В. В. Федосеев, А. Н. Гармаш, И. В. Орлова

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И ПРИКЛАДНЫЕ МОДЕЛИ

УЧЕБНИК ДЛЯ БАКАЛАВРОВ

3-е издание, переработанное и дополненное

Под редакцией **В. В. Федосеева**

Рекомендовано Министерством образования Российской Федерации в качестве учебника для студентов высших учебных заведений, обучающихся по направлениям подготовки «Экономика» и «Менеджмент»
Рекомендовано Учебно-методическим центром «Профессиональный учебник» в качестве учебника для студентов высших учебных заведений, обучающихся по направлениям подготовки «Экономика» и «Менеджмент»

Москва ▪ Юрайт ▪ 2013

УДК 51
ББК 22.1я73
Э40

***Авторы*:**

Федосеев Владилен Валентинович — кандидат экономических наук, заведующий кафедрой экономики и финансового права филиала РГСУ в г. Люберцы, руководитель авторского коллектива, редактор издания;

Гармаш Александр Николаевич — кандидат экономических наук, профессор кафедры экономико-математических методов и моделей Заочного финансово-экономического института Финансового университета при Правительстве РФ;

Орлова Ирина Владленовна — кандидат экономических наук, профессор кафедры экономико-математических методов и моделей Заочного финансово-экономического института Финансового университета при Правительстве РФ.

***Рецензенты*:**

кафедра экономических информационных систем и информационных технологий Московского государственного университета экономики, статистики и информатики (МЭСИ);

Ершов А. Т. — кандидат физико-математических наук, доцент кафедры прикладной математики Государственного университета управления.

Э40 **Экономико-математические методы и прикладные модели** : учебник для бакалавров / В. В. Федосеев, А. Н. Гармаш, И. В. Орлова ; под ред. В. В. Федосеева. — 3-е изд., перераб. и доп. — М. : Издательство Юрайт, 2013. — 328 с. — Серия : Бакалавр. Базовый курс.

ISBN 978-5-9916-2499-2

Изложена система экономико-математических и математико-статистических методов и моделей для решения широкого класса теоретических и прикладных задач анализа и прогнозирования социально-экономических процессов. Теоретическое рассмотрение указанных моделей сопровождается конкретными числовыми примерами. Приведены вопросы, задания и упражнения для контроля усвоения изучаемых тем.

Соответствует Федеральному государственному образовательному стандарту высшего профессионального образования третьего поколения.

Для студентов и аспирантов экономических направлений и специальностей, преподавателей дисциплин экономико-математического цикла, а также для практических работников в области финансово-экономической деятельности.

УДК 51
ББК 22.1я73

ISBN 978-5-9916-2499-2

Оглавление

Предисловие

Предлагаемый вниманию читателей учебник является третьим изданием учебного пособия с тем же названием. Он подготовлен в соответствии с программой дисциплины «Экономико-математические методы и прикладные модели», читаемой авторами во Всероссийском заочном финансово-экономическом институте (ВЗФЭИ), и требованиями Федерального государственного образовательного стандарта высшего профессионального образования третьего поколения для экономических специальностей при изучении дисциплин экономико-математического цикла. В результате освоения дисциплин этого цикла на основе предлагаемого учебника студенты должны обладать следующими компетенциями.

Знать:

- общие понятия и этапы математического моделирования социально-экономических систем и процессов;
- основы оптимального (математического) программирования;
- сущность методов математико-статистического анализа и прогнозирования экономической динамики;
- основные понятия балансовых методов в экономике;
- суть эконометрических методов анализа социально-экономических процессов;
- основные типовые модели макро- и микроэкономики.

Уметь:

- формулировать экономико-математические модели (ЭММ) реальных экономических процессов и задач;
- выбирать конкретное математическое обеспечение для рассматриваемых типов ЭММ;
- решать задачи на основе сформулированных моделей как аналитическими методами, так и с использованием ЭВМ;
- давать экономическую интерпретацию как параметров модели, так и полученных результатов решения.

Владеть:

- методами постановки экономических задач, включая задачи макроэкономики;
- методами решения оптимизационных задач, в том числе задач линейного программирования, балансовых и эконометрических задач, а также задач математико-статистического анализа экономических процессов;
- математическим аппаратом исследования широкого класса типовых и прикладных задач экономического анализа и принятия решений.

Круг вопросов, рассматриваемых в указанных требованиях Государственных образовательных стандартов для дисциплин экономико-математического цикла, определил структуру данной книги. В главе 1 «Основные понятия математического моделирования социально-экономических систем» раскрываются общие понятия системного анализа и моделирования систем и процессов в экономике, рассматривается сущность основных этапов экономико-математического моделирования, приводится краткая классификация экономико-математических методов и моделей.

В главе 2 «Основы линейного программирования» даются примеры экономических задач, которые в процессе экономико-математического моделирования сводятся к задачам линейного программирования. Приводятся основные сведения о математическом аппарате линейного программирования. Излагается геометрический метод решения простейших задач линейного программирования. Основное внимание уделено изложению алгоритмов симплексного метода решения задач линейного программирования, включая симплексный метод с искусственным базисом. Рассматриваемые методы и алгоритмы иллюстрируются на конкретных экономических задачах.

Глава 3 «Оптимальные экономико-математические модели» посвящена вопросам применения методов математического программирования для решения ряда оптимизационных экономических задач. В параграфе 3.1 рассматриваются вопросы применения теории двойственности линейного программирования для анализа оптимальных решений. В параграфе 3.2 изучаются специальные задачи линейного программирования на примере открытых и закрытых транспортных задач. Отдельные параграфы посвящены методам дискретного (целочисленного) программирования, задачам многокритериальной (векторной) оптимизации, основным понятиям нелинейного и динамического программирования, сетевым моделям управления. Приводится также ряд сведений о методах имитационного моделирования.

В главе 4 «Методы и модели анализа динамики экономических процессов» изучаются основные понятия временных рядов экономических показателей на примере одномерных временных рядов. Рассматриваются методы выявления и устранения аномальных наблюдений, методы определения наличия тренда во временных экономических рядах. Исследуются методы механического сглаживания рядов, включая метод экспоненциального сглаживания. Приводятся формулы и примеры расчета основных показателей динамики развития экономических систем. Особое внимание уделено анализу сезонности в экономических процессах, а также исследованию явления автокорреляции во временных экономических рядах.

