в этой книге. Ему необходимо предоставить возможность работы над литературой без отрыва на другие работы.

В группе, состоящей из двух человек, желательно, чтобы один из специалистов имел склонность к анализу фактических данных, а другой — к теории. Большинство ученых и инженеров принадлежит к тому или иному типу специалистов, и лишь немногие сочетают в себе оба качества. Операционная группа, все сотрудники которой увлекаются лишь анализом фактических данных или погружены исключительно в теорию, была бы так сказать «неуравновешенной». Исследователи, занятые только фактическим материалом, имеют тенденцию игнорировать этап построения модели при проведении исследования, а теоретики игнорируют фактические данные, необходимые для обоснования их моделей. Только сочетание специалистов обоих типов может обеспечить всесторонний подход к решению операционных задач.

Поскольку проблема обеспечения контактов и связи между сотрудниками научного коллектива проще всего решается при условии, когда они с уважением относятся к различным отраслям знаний, следует с осторожностью относиться к ученому, допустим физику, хмурящемуся, когда говорят о его совместной работе с инженерами, вздрагивающему при упоминании о сотрудничестве с экономистами и стонущему при мысли о контактах с психологом или социологом. Такой ученый слишком ограничен, чтобы оценить вклады в науку, сделанные представителями других отраслей знаний.

Если речь идет об организации расширенной операционной группы или о ее последующем расширении, то можно рекомендовать следующий ориентировочный состав группы: (1) один сотрудник фирмы, (2) один научный работник (представитель естественных наук), (3) один инженер, (4) один математик или статистик, (5) один научный работник (представитель биехиристических наук: биологии, психологии или общественных наук), (6) один логик или специалист по научной методологии. Расширение за пределы этого списка имеет целью создать максимально возможное представительство различных отраслей науки и техники, причем с учетом требований, предъявляемых к личным качествам специалистов с точки зрения коллективной работы.

Если ни один из сотрудников группы не имеет специальной операционной подготовки или опыта работы в этой области и такой специалист не зачисляется в штат, то необходимо прибегнуть к помощи консультантов или заключить на некоторый срок контракт с опытным операционистом. Существует ряд консультационных фирм по вопросам организационно-о управления, а также научно-исследовательских и учебных институтов, которые могут предоставить квалифицированную помощь. Регулярные консультации с опытным операционистом по проводимой работе обеспечивают один из наиболее эффективных способов подготовки специалистов, начинающих заниматься исследованием операций. Кроме того, при этих условиях организация, финансирующая работу, получает гарантию, что эффективные средства решения проблем не будут по неведению упущены из вида.

Требование смешанного в отношении специальностей состава операционной группы не является произвольным. Оно основано на строгом научном наблюдении: многие объекты характеризуются самыми различными проявлениями и поэтому изучались независимо многими науками. Привлекая богатый арсенал различных методов, разработанных для изучения некоторой конкретной системы, можно получить наиболее ценные результаты. Ярким подтверждением этого является поразительное развитие знаний в области систем связи и автоматического управления, обусловленное совместной работой ученых и инженеров, принадлежащих самым различным отраслям знаний и создавших такую науку, как кибернетика. Было установлено, что физики, химики, биологи, психологи, специалисты общественных наук, электротехники, философы и представители многих других отраслей знаний — каждый со своей стороны подходит к изучению связи и управления в различных явлениях и объектах. Их объединение привело к появлению «общего языка», а синтез их знаний дал одно из наиболее фундаментальных достижений современной науки.

Ученые, получившие одинаковую подготовку, не могут дополнять друг друга в той мере, в какой это обеспечивают представители различных областей. Кроме того, классы задач, для решения которых применяется исследование операций, требуют привлечения специалистов из различных областей, ибо при системном подходе, характерном для этой дисциплины, объектом

изучения являются самые разнообразные стороны функционирования исследуемой системы. Пример профессора Миттена, изложенный в главе 3, ярко иллюстрирует справедливость этого утверждения. Объединение трех специалистов столь различного профиля в одну группу для проведения операционного исследования не может не показаться на первый взгляд странным, однако в дальнейшем становится ясно, что такой подбор сотрудников группы основан на прочном логическом фундаменте. Гораздо большую опасность представляет недостаток представителей, чем их избыток.

Приведем вкратце примеры нескольких операционных групп, организованных рядом фирм, с которыми Кейсовский технологический институт имел договоры о сотрудничестве на период проведения первых операционных исследований.

Владелец станкостроительной фирмы выделил одного финансового работника, бывшего помощником казначея, для работы с двумя операционистами, сотрудниками внешней организации. К работе периодически привлекались и другие сотрудники фирмы, но этот финансист был центральной фигурой, на которой «замыкались» все полученные результаты. Он был хорошо осведомлен о всех сторонах деятельности фирмы и имел беспрепятственный доступ во все ее подразделения.

Химическая компания выделила начальника ОТК в качестве руководителя операционной группы. Он имел степень кандидата наук по физической химии и привлек к работе сотрудников своего отдела — специалистов по контролю качества и статистиков. Группа была усилена сотрудником бухгалтерии, имевшим также подготовку по химии.

Компания, выпускающая кухонное оборудование, выделила одного из помощников вице-президента, отвечавшего за перспективное планирование, в качестве руководителя своей операционной группы. Он не имел солидной научной или технической подготовки, но обладал богатым опытом анализа деятельности компании. В его распоряжение был выделен молодой инженер, а начальник отдела исследования сбыта оказывал постоянную помощь в работе этой группы.

Фирма, производящая электрические лампы, организовала операционную группу в составе трех человек во главе с главным ревизором фирмы, в которую были включены помощник начальника отдела сбыта и специалист по исследованию рынка.

Нефтяная компания поставила во главе операционной группы специалиста по административным вопросам коммерческой деятельности и включила в нее инженера-механика, работавшего в исследовательском секторе производственного отдела.

Фармацевтическая фирма организовала операционную группу, наняв на работу квалифицированного статистика и выделив в его распоряжение специалиста по организации производства, проработавшего в этой фирме несколько лет.

Компания по производству автомобильных частей сформировала группу в составе начальника сектора планирования одного из производственных циклов и начальника ОТК.

Операционная группа металлургической компании состояла из одного из ведущих работников бухгалтерии и начальника ОТК.

Список этих примеров можно значительно продолжить. Однако уже ясно, что начать проведение операционных исследований не так уж сложно, как может показаться на первый взгляд. Однако подготовка отобранного в операционное подразделение персонала, которая обеспечивала бы эффективную работу этих специалистов, действительно представляет трудную

задачу. Естественно, что успех любой программы подготовки в значительной мере зависит от деловых качеств отобранных кандидатов. В каждом из упомянутых примеров удалось создать работоспособную, эффективную операционную группу.

ПОДГОТОВКА ОПЕРАЦИОНИСТОВ

Подготовку по исследованию операций можно обеспечить либо путем обучения на специальных курсах, либо путем специализации в этой области в рамках учебной программы вуза. При этом проходящий обучение работник или зачисленный в штат операционист должен изучить организацию, в которой ему предстоит работать, конечно, при условии, что он с ней еще недостаточно знаком. Рассмотрим именно такую ситуацию.

Изучение организации требует затрат времени. В течение этого времени должно возникнуть чувство взаимного уважения между специалистом, ничего не знающим о данной организации, но владеющим методами, которые могут оказаться для нее полезными, и специалистом, досконально знакомым с внутренней структурой организации, но не располагающим средствами для решения задач, которые представляются ему важными с точки зрения интересов организации.

Эту сторону организационной подготовки можно ускорить. В распоряжение нового работника следует предоставить как можно больше информации об организации, ему должна быть предоставлена «свобода перемещений» в рамках данной организации, возможность встречаться с ее руководителями и проводить наблюдения за деятельностью различных подразделений. Если предприятия и учреждения компании расположены в различных географических пунктах, то этот этап подготовки должен включать посещение ряда пунктов.

Все эти замечания могут показаться очевидными, но, к сожалению, слишком часто руководство поручает новому специалисту решение сложной задачи прежде, чем закончится период ознакомления. Это напоминает случай с одним агентом по сбыту, которому просто вручили книгу заказов и приказали разместить заказы.