В главе 5 «Модели прогнозирования экономических процессов» рассматриваются методологические вопросы экономического прогнозирования, в том числе такие принципы разработки прогнозов, как системность, адекватность и альтернативность. Исследуются проблемы экономического прогнозирования на основе принципов экстраполяции с использованием кривых роста; при этом анализируются основные типы кривых роста, методы выбора наилучших из них, описывается порядок определения параметров кривых роста на основе одномерных временных рядов экономических показателей. Особое внимание уделено оценке адекватности и точности трендовых моделей на основе кривых роста. Отдельный параграф посвящен вопросам составления точечных и интервальных прогнозов экономической динамики на базе рассматриваемых трендовых моделей. Приводятся также основные сведения об адаптивных методах и моделях прогнозирования.

Глава 6 «Балансовые модели» посвящена проблеме применения балансовых методов в экономико-математическом моделировании. Рассмотрены основные понятия балансового метода в экономических исследованиях, описана принципиальная схема межотраслевого баланса. Изучается экономико-математическая модель межотраслевого баланса в статической постановке, описывается порядок расчета на ее основе коэффициентов прямых, косвенных и полных материальных затрат. Приводятся примеры использования балансовых моделей для анализа экономических показателей, а также кратко обсуждаются вопросы разработки и применения динамических межотраслевых балансовых моделей.

В главе 7 «Эконометрические модели» рассмотрены общие понятия об эконометрических моделях, параметры которых оцениваются с помощью методов математической статистики. Изучаются такие наиболее распространенные эконометрические модели, как регрессионные факторные модели и производственные функции. Описан порядок решения основных задач регрессионного анализа (установление формы связи результативного признака с влияющими факторами, определение тесноты этой связи, анализ влияния отдельных факторов) на примере линейных моделей. Рассмотрены конкретные примеры решения этих задач с использованием линейных моделей парной и множественной регрессии.

Глава 8 «Некоторые прикладные и теоретические модели микро- и макроэкономических процессов» посвящена рассмотрению ряда прикладных задач маркетинга, менеджмента и других областей управления в экономике: моделирование спроса и потребления, научное управление запасами, аналитическое моделирование систем массового обслуживания, принятие решений на основе теории игр. В заключительном параграфе главы рассматриваются динамические модели макроэкономики на примере моделей Кейнса, Самуэльсона — Хикса и Солоу.

При подготовке третьего издания в учебник внесены следующие изменения и дополнения:

глава 3 — полностью переработан параграф 3.6, посвященный сетевым моделям планирования и управления;

глава 4 — переработан параграф 4.4, при этом детальное описание итерационных методов исследования сезонности, достаточно сложных для изучения и применения на практике, заменено описанием статистических методов анализа сезонных колебаний в экономических процессах, как без тренда, так и с наличием тренда;

глава 5 — при рассмотрении в параграфе 5.4 адаптивных методов прогнозирования иллюстрирующие теоретические выводы примеры заменены другими, более экономически содержательными;

глава 8 — расширен учебный материал главы путем добавления параграфа, связанного с изучением некоторых основных макроэкономических моделей, в связи с чем внесены корректировки в название главы.

Внесены также необходимые исправления и дополнения в материалы других глав.

Авторы надеются, что настоящее издание, как и предыдущие, окажется полезным большому кругу читателей. В то же время авторы будут весьма признательны за все замечания и предложения по структуре и содержанию этой книги, которые следует направлять в адрес издательства Юрайт.

Глава 1
ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ

— Социально-экономические системы, методы их исследования и моделирования

— Этапы экономико-математического моделирования

— Классификация экономико-математических методов и моделей

В результате изучения этой главы студенты должны:

***знать*:**
- особенности социально-экономических систем как объекта моделирования;
- этапы экономико-математического моделирования;
- основные подходы к классификации экономико-математических методов и моделей;

***уметь*:**
- формулировать такие понятия, как «метод» и «модель»;
- дать характеристику основных этапов экономико-математического моделирования;
- различать типы экономико-математических моделей по всем основным классификационным рубрикам;

***владеть*:**
- основными принципами системного подхода к анализу экономических процессов;
- общими понятиями математического моделирования в экономике.

1.1. Социально-экономические системы, методы их исследования и моделирования

Использование в данном учебнике термина «социально-экономическая система» требует, строго говоря, некоторого предварительного обсуждения. Если понятие «экономическая система» более или менее сложилось и в широком смысле трактуется как система общественного производства и потребления материальных благ, то социальные аспекты жизни общества весьма многогранны и не всегда доступны для детального анализа, моделирования и прогнозирования.

Вместе с тем некоторые социальные проблемы являются объектом исследования для практических работников тех специальностей финансово-экономического профиля, на которые ориентирован этот учебник. В качестве примера можно привести проблему анализа и прогнозирования покупательского спроса в маркетинге, задачу анализа распределения работников по уровню заработной платы в экономике и социологии труда и др. Многие из такого рода проблем могут быть решены с использованием экономико-математических методов и моделей, и именно такие проблемы имеют в виду авторы учебника, используя термин «социально-экономическая система».

Рассмотрим ряд основных понятий, связанных с системным анализом и моделированием социально-экономических систем, чтобы с их помощью более полно раскрыть суть такого ключевого понятия, как экономико-математические методы. Термин *экономико-математические методы* понимается в свою очередь как обобщающее название комплекса экономических и математических научных дисциплин, объединенных для изучения социально-экономических систем и процессов.

Под *социально-экономической системой* будем понимать сложную вероятностную динамическую систему, охватывающую процессы производства, обмена, распределения и потребления материальных и других благ. Она относится к классу кибернетических систем, т.е. систем управляемых. Рассмотрим, прежде всего, понятия, связанные с такими системами и методами их исследования.

Центральным понятием кибернетики является понятие «система». Единого определения этого понятия нет; возможна такая формулировка: *системой* называется комплекс взаимосвязанных элементов вместе с отношениями между элементами и между их атрибутами. Исследуемое множество элементов можно рассматривать как систему, если выявлены следующие четыре признака:

— целостность системы, т.е. принципиальная несводимость свойств системы к сумме свойств составляющих ее элементов;

— наличие цели и критерия исследования данного множества элементов;

— наличие более крупной, внешней по отношению к данной, системы, называемой «средой»;

— возможность выделения в данной системе взаимосвязанных частей (подсистем).

Основным методом исследования систем является *метод моделирования*, т.е. способ теоретического анализа и практического действия, направленный на разработку и использование моделей. При этом под *моделью* будем понимать образ реального объекта (процесса) в материальной или идеальной форме (т.е. описанный знаковыми средствами на каком-либо языке), отражающий существенные свойства моделируемого объекта (процесса) и замещающий его в ходе исследования и управления. Метод моделирования основывается на принципе аналогии, т.е. возможности изучения реального объекта не непосредственно, а через рассмотрение подобного ему и более доступного объекта, его модели. В дальнейшем мы будем говорить только об экономико-математическом моделировании, т.е. об описании знаковыми математическими средствами социально-экономических систем.