Подготовка во время проведения конкретных исследований. Фирма может первоначально подойти к решению операционных задач, используя группы в составе не имеющих специальной подготовки сотрудников фирмы и квалифицированных операционистов, привлекаемых на договорных началах из консультационных фирм. В таких условиях сотрудники фирмы, проходящие операционную подготовку непосредственно в ходе работы над проектом, должны быть освобождены от своих повседневных обязанностей в такой мере, чтобы они могли уделять необходимое время операционному проекту.

Нужно также предоставить соответствующее время для работы над литературой, рекомендованной опытными операционистами, и для обсуждения прочитанного материала. Следует, кроме того, выделять время на посещение конференций, проводимых различными институтами, учреждениями и профессиональными обществами. Число таких конференций, посвященных исследованию операций, возрастает с каждым годом.

По мере того, как неподготовленные сотрудники приобретают опыт, следует уменьшать объем работы, выполняемой внештатными специалистами, которые на заключительном этапе подготовки сотрудников фирмы

используются в роли чистых консультантов. Однако и после завершения подготовки иногда целесообразно прибегать к помощи консультантов, чтобы операционисты фирмы были в курсе последних достижений в области исследования операций.

Подготовку операционистов в период работы над первыми операционными задачами, решаемыми в данной организации, можно дополнить подготовкой по специальным учебным программам. Можно заключить договоры, и такие договоры неоднократно заключались, с учебными институтами на организацию специальных учебных курсов по исследованию операций для специалистов, выделяемых в операционные группы. Лекции и практические занятия по программе таких курсов проводятся непосредственно в тех фирмах, с которыми заключены договоры. Можно также организовать курсы и для руководства фирмы. Эти курсы предназначены для того, чтобы дать руководству информацию о возможностях исследования операций: разъяснить сущность операционных методов, дать представление о задачах, которые решались или могут решаться этими методами, и изложить требования, предъявляемые к руководству при использовании исследования операций. Прослушав цикл лекций такого характера, руководители становятся достаточно осведомленными и приобретают интерес к исследованию операций, что позволяет им принять активное участие в развитии операционных исследований в данной фирме.

Если организация таких специальных курсов по каким-либо причинам невозможна, то их можно заменить посещением «краткосрочных» курсов по исследованию операций. В Массачусетском технологическом институте читается такой двухнедельный курс лекций [8], а Кейсовский технологический институт проводит аналогичный цикл лекций дважды в течение каждого учебного года. Эти лекции представляют собой насыщенную программу обучения для специалистов-производственников, предназначенных для работы в операционных группах. Подобные курсы не выпускают «законченных» операционистов-практиков, но обеспечивают базу, на основе которой можно совершенствоваться, работая над литературой и решая практические задачи. Ряд учебных институтов проводит примерно аналогичные циклы лекций по смежным с исследованием операций предметам, в частности по прикладной математике и вычислительной технике.

Другой способ обеспечения подготовки операционистов, так сказать без отрыва от производства, заключается в направлении выделенных сотрудников на стажировку в какую-либо организацию, специализирующуюся по исследованию операций. Договоры на стажировку можно заключить с некоторыми операционными группами, созданными при учебных институтах.

Учебные программы подготовки по исследованию операций. В поле зрения появляется все большее число учебных программ по исследованию операций. В Кейсовском технологическом институте разработана обширная учебная программа подготовки кандидатов и докторов наук по исследованию операций. Примерно аналогичная программа введена в сентябре 1955 года в Университете Джонса Гопкинса. В Массачусетском технологическом институте принимаются к защите работы по исследованию операций на соискание ученых степеней, а Колумбийский университет имеет программу подготовки специалистов по операционной технике, являющуюся по существу программой по исследованию операций. Такие учебные заведения, как Пенсильванский университет, Американский университет, Технологический институт Иллинойс, Технологический институт Карнеги, Университет Джорджа Вашингтона, Университет штата Огайо, Ренссилир

и Университет штата Калифорния (Лос-Анжелос), предлагают специальные программы подготовки студентов или аспирантов, включающие обширный курс исследования операций *.