Практическими задачами экономико-математического моделирования являются:

— анализ экономических объектов и процессов;

— экономическое прогнозирование, предвидение развития экономических процессов;

— выработка управленческих решений на всех уровнях хозяйственной иерархии.

Следует, однако, иметь в виду, что далеко не во всех случаях данные, полученные в результате экономико-математического моделирования, могут использоваться непосредственно как готовые управленческие решения. Они скорее могут быть рассмотрены как «консультирующие» средства. Принятие управленческих решений остается за человеком. Таким образом, экономико-математическое моделирование является лишь одним из компонентов (пусть очень важным) в человеко-машинных системах планирования и управления экономическими системами.

Важнейшим понятием при экономико-математическом моделировании, как и при всяком моделировании, является понятие *адекватности модели*, т.е. соответствия модели моделируемому объекту или процессу. Адекватность модели — в какой-то мере условное понятие, так как полного соответствия модели реальному объекту быть не может, что характерно и для экономико-математического моделирования. При моделировании имеется в виду не просто адекватность, но соответствие модели тем свойствам, которые считаются существенными для исследования. Проверка адекватности экономико-математических моделей является весьма серьезной проблемой, тем более что ее осложняет трудность измерения экономических величин. Однако без такой проверки применение результатов моделирования в управленческих решениях может не только оказаться мало полезным, но и принести существенный вред.

Социально-экономические системы относятся, как правило, к так называемым *сложным системам*. Сложные системы в экономике обладают рядом свойств, которые необходимо учитывать при их моделировании, иначе невозможно говорить об адекватности построенной экономической модели. Важнейшие из этих свойств:

— эмерджентность как проявление в наиболее яркой форме свойства целостности системы, т.е. наличие у экономической системы таких свойств, которые не присущи ни одному из составляющих систему элементов, взятому в отдельности, вне системы. Эмерджентность есть результат возникновения между элементами системы так называемых синергических связей, которые обеспечивают увеличение общего эффекта до величины большей, чем сумма эффектов элементов системы, действующих независимо. Поэтому социально-экономические системы необходимо исследовать и моделировать в целом;

— массовый характер экономических явлений и процессов. Закономерности экономических процессов не обнаруживаются на основании небольшого числа наблюдений. Поэтому моделирование в экономике должно опираться на массовые наблюдения;

— динамичность экономических процессов, заключающаяся в изменении параметров и структуры экономических систем под влиянием среды (внешних факторов);

— случайность и неопределенность в развитии экономических явлений. Поэтому экономические явления и процессы носят в основном вероятностный характер, и для их изучения необходимо применение экономико-математических моделей на базе теории вероятностей и математической статистики;

— невозможность изолировать протекающие в экономических системах явления и процессы от окружающей среды, чтобы наблюдать и исследовать их в чистом виде;

— активная реакция на появляющиеся новые факторы, способность социально-экономических систем к активным, не всегда предсказуемым действиям в зависимости от отношения системы к этим факторам, способам и методам их воздействия.

Выделенные свойства социально-экономических систем, естественно, осложняют процесс их моделирования, однако эти свойства следует постоянно иметь в виду при рассмотрении различных аспектов экономико-математического моделирования, начиная с выбора типа модели и кончая вопросами практического использования результатов моделирования.

1.2. Этапы экономико-математического моделирования

Процесс моделирования, в том числе и экономико-математического, включает в себя три структурных элемента: объект исследования; субъект (исследователь); модель, опосредующая отношения между познающим субъектом и познаваемым объектом. Рассмотрим общую схему процесса моделирования, в котором можно выделить четыре стадии.

Пусть имеется некоторый объект, который мы хотим исследовать методом моделирования. На *первой стадии* мы конструируем (или находим в реальном мире) другой объект — модель исходного объекта-оригинала. Стадия построения модели предполагает наличие определенных сведений об объекте-оригинале. Познавательные возможности модели определяются тем, что модель отображает лишь некоторые существенные черты исходного объекта, поэтому любая модель замещает оригинал в строго ограниченном смысле. Из этого следует, что для одного объекта может быть построено несколько моделей, отражающих определенные стороны исследуемого объекта или характеризующих его с разной степенью детализации.

На *второй стадии* процесса моделирования модель выступает как самостоятельный объект исследования. Например, одну из форм такого исследования составляет проведение модельных экспериментов, при которых целенаправленно изменяются условия функционирования модели и систематизируются данные о ее «поведении». Конечным результатом этой стадии является совокупность знаний о модели в отношении существенных сторон объекта-оригинала, которые отражены в данной модели.

Третья стадия заключается в переносе знаний с модели на оригинал, в результате чего мы формируем множество знаний об исходном объекте и при этом переходим с языка модели на язык оригинала. С достаточным основанием переносить какой-либо результат с модели на оригинал можно лишь в том случае, если этот результат соответствует признакам сходства оригинала и модели (другими словами, признакам *адекватности*).

На *четвертой стадии* осуществляются практическая проверка полученных с помощью модели знаний и их использование как для построения обобщающей теории реального объекта, так и для его целенаправленного преобразования или управления им. В итоге мы снова возвращаемся к проблематике объекта-оригинала.

Моделирование представляет собой циклический процесс, т.е. за первым четырехстадийным циклом может последовать второй, третий и т.д. При этом знания об исследуемом объекте расширяются и уточняются, а первоначально построенная модель постепенно совершенствуется. Таким образом, в методологии моделирования заложены большие возможности самосовершенствования.

Перейдем теперь непосредственно к процессу экономико-математического моделирования, т.е. к описанию экономических и социальных систем и процессов в виде экономико-математических моделей. Эта разновидность моделирования обладает рядом существенных особенностей, связанных как с объектом моделирования, так и с применяемыми аппаратом и средствами моделирования. Поэтому целесообразно более детально проанализировать последовательность и содержание этапов экономико-математического моделирования, выделив следующие шесть этапов: постановка экономической проблемы и ее качественный анализ; построение математической модели; математический анализ модели; подготовка исходной информации; численное решение; анализ численных результатов и их применение. Рассмотрим каждый из этапов более подробно.

1. Постановка экономической проблемы и ее качественный анализ. На этом этапе требуется сформулировать сущность проблемы, принимаемые предпосылки и допущения. Необходимо выделить важнейшие черты и свойства моделируемого объекта, изучить его структуру и взаимосвязь его элементов, хотя бы предварительно сформулировать гипотезы, объясняющие поведение и развитие объекта.

2. Построение математической модели. Это этап формализации экономической проблемы, т.е. выражения ее в виде конкретных математических зависимостей (функций, уравнений, неравенств и др.). Построение модели подразделяется в свою очередь на несколько стадий. Сначала определяется тип экономико-математической модели, изучаются возможности ее применения в данной задаче, уточняются конкретный перечень переменных и параметров и форма связей. Для некоторых сложных объектов целесообразно строить несколько разноаспектных моделей; при этом каждая модель выделяет лишь некоторые стороны объекта, а другие стороны учитываются агрегированно и приближенно. Оправданно стремление построить модель, относящуюся к хорошо изученному классу математических задач, что может потребовать некоторого упрощения исходных предпосылок модели, не искажающего основных черт моделируемого объекта. Однако возможна и такая ситуация, когда формализация проблемы приводит к неизвестной ранее математической структуре.

3. Математический анализ модели. На этом этапе чисто математическими приемами исследования выявляются общие свойства модели и ее решений. В частности, важным моментом является доказательство существования решения сформулированной задачи. При аналитическом исследовании выясняется, единственно ли решение, какие переменные могут входить в решение, в каких пределах они изменяются, каковы тенденции их изменения и т.д. Однако модели сложных экономических объектов с большим трудом поддаются аналитическому исследованию; в таких случаях переходят к численным методам исследования.

4. Подготовка исходной информации. В экономических задачах это, как правило, наиболее трудоемкий этап моделирования, так как дело не сводится к пассивному сбору данных. Математическое моделирование предъявляет жесткие требования к системе информации; при этом надо принимать во внимание не только принципиальную возможность подготовки информации требуемого качества, но и затраты на подготовку информационных массивов. В процессе подготовки информации используются методы теории вероятностей, теоретической и математической статистики для организации выборочных обследований, оценки достоверности данных и т.д. При системном экономико-математическом моделировании результаты функционирования одних моделей служат исходной информацией для других.

5. Численное решение. Этот этап включает разработку алгоритмов численного решения задачи, подготовку программ на ЭВМ и непосредственное проведение расчетов; при этом значительные трудности вызываются большой размерностью экономических задач. Обычно расчеты на основе экономико-математической модели носят многовариантный характер. Многочисленные модельные эксперименты, изучение поведения модели при различных условиях стало возможно проводить благодаря высокому быстродействию современных ЭВМ. Численное решение существенно дополняет результаты аналитического исследования, а для многих моделей является единственно возможным.

6. Анализ численных результатов и их применение. На этом этапе, прежде всего, решается важнейший вопрос о правильности и полноте результатов моделирования и применимости их как в практической деятельности, так и в целях усовершенствования модели. Поэтому в первую очередь должна быть проведена проверка адекватности модели по тем свойствам, которые выбраны в качестве существенных (другими словами, должны быть произведены верификация и валидация модели)[1]. Применение численных результатов моделирования в экономике направлено на решение практических задач (анализ экономических объектов, экономическое прогнозирование развития хозяйственных и социальных процессов, выработка управленческих решений на всех уровнях хозяйственной иерархии).

Перечисленные этапы экономико-математического моделирования находятся в тесной взаимосвязи, в частности, могут иметь место возвратные связи этапов. Так, на этапе построения модели может выясниться, что постановка за-

[1] *Верификация* модели — проверка правильности структуры (логики) модели; *валидация* модели — проверка соответствия данных, полученных на основе модели, реальному процессу.

дачи или противоречива, или приводит к слишком сложной математической модели; в этом случае исходная постановка задачи должна быть скорректирована. Наиболее часто необходимость возврата к предшествующим этапам моделирования возникает на этапе подготовки исходной информации. Если необходимая информация отсутствует или затраты на ее подготовку слишком велики, приходится возвращаться к этапам постановки задачи и ее формализации, чтобы приспособиться к доступной исследователю информации.

Выше уже было сказано о циклическом характере процесса моделирования. Недостатки, которые не удается исправить на тех или иных этапах моделирования, устраняются в последующих циклах. Однако результаты каждого цикла имеют и вполне самостоятельное значение. Начав исследование с построения простой модели, можно получить полезные результаты, а затем перейти к созданию более сложной и более совершенной модели, включающей в себя новые условия и более точные математические зависимости.

1.3. Классификация экономико-математических методов и моделей

Суть экономико-математического моделирования заключается в описании социально-экономических систем и процессов в виде экономико-математических моделей. В параграфе 1.1 кратко рассмотрен смысл понятий «метод моделирования» и «модель». Исходя из этого, экономико-математические методы следует понимать как инструмент, а экономико-математические модели — как продукт процесса экономико-математического моделирования.

Рассмотрим вопросы классификации экономико-математических методов. Эти методы, как отмечено выше, представляют собой комплекс экономико-математических дисциплин, являющихся сплавом экономики, математики и кибернетики. Поэтому классификация экономико-математических методов сводится к классификации научных дисциплин, входящих в их состав. Хотя общепринятая классификация этих дисциплин пока не выработана, с известной степенью приближения в составе экономико-математических методов можно выделить следующие разделы:

— *экономическая кибернетика*: системный анализ экономики, теория экономической информации и теория управляющих систем;

— *математическая статистика*: экономические приложения данной дисциплины — выборочный метод, дисперсионный анализ, корреляционный анализ, регрессионный анализ, многомерный статистический анализ, факторный анализ, теория индексов и др.;

— *математическая экономика* и изучающая те же вопросы с количественной стороны *эконометрика*: теория экономического роста, теория производственных функций, межотраслевые балансы, национальные счета, анализ спроса и потребления, региональный и пространственный анализ, глобальное моделирование и др.;

— *методы принятия оптимальных решений*, в том числе исследование операций в экономике. Это наиболее объемный раздел, включающий в себя следующие дисциплины и методы: оптимальное (математическое) программирование, в том числе методы ветвей и границ, сетевые методы планирования и управления, программно-целевые методы планирования и управления, теория и методы управления запасами, теория массового обслуживания, теория игр, теория и методы принятия решений, теория расписаний. В оптимальное (математическое) программирование входят в свою очередь: линейное программирование, нелинейное программирование, динамическое программирование, дискретное (целочисленное) программирование, дробно-линейное программирование, параметрическое программирование, сепарабельное программирование, стохастическое программирование, геометрическое программирование;

— *методы и дисциплины, специфичные отдельно как для централизованно планируемой экономики, так и для рыночной (конкурентной) экономики*. К первым можно отнести теорию оптимального функционирования экономики, оптимальное планирование, теорию оптимального ценообразования, модели материально-технического снабжения и др. Ко вторым — методы, позволяющие разработать модели свободной конкуренции, модели капиталистического цикла, модели монополии, модели индикативного планирования, модели теории фирмы и т.д. Многие из методов, разработанных для централизованно планируемой экономики, могут оказаться полезными и при экономико-математическом моделировании в условиях рыночной экономики;

— *методы экспериментального изучения экономических явлений*. К ним относят, как правило, математические методы анализа и планирования экономических экспериментов, методы машинной имитации (имитационное моделирование), деловые игры. Сюда можно отнести также и методы экспертных оценок, разработанные для оценки явлений, не поддающихся непосредственному измерению.