Пока что эти учебные заведения выпускают ежегодно лишь незначительное количество специалистов по исследованию операций. Их число явно недостаточно для удовлетворения потребностей, однако эти операционисты представляют собой потенциальных преподавателей, которые в будущем смогут обеспечить подготовку все большего количества специалистов в этой области по мере увеличения числа учебных заведений, где вводится преподавание исследования операций. Промышленные организации могут оказать финансовую помощь в деле развития образования операционистов, предоставляя стипендии студентам, выбирающим исследование операций в качестве основной профессии. Эти студенты не берут на себя обязательство поступить на работу в фирму или компанию, предоставившую стипендию, однако аналогичные затраты на развитие других отраслей знаний вполне себя оправдали и наверняка оправдают и в этой области.

ОРГАНИЗАЦИЯ

В отношении организационной стороны деятельности операционных групп в составе фирмы чаще всего задают следующие вопросы:

1. Какое место должна занимать группа в организационной структуре фирмы?
2. Как осуществлять административное руководство группой?
3. С каких задач она должна начинать и к каким переходить в дальнейшем?

Рассмотрим поочередно все эти вопросы.

Кому должна подчиняться операционная группа? Существует две принципиально различные точки зрения по этому вопросу. Одни считают, что группа может непосредственно подчиняться только руководителю самого высокого уровня. Другие убеждены, что стационарная группа должна отстоять от высшего уровня на несколько звеньев и подчиняться руководителю того уровня иерархии, где выполняется «реальная» работа, ибо именно на этом уровне формируются фактические данные и на него влияет решение «реальных» задач. Промежуточная позиция сводится к тому, что группу ставят на самый низкий из возможных уровней иерархической лестницы, обеспечивая обязательно беспрепятственный доступ к любой информации на всех уровнях. На практике успешно функционирующие операционные группы находились в непосредственном подчинении и президентов компаний, и первых вице-президентов, и других вице-президентов, и начальников отделов, и эквивалентных им административных работников.

По нашему опыту, наиболее эффективной является такая организация, когда операционная группа подчиняется административному правлению, состоящему из руководителей основных подразделений компании. Это постоянно действующий орган постоянного состава, рассматривающий все проекты, выполняемые группой и заслушивающий ее отчеты о ходе работ периодически раз в месяц или в два месяца. Один из членов правления, руководитель того подразделения, которое больше всего затрагивает выполняемый проект, выступает в роли председательствующего, и именно к нему должны обращаться сотрудники группы за помощью, когда в этом возникает необходимость.

* Сведения по этому вопросу, относящиеся к Англии, содержатся в [8].

Кроме того, весьма полезно организовать рабочий комитет в составе административных работников, подчиненных руководителям основных подразделений, которых больше всего будут затрагивать результаты проводимых исследований. Этот комитет также должен регулярно встречаться с сотрудниками группы, но состав комитета целесообразно менять по мере изменения объектов исследования.

При наличии таких органов отпадает необходимость частого представления письменных отчетов. В идеальном случае подготавливается лишь один отчет по каждому проекту, содержащий сведения о том, что было фактически сделано к моменту завершения проекта, и служащий в качестве архивного, справочного материала. В той мере, в какой можно регулярно устно отчитываться перед заинтересованными руководителями, получаемые результаты будут усваиваться постепенно, и их можно планомерно внедрять по мере появления. Такие встречи приводят к тому, что лица, которым в конечном счете пришлось бы проводить рекомендуемые мероприятия, становятся реальными участниками принимаемых решений, что сводит к минимуму сопротивление предлагаемым изменениям. Иногда целесообразно представлять краткие письменные доклады (записки), более детально освещающие результаты или методы, в которых трудно разобраться в течение одного заседания. Сообщения для правления советников и рабочего комитета необходимо готовить так же тщательно, как письменные отчеты. Следует также широко прибегать к помощи наглядных средств (плакаты, графики и т. п.).

Сама операционная группа должна быть невелика по численности, а тенденцию к расширению состава только вследствие того, что велик объем работы, следует всячески пресекать. В то же время группа должна быть достаточно велика, чтобы в ней были представлены все необходимые специальности и чтобы обеспечивалось взаимное стимулирование и «генерация» идей в рамках самой группы.