Перейдем теперь к вопросам классификации экономико-математических моделей, другими словами, математических моделей социально-экономических систем и процессов. Единой системы классификации таких моделей в настоящее время также не существует, однако обычно выделяют более 10 основных признаков их классификации, или классификационных рубрик. Рассмотрим некоторые из этих рубрик.

По общему целевому назначению экономико-математические модели делятся на *теоретико-аналитические*, используемые при изучении общих свойств и закономерностей экономических процессов, и *прикладные*, применяемые в решении конкретных экономических задач анализа, прогнозирования и управления. Различные типы прикладных экономико-математических моделей как раз и рассматриваются в данном учебнике.

По степени агрегирования объектов моделирования модели разделяются на *макроэкономические* и *микроэкономические*. Хотя между ними и нет четкого разграничения, к первым из них относят модели, отражающие функционирование экономики как единого целого, в то время как микроэкономические модели связаны, как правило, с такими звеньями экономики, как предприятия и фирмы.

По конкретному предназначению, т.е. по цели создания и применения, выделяют *балансовые* модели, выражающие требование соответствия наличия ресурсов и их использования; *трендовые* модели, в которых развитие моделируемой экономической системы отражается через тренд (длительную тенденцию) ее основных показателей; *оптимизационные* модели, предназначенные для выбора наилучшего варианта из определенного числа вариантов производства, распределения или потребления; *имитационные* модели, предназначенные для использования в процессе машинной имитации изучаемых систем или процессов, и др.

По типу информации, используемой в модели, экономико-математические модели делятся на *аналитические*, построенные на априорной информации, и *идентифицируемые*, построенные на апостериорной информации.

По учету фактора времени модели подразделяются на *статические*, в которых все зависимости отнесены к одному моменту времени, и *динамические*, описывающие экономические системы в развитии.

По учету фактора неопределенности модели распадаются на *детерминированные*, если в них результаты на выходе однозначно определяются управляющими воздействиями, и *стохастические* (*вероятностные*), если при задании на входе модели определенной совокупности значений на ее выходе могут получаться различные результаты в зависимости от действия случайного фактора.

Экономико-математические модели могут классифицироваться также **по характеристике математических объектов**, включенных в модель, другими словами, **по типу математического аппарата**, используемого в модели. По этому признаку могут быть выделены *матричные* модели, модели *линейного* и *нелинейного программирования*, *корреляционно-регрессионные* модели, модели *теории массового обслуживания*, модели *сетевого планирования и управления*, модели *теории игр* и т.д.

Наконец, **по типу подхода к изучаемым социально-экономическим системам** выделяют *дескриптивные* и *нормативные* модели. При дескриптивном (описательном) подходе получаются модели, предназначенные для описания и объяснения фактически наблюдаемых явлений или для прогноза этих явлений; в качестве примера дескриптивных моделей можно привести названные ранее балансовые и трендовые модели. При нормативном подходе интересуются не тем, каким образом устроена и развивается экономическая система, а как она должна быть устроена и как должна действовать в смысле определенных критериев. В частности, все оптимизационные модели относятся к типу нормативных; другим примером могут служить нормативные модели уровня жизни.

Рассмотрим в качестве примера экономико-математическую модель межотраслевого баланса (ЭММ МОБ). С учетом приведенных выше классификационных рубрик это прикладная, макроэкономическая, аналитическая, дескриптивная, детерминированная, балансовая, матричная модель; при этом существуют как статические, так и динамические ЭММ МОБ.

Вопросы и задания

1. В чем заключается смысл системного подхода к анализу социально-экономических систем и процессов?
2. Сформулируйте понятия «модель» и «метод моделирования».
3. Каковы важнейшие особенности социально-экономических систем как объектов моделирования?
4. Дайте характеристику этапов экономико-математического моделирования.
5. Укажите основные научные дисциплины и методы, входящие в состав экономико-математических методов.
6. Назовите основные классификационные признаки экономико-математических моделей и приведите примеры моделей, входящих в ту или иную классификационную рубрику.

Глава 2
ОСНОВЫ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

— Принцип оптимальности в планировании и управлении, общая задача оптимального программирования
— Формы записи задачи линейного программирования и ее экономическая интерпретация
— Математический аппарат линейного программирования
— Геометрическая интерпретация задачи
— Симплексный метод решения задачи

После изучения данной главы студенты должны:

***знать*:**
- общие принципы оптимального планирования и управления в экономике;
- основные понятия линейного программирования;
- методы решения задач линейного программирования (ЗЛП);

***уметь*:**
- формулировать общую постановку ЗЛП;
- представить ЗЛП в различных формах записи;
- дать экономическую интерпретацию полученных результатов на всех этапах графического и симплексного методов решения ЗЛП;

***владеть*:**
- математическим аппаратом линейного программирования;
- практическими навыками формулирования и решения ЗЛП.

2.1. Принцип оптимальности в планировании и управлении, общая задача оптимального программирования

Линейное программирование — это особый раздел оптимального программирования. В свою очередь *оптимальное (математическое) программирование* — раздел прикладной математики, изучающий задачи условной оптимизации. В экономике такие задачи возникают при практической реализации принципа оптимальности в планировании и управлении.

Необходимым условием использования оптимального подхода к планированию и управлению (принципа оптимальности) является гибкость, альтернативность производственно-хозяйственных ситуаций, в условиях которых приходится принимать планово-управленческие решения. Именно такие ситуации, как правило, и составляют повседневную практику хозяйствующего субъекта (выбор производственной программы, прикрепление к поставщикам, маршрутизация, раскрой материалов, приготовление смесей, загрузка контейнеров и т.д.).

Суть принципа оптимальности состоит в стремлении выбрать такое планово-управленческое решение $\bar{X}=(x_1, x_2, ..., x_n)$, где x_j (j = 1, ..., n) — его компоненты, которое наилучшим образом учитывало бы внутренние возможности и внешние условия производственной деятельности хозяйствующего субъекта.

Слова «наилучшим образом» здесь означают выбор некоторого критерия оптимальности, т.е. некоторого экономического показателя, позволяющего сравнивать эффективность тех или иных планово-управленческих решений. Традиционные критерии оптимальности: «максимум прибыли», «минимум затрат», «максимум объема работ (услуг)» и др.

Слова «учитывало бы внутренние возможности и внешние условия производственной деятельности» означают, что на выбор управленческого решения (поведения) накладывается ряд условий, т.е. выбор $\bar{X}$ осуществляется из некоторой области возможных (допустимых) решений D; эту область называют также областью определения задачи.

Таким образом, реализовать на практике принцип оптимальности в планировании и управлении — это значит решить экстремальную задачу вида

$$\max(\min) f(\bar{X}) \qquad (2.1)$$

$$\bar{X} \in D, \qquad (2.2)$$

где $f(\bar{X})$ — математическая запись критерия оптимальности — целевая функция задачи (модели) оптимизации.

Задачу условной оптимизации (2.1), (2.2) обычно записывают в виде:

найти максимум или минимум функции

$$f(\bar{X}) = f(x_1, x_2, ..., x_n) \qquad (2.3)$$

при ограничениях

$$\varphi_1(x_1, x_2, ..., x_n)\ \{\le, =, \ge\}\ b_1,$$
$$\varphi_2(x_1, x_2, ..., x_n)\ \{\le, =, \ge\}\ b_2, \qquad (2.4)$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$\varphi_m(x_1, x_2, ..., x_n)\ \{\le, =, \ge\}\ b_m,$$
$$x_j \ge 0, j = 1, 2, ..., n. \qquad (2.5)$$

Условие (2.5) необязательно, но его всегда при необходимости можно добиться. Обозначение $\{\le, =, \ge\}$ говорит о том, что в конкретном ограничении возможен один из знаков $\le, =$ или $\ge$. Более компактная запись:

$$\max(\min) f(x_1, x_2, ..., x_n) \qquad (2.6)$$
$$\varphi_i(x_1, x_2, ..., x_n)\ \{\le, =, \ge\}\ b_i, i = 1, 2, ..., m, \qquad (2.7)$$
$$x_j \ge 0, j = 1, 2, ..., n. \qquad (2.8)$$

Задача (2.6)—(2.8) — общая задача оптимального (математического) программирования, иначе — математическая модель задачи оптимального программирования, в основе построения (разработки) которой лежат принципы оптимальности, системности и адекватности.

Вектор $\bar{X}$ (набор управляющих переменных x_j, $j = 1, 2, ..., n$) называется допустимым решением или планом задачи оптимального программирования, если его компоненты x_j удовлетворяют системе ограничений. А тот план $\bar{X}$ (допустимое решение), который доставляет максимум или минимум целевой функции $f(x_1, x_2, ..., x_n)$ называется оптимальным планом (оптимальным поведением, или просто решением) задачи оптимального программирования.

Таким образом, выбор оптимального управленческого поведения в конкретной производственной ситуации связан с проведением с позиций системности, адекватности и оптимальности экономико-математического моделирования и решением задачи оптимального программирования.

Решить задачу оптимального программирования (получить решение по оптимизационной экономико-математической модели) — это значит:

— во-первых, найти оптимальный план $\bar{X}^* = (x_1^*, x_2^*, ..., x_n^*)$, который, с учетом интерпретации его компонент, и определяет оптимальное поведение в рассматриваемой ситуации;

— во-вторых, найти оптимальное значение (максимум или минимум) целевой функции $f(\bar{X}^*) = (x_1^*, x_2^*, ..., x_n^*)$, которое и представляет собой экономическую оценку последствий предлагаемого решения (поведения).

Иногда невозможно получить решение по оптимизационной модели: область допустимых решений может оказаться пустым множеством (задача противоречива) или целевая функция является неограниченной на области определения.

Первый случай связан с некорректностями в постановке экономической задачи или (и) разработанной экономико-математической модели (ЭММ). Например, с имеющимся объемом ресурсов заведомо невозможно выполнить даже те минимальные объемы работ, которые закладываются в ограничения как необходимые минимальные плановые задания. Если в данной ситуации все же необходимо найти решение задачи, то следует построить непустое множество допустимых решений, исключив одно или несколько ограничений, т.е. фактически соблюсти принцип альтернативности.

Второй случай обычно означает, что ЭММ разработана некорректно и некоторые существенные ограничения в ней отсутствуют.

Задачи оптимального программирования *в наиболее общем виде* классифицируют по следующим признакам.

1. По характеру взаимосвязи между переменными:

а) линейные;

б) нелинейные.

В случае (а) все функциональные связи в системе ограничений и функция цели — линейные функции, наличие нелинейности в хотя бы одном из упомянутых элементов приводит к случаю (б).

2. По характеру изменения переменных:

а) непрерывные;

б) дискретные.

В случае (а) значения каждой из управляющих переменных могут заполнять сплошь некоторую область действительных чисел, в случае (б) все или хотя бы одна переменная могут принимать только целочисленные значения.

3. По учету фактора времени:

а) статические;

б) динамические.

В задачах (а) моделирование и принятие решений осуществляются в предположении о независимости от времени элементов модели в течение периода времени, на который принимается планово-управленческое решение. В случае (б) такое предположение достаточно аргументированно принято не может быть и необходимо учитывать фактор времени.

4. По наличию информации о переменных:

а) задачи в условиях полной определенности (детерминированные);

б) задачи в условиях неполной информации;

в) задачи в условиях неопределенности.

В задачах (б) отдельные элементы являются вероятностными величинами, однако известны или дополнительными статистическими исследованиями могут быть установлены их законы распределения вероятностей. В случае (в) можно сделать предположение о возможных исходах случайных элементов, но нет возможности сделать вывод о вероятностях исходов.

5. По числу критериев оценки альтернатив:

а) простые, однокритериальные задачи;

б) сложные, многокритериальные задачи.

В задачах (а) экономически приемлемо использование одного критерия оптимальности или удается специальными процедурами (например, «взвешиванием приоритетов») свести многокритериальный поиск к однокритериальному; примеры многокритериальных задач рассмотрены в главе 3.

Сочетание признаков 1—5 позволяет группировать (классифицировать) в самом общем виде задачи и методы оптимального программирования, например: 1а)2а)3а)4а)5а) — задачи и методы линейного программирования, 1б)2а)3а)4а)5а) — задачи и методы нелинейного программирования, 1а)2б)3а)4а)5а) — задачи и методы целочисленного (дискретного) линейного программирования и т.д.