На первых порах деятельности операционной группы необходимо гарантировать высокое постоянство заданий отчасти для того, чтобы упростить задачи подготовки специалистов, а отчасти для облегчения задачи накопления максимального количества информации об организации, обслуживаемой группой.

Если все сотрудники группы имеют одинаковый опыт или в равной мере им не обладают, не следует назначать кого-либо руководителем. Если же один из сотрудников имеет больший опыт, чем все остальные, то формально его можно назначить руководителем группы, хотя такое назначение обычно не требуется, пока численность группы не превысит пяти человек. Внутренняя организация операционной группы должна быть максимально демократичной в пределах допустимых в настоящее время «свобод».

Почти в каждой организации, где созданы операционные группы, различные руководители и инженерно-технические работники неизменно начинают обращаться к различным сотрудникам группы за советами или помощью в решении тех задач, по которым эти сотрудники являются специалистами.

Обеспечение таких «услуг» может играть важную роль в деятельности группы, но текущим консультациям следует уделять лишь незначительное время. Такие возможности оказать услуги фирме позволяют более быстро добиться признания роли операционистов и упрочить положение группы в организационной структуре. Однако группе должно быть предоставлено право отказа, если обращения за советами или помощью становятся слишком частыми.

С решения какого рода задач рационально начинать работу операционной группы? Вообще говоря, исходные задачи не должны быть слишком сложны или требовать очень много времени, а их решения должны быть реально ощутимыми. При таком подходе сама группа и руководство приобретают уверенность в себе и в методах исследования операций. Постепенно можно переходить к более сложным задачам.

Задачи управления производством и запасами обычно отвечают таким требованиям. Методы и средства их решения достаточно развиты. Имеются многочисленные примеры успешного практического применения решений задач этого класса, описанные в литературе.

В конечном счете группа должна перейти к решению задач, связанных с исследованием спроса и организационной структуры фирмы, а также к задачам стратегического характера. В ходе работы группы будут ставиться новые, все более общие задачи. Если первая задача успешно решена, то маловероятно, что будет ощущаться недостаток задач в дальнейшем.

В заключение следует отметить, что развитие операционных групп, предназначенных для исследования задач организационного управления, все еще связано с проявлением творческой инициативы и изобретательности. Идеальная группа должна обладать следующими качествами:

1. Глубоко понимать сущность научного метода и уметь его применять.
2. Проявлять подлинный интерес к задачам организационного управления.
3. Обладать способностью устанавливать контакты с персоналом фирмы и уметь завоевать его доверие.
4. Иметь опыт практического использования методов исследования операций и быть в курсе последних достижений в этой области.

Известно, что такие идеальные группы не существуют и вряд ли вообще появятся. Но по мере накопления опыта можно будет высказать более определенные утверждения относительно того, как приблизиться к идеалу. В конечном счете следует ожидать, что удастся применить сами методы исследования операций для создания промышленных операционных групп.

ЛИТЕРАТУРА

1. A c k o f f R. L., The Design of Social Research, University of Chicago Press, Chicago, 1953.
2. B u s h G. P. and H a t t e r y L. H., Teamwork in Research, American University Press, Washington, 1953.
3. G a n d e r R. S., Training for Operations Research at the University of Birmingham, J. Opns. Res. Soc. Am., 3, 215—216 (1955).
4. H e r t z D. B. and R u b i n s t e i n A. H., Research Operations in Industry, King's Crown Press, New York, 1953.
5. H e r t z D. B. and R u b i n s t e i n A. H. (eds.), Selection, Training, and Use of Personnel in Industrial Research, King's Crown Press, New York, 1953.
6. M c C l o s k e y J. F., Training for Operations Research, J. Opns. Res. Soc. Am., 2, 386—392 (1954).
7. M c C l o s k e y J. F. and T r e f e t h e n F. N., Operations Research for Management, The Johns Hopkins Press, Baltimore, 1954.
8. M o r s e P. M., Report on the First Summer Program in Operations Research at the Massachusetts Institute of Technology, J. Opns. Res. Soc. Am., 1, 303—305 (1953).
9. Report of The Education Committee, J. Opns. Res. Soc. Am., 1, 248—251 (1953).
10. R i n e h a r t R. F., Threats to the Growth of Operations Research in Business and Industry, J. Opns. Res. Soc. Am., 2, 229—233 (1953).
11. S c h u l t z A n d r e w, Jr., A Discussion of Education for Operations Research, J. Opns. Res. Soc. Am., 3, 91—100 (1955).

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Автономная организация 75
Алгоритм Джонсона 363
Альтернативные стратегии 99
Анализ потока автомашин 323
— тенденций рынка 52
Аукционный торг 24, 427

«Взвешивание» 102, 107
Венгерский метод 267, 281
Вероятность получения контракта 432
Время обслуживания 307
— ожидания обслуживания 307
Выборка 82, 442
Врожденный план 253

Гомоморфная модель 134
График 136
— сборки 45

Двойственная задача о назначении 266
Диаграмма Ганта 346, 357
Доминирование 421

Задача баланса 35
— коммивояжера 361
— массового обслуживания 155
— о назначении 214, 263, 278
— о случайном блуждании 147
— организационно-управленческая 13
— оценки 100
— разработки 100
— распределения 154
— состязания (конкуренции) 157
— транспортная 265
— управления запасами 153
Заказы на запчасти 41
Закрытые торги 430
Закупка и хранение сырья 45
Замена 380
—, затраты при замене 370
— оборудования 55
—, скорость замены 380, 383
—, стоимость замены 387

Игра без седловой точки 404
— конкурентная 398
— прямоугольная 400
— с ненулевой суммой 399
— с нулевой суммой 399
— — n лиц 417
— с седловой точкой 403
Изменение машинного времени 246

Изменение прибыли на изделие 249
— производительности 247
Изобразительные модели 135
Изоморфные модели 135
Исследование операций 12, 19, 25
Итеративный метод 414
Итерации 21

Календарное планирование 345
Квадратичное программирование 162
Коалиции 417
Консультанты 14
Контрстратегии 101
Корректировка параметров 454
— решения 453
Кривые выхода из строя 378
Критерий оптимальности плана 238
— эффективности 21

Линия Пуассона 339

Максимизация ожидаемого дохода 429
— разности доходов 427
Матрица выигрышей 418
— назначения 261
— оценок 261
Мера надежности оценок 122
— эффективности 108
Метод Акерса — Фридмена 356
— венгерский 267, 281
— Монте-Карло 22, 147, 311
— Ньютона 158
— сокращенной матрицы 267
— транспортный 217, 252
— упорядочения 125
— Флада 267
Модели-аналоги 136
Модели геометрического подобия 135
— гомоморфные 134
— замены и ремонта 156
— изоморфные 134
— иерархических систем 162
— запасов при двух уровнях цен 191
— — — трех уровнях цен 194
— — — числе уровней цен выше трех 197
— упорядочения 346
Молина уравнение 328

Наблюдение 182
Наклон матрицы 363
Наклонное решение 363

- Общая задача линейного программирования 251
— информационная модель 21
— эффективность 102, 112
Общее решение прямоугольной игры 407
Обязательства 420
Операционная игра 359
— модель 21
Операционный эксперимент 359
Определение Неймана — Моргенштерна 421
Оптимальная стратегия 131
Оптимальное решение 14
Основные уравнения цен 190
— фонды 199
Оценка выборок 444
— минимаксная 447
- План игры 400
Платежная матрица 399
Плотность вероятности 108
Поллачека — Кроммелина уравнение 328
Постоянные ресурсы 199
Принцип минимакса и максимина 401
Проблема двойственности 257
— оптимизации 341
— очередей 301, 309
Простая сортировка 71
Процессы замены 23
— комбинированные 24
— обслуживания 23
— распределения 23
— создания и хранения запасов 23
— состязательные 24
Прямоугольные игры 400
Пуассона линия 339
— точка 338
- Разбалансированная линия 360
Распределение Пуассона 326
Расходы на подготовку 34
— на производственные операции 35
— на сырье 35
— на хранение 35, 38
Расчет экономического размера партии 38
Решение игры 400
— оптимальное 14
— субоптимальное 14
- Сбор информации 81
— платы за проезд 320
Седловая точка 403
Символические модели 136
Симплексный метод 213, 231, 283
— — «модифицированный» 213
Система «человек — машина» 15, 18
Скорость замены 380, 383
Старение 45
Стоимость замены 387
Субоптимальное решение 14
- Теорема двойственности 214, 237
— Кёнига 266, 267
Теоремы прямоугольной игры 406
Теория балансирования линий 23
— очередей 23
— расписаний 23
— торгов 24, 427
Тождественная матрица 237
Точка Пуассона 338
Транспортная задача 265
Транспортный метод 217
- Упорядочение 345
—, модели упорядочения 346
Уравнение затрат 372
— Молина 328
— планирования 36
— Поллачека — Кроммелина 328
— цен 165, 169, 190
— Эрланга 327
- Функция оптимальная 109
— риска 458
— эффективности 108
- Цель 70
Цена игры 401
Циклограмма 346
- Численное решение 21
Чистая стратегия 400
- Экономически выгодный размер партии 200, 208
Эрланга уравнение 327