Новые возможности для широкого практического применения методов оптимального программирования представляют современные офисные средства. Широкий круг специалистов в своей повседневной практике использует необходимый компонент финансово-экономических расчетов — Microsoft Excel (MS Excel), который содержит специальное средство — надстройку **Поиск решения**, позволяющую реализовывать модели линейной, нелинейной и дискретной оптимизации. Технология оптимизации с помощью надстройки **Поиск решения** с решением некоторых типовых задач оптимального программирования в среде MS Excel подробно рассмотрена, например, в литературе[1,2].

Рассмотрим **пример задачи оптимального программирования**.

Постановка задачи. Предлагается n инвестиционных проектов, тщательная экономическая экспертиза которых позволяет получить для каждого из проектов достаточно убедительные экономические оценки ожидаемого эффекта от их реализации $c_1, c_2, ..., c_n$ и необходимых капиталовложений $p_1, p_2, ..., p_n$. Общий объем возможных капиталовложений ограничен величиной B. Необходимо так распорядиться имеющимися финансовыми ресурсами, чтобы максимизировать суммарный эффект от инвестиций.

Математическая модель. Введем необходимые обозначения, пусть x_j ($j = 1, 2, ..., n$):

$$x_j = \begin{cases} 1, \text{если } j\text{-й проект следует финансировать,} \\ 0, \text{в противном случае.} \end{cases}$$

Таким образом, формально инвестиционный план — это вектор $\bar{X} = (x_1, x_2, ..., x_n)$. С учетом этих обозначений задача по критерию «максимум экономического эффекта» математически запишется следующим образом:

$$\max f(\bar{X}) = \sum_{j=1}^{n} c_j x_j$$

$$\sum_{j=1}^{n} p_j x_j \le B,$$

$$x_j = \begin{cases} 1 \\ 0 \end{cases}, \; j = 1, 2, ..., n.$$

Приведенная задача (модель) является задачей дискретного линейного программирования с булевыми переменными (переменные, которые могут принимать только два значения: 1 и 0, иначе «да» или «нет»), т.е. относится к классу задач 1а)2б)3а)4а)5а). Эта задача может быть решена, например, известным методом Балаша.

---

1 *Мур Дж., Уэдерфорд Л.* Экономическое моделирование в Microsoft Excel. М. : Издательский дом «Вильямс», 2004.

2 *Гармаш А. Н., Орлова И. В.* Математические методы в управлении: учеб. пособие. М. : Вузовский учебник, 2012.

Выбору метода решения конкретной задачи оптимального программирования предшествует ее классификация, т.е. отнесение к одному из классов оптимизационных задач, начиная с приведенных самых общих признаков (например, задача дискретного линейного программирования с булевыми переменными).

Развитие и совершенствование методов решения задач оптимального программирования идет от случаев типа (а) к случаям типа (б), (в).

Наиболее изучены задачи линейного программирования, для которых разработан универсальный метод решения — метод последовательного улучшения плана (симплекс-метод), т.е. любая задача линейного программирования решается (о модели линейной оптимизации говорят, что она реализуется) этим методом.

Именно эти задачи в дальнейшем рассматриваются в данной главе.

2.2. Формы записи задачи линейного программирования и ее экономическая интерпретация

Как отмечено выше, среди широкого класса задач оптимального программирования имеются важные подклассы задач, для которых разработаны эффективные методы решения. Наиболее изученным подклассом задач являются *задачи линейного программирования.*

В задаче линейного программирования (ЗЛП) требуется найти экстремум (максимум или минимум) линейной целевой функции $f(\bar{X})$:

$$\max(\min) f(\bar{X}) = c_1 x_1 + c_2 x_2 + \dots c_n x_n \tag{2.9}$$

при ограничениях (условиях):

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \{\le, =, \ge\} b_1, \\ &a_{21}x_1 + a_{21}x_2 + \dots + a_{2n}x_n \{\le, =, \ge\} b_2, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \end{aligned} \tag{2.10}$$

$$\begin{aligned} &a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \{\le, =, \ge\} b_m, \\ &x_j \ge 0, j = 1, 2, \dots, n. \end{aligned} \tag{2.11}$$

где a_{ij}, b_i, c_j ($i = 1, 2, ..., m, j = 1, 2, ..., n$) — заданные постоянные величины. Так записывается общая задача линейного программирования в развернутой форме. Знак $\{\le, =, \ge\}$ означает, что в конкретной ЗЛП возможно ограничение типа равенства или неравенства (в ту или иную сторону).

Систему ограничений (2.10) называют *функциональными ограничениями* ЗЛП, а ограничения (2.11) — *прямыми*.

Вектор $\overline{X} = (x_1, x_2, ..., x_n)$, компоненты которого x_j удовлетворяют системе ограничений (2.10) и (2.11), называется *допустимым решением* или *планом* ЗЛП, т.е. ограничения (2.10), (2.11) определяют *область допустимых решений* или *планов* задачи линейного программирования (*область определения* ЗЛП).

План (допустимое решение), который доставляет максимум или минимум целевой функции (2.9), называется *оптимальным планом* (*оптимальным решением*) ЗЛП.

Канонической формой записи задачи линейного программирования (КЗЛП) называют задачу вида (запись с использованием знаков суммирования):

$$\max(\min) f(\overline{X}) = \sum_{j=1}^{n} c_j x_j \tag{2.12}$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i,\ i = 1, 2, ..., m; \tag{2.13}$$

$$x_j \ge 0,\ j = 1, 2, ..., n;\ b_i \ge 0, i = 1, 2, ..., m. \tag{2.14}$$

Векторная форма записи КЗЛП имеет вид

$$\begin{aligned}
&\max(\min) f(\overline{X}) = CX\\
&A_1 x_1 + A_2 x_2 + ... + A_n x_n = B,\\
&X \ge 0, B \ge 0,
\end{aligned}$$

где CX — скалярное произведение векторов $C = (c_1, c_2, ..., c_n)$, $X = (x_1, x_2, ..., x_n)$; A_j и B — вектор-столбцы:

$$A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ ... \\ a_{m1} \end{pmatrix}, A_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ ... \\ a_{m2} \end{pmatrix}, ..., A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ ... \\ a_{mn} \end{pmatrix}; B = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ b_m \end{pmatrix}.$$

Матричная форма записи КЗЛП:

$$\max(\min) f(\bar{X}) = CX,$$
$$AX = B,$$
$$X \geq 0,\ B \geq 0$$

Здесь $C = (c_1, c_2, ..., c_n)$ — вектор-строка; $A = (a_{ij})_{m \times n}$ — матрица размерности $m \times n$, столбцами которой являются вектор-столбцы A_j; X и B — вектор-столбцы:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix};\ B = \begin{pmatrix} b_1 \\ b_2 \\ ... \\ b_m \end{pmatrix}.$$

Иногда используется стандартная форма записи ЗЛП:

$$\max(\min) f(\bar{X}) = \sum_{j=1}^{n} c_j x_j$$
$$\sum_{j=1}^{n} a_{ij} x_j \leq (\geq)\ b_i,\ i = 1, 2, ..., m;$$
$$x_j \geq 0,\ j = 1, 2, ..., n.$$

Имеет место утверждение, что любую ЗЛП можно привести к каноническому виду.