Уэст Черчмен, Рассел Акоф, Ленард Арноф
ВВЕДЕНИЕ В ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

М., 1968 г., 488 стр. с илл.

Редактор *Д. С. Фурманов*
Техн. редактор *С. Я. Шкляр*
Корректоры *С. Д. Кайсер* и *В. П. Сорокина*

Сдано в набор 11/VII 1967 г. Подписано к печати 21/XII 1
Бумага 70 × 108¹/₁₆. Физ. печ. л. 30,5.
Условн. печ. л. 42,70.
Уч.-изд. л. 38,72. Тираж 13000 экз.
Цена книги 2 р. 99 к. Заказ № 1177.

Издательство «Наука».
Главная редакция
физико-математической литературы.
Москва, В-71, Ленинский проспект, 15.

Московская типография № 16 Главполиграфпрома
Комитета по печати при Совете Министров СССР.
Москва, Трехпрудный пер., 9.

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
Москва, В-71, Ленинский проспект, 15

ИМЕЮТСЯ В ПРОДАЖЕ:

Березин И. С. и Жидков Н. П., Методы вычислений, том II, 1960, 620 стр., цена 1 р. 41 к.

Волынский Б. А. и Бухман В. Е., Модели для решения краевых задач, под ред. Л. А. Люстерника, 1960, 452 стр., цена 1 р. 38 к.

Джурин Э., Импульсные системы автоматического регулирования, перев. с англ., 1963, 456 стр., цена 1 р. 51 к.

Каган В. Ф., Субпроективные пространства, 1961, 218 стр., цена 68 коп.

Климов Г. П., Стохастические системы обслуживания, 1966, 244 стр., цена 70 коп.

Коуден Д., Статистические методы контроля качества, перев. с англ., 1961, 624 стр., цена 2 р. 88 к.

Об основаниях геометрии. Сборник классических работ по геометрии Лобачевского и развитию ее идей. К столетию со дня смерти Лобачевского, под ред. А. П. Нордена, 1956, 527 стр., цена 1 р. 87 к.

Проблемы кибернетики, вып. 9, 1963, 360 стр., цена 1 р. 59 к.

Проблемы кибернетики, вып. 13, 1965, 248 стр., цена 1 р. 43 к.

Проблемы кибернетики, вып. 15, 1965, 292 стр., цена 1 р. 06 к.

Проблемы кибернетики, вып. 16, 1966, 248 стр., цена 1 р. 48 к.

Проблемы кибернетики, вып. 17, 1966, 288 стр., цена 1 р. 54 к.

Перечисленные выше книги требуют в магазинах Книготорга. Письменный заказ можно направить также в ближайший отдел «Книга — почтой» республиканского, областного, краевого книготорга.

Литература будет выслана наложенным платежом. При отсутствии этих книг на месте следует обратиться по адресу: Москва, К-31, Петровка, 15, магазин № 8, отдел «Книга — почтой».



У. ЧЕРЧМЕН,
Р. АКОФ, ЛАРНОФ

ВВЕДЕНИЕ В ИССЛЕДОВАНИЕ ОПЕРАЦИЙ