Приведение ЗЛП к каноническому виду осуществляется введением в левую часть соответствующего ограничения вида (2.10) k-й дополнительной (вспомогательной) переменной $x_{n+k} \geq 0$ со знаком «−» в случае ограничения типа «≥» и знаком «+» в случае ограничения типа «≤».

Если на некоторую переменную x_r не накладывается условие неотрицательности, то делают замену переменных $x_r = x_r^1 - x_r^2$, $x_r^1 \geq 0$, $x_r^2 \geq 0$. В преобразованной задаче все переменные неотрицательные.

Переход в случае необходимости к задаче на максимум достигается изменением знака у целевой функции.

К математическим задачам линейного программирования приводят исследования конкретных производственно-хозяйственных ситуаций, которые в том или ином виде интерпретируются как задачи об оптимальном использовании ограниченных ресурсов (задача о раскрое, смесях, рационе и т.д.).

Пример 2.1 (задача о смесях). Стандартом предусмотрено, что октановое число автомобильного бензина А-76 должно быть не ниже 76, а содержание серы в нем — не более 0,3%. Для изготовления такого бензина на заводе используется смесь из четырех компонентов. Данные о ресурсах смешиваемых компонентов, их себестоимости и их октановом числе, а также о содержании серы приведены в таблице[1]:

Характеристика	Компонент автомобильного бензина			
	№ 1	№ 2	№ 3	№ 4
Октановое число	68	72	80	90
Содержание серы, %	0,35	0,35	0,3	0,2
Ресурсы, т	700	600	500	300
Себестоимость, ден. ед./т	40	45	60	90

Требуется определить, сколько тонн каждого компонента следует использовать для получения 1000 т автомобильного бензина А-76, чтобы его себестоимость была минимальной.

Решение. Для решения этой задачи сформулируем ее экономико-математическую модель, т.е. сформулируем задачу математически. Введем необходимые обозначения: пусть x_j (j = 1, 2, 3, 4) — количество в смеси компонента с номером j. Таким образом, формально план производства смеси представляет собой вектор $\bar{X} = (x_1, x_2, x_3, x_4)$.

С учетом введенных обозначений математическая модель рассматриваемой задачи по критерию «минимум суммарных затрат на изготовление смеси» имеет вид

$$\min f(\bar{X}) = 40x_1 + 45x_2 + 60x_3 + 90x_4$$

$$0{,}35x_1 + 0{,}35x_2 + 0{,}3x_3 + 0{,}2x_4 \le 0{,}3 \cdot 1000, \quad (1)$$

$$68x_1 + 72x_2 + 80x_3 + 90x_4 \ge 76 \cdot 1000, \quad (2)$$

$$x_1 + x_2 + x_3 + x_4 = 1000, \quad (3)$$

$$0 \le x_1 \le 700,$$
$$0 \le x_2 \le 600,$$
$$0 \le x_3 \le 500,$$
$$0 \le x_4 \le 300.$$

[1] *Мельник М. М.* Экономико-математические методы и модели в планировании и управлении МТС : учеб. М. : Высш. шк., 1990.

Функциональное ограничение (3) отражает необходимость получения заданного количества смеси (1000 т), (1) и (2) — ограничения по октановому числу и содержанию серы в смеси, остальные — ограничения по запасам соответствующих ресурсов (компонентов). Прямые ограничения очевидны, но принципиально важны для выбора метода решения. Полученная математическая задача (модель) — задача линейного программирования. Она может быть решена симплекс-методом, который рассмотрен в данной главе ниже (для автоматизации расчетов можно использовать надстройку **Поиск решения** MS Excel[1]). В результате решения ЗЛП получим оптимальное решение:

$$\bar{X}^* = (x_1^*, x_2^*, x_3^*, x_4^*),$$

где $x_1^* = 571{,}429$; $x_2^* = 0$; $x_3^* = 142{,}857$; $x_4^* = 285{,}714$.

Подставляя найденное решение в целевую функцию (ЦФ), имеем

$$\min f(\bar{X}) = f(\bar{X}^*) = 40 * 571{,}429 + 45 * 0 + 60 * 142{,}857 + \\ + 90 * 285{,}714 = 57\,143.$$

Таким образом, предлагаются следующие рекомендации по производству смеси:

— для производства смеси компонент № 1 берется в количестве 571,429 т;

— компонент № 3 берется в количестве 142,857 т;

— компонент № 4 берется в количестве 285,714 т.

При таком смешении ожидаются минимальные затраты на производство смеси, равные 57 143 (ден. ед.).

Пример 2.2 (задача об оптимальном использовании ограниченных ресурсов).

На участок строящейся дороги[2] необходимо вывезти 20 000 м^3 каменных материалов. В районе строительства имеются три карьера с запасами 8000 м^3, 9000 м^3 и 10 000 м^3. Для погрузки материалов используются экскаваторы, имеющие производительность 250 м^3 в смену в карьерах 1, 2 и 500 м^3 в смену в карьере 3.

Эти карьеры обеспечивают каменными материалами также ряд других строящихся объектов. На погрузку материалов для рассматриваемого участка выделен для экскаваторов общий лимит 60 машино-смен с правом использования его по усмотрению строителей.

---

[1] *Гармаш А. Н., Орлова И. В.* Математические методы в управлении : учеб. пособие. М. : Вузовский учебник, 2012.

[2] *Золотарь И. А.* Экономико-математические методы в дорожном строительстве. М. : Транспорт, 1974.

Транспортные затраты на перевозку материалов характеризуются следующими показателями: на перевозку 1000 м^3 материалов из карьера 1 требуется 100 автомобиле-смен, из карьера 2 — 1350 автомобиле-смен, из карьера 3 — 1700 автомобиле-смен.

Требуется найти оптимальный план перевозок, обеспечивающий минимальные транспортные затраты.

Решение. Сформулируем экономико-математическую модель задачи. Примем за единицу измерения количества материалов 10 000 м^3 . Обозначим через x_1 — объем добычи материалов в карьере 1, x_2 — в карьере 2, x_3 — в карьере 3. Таким образом, формально план перевозок представляет собой вектор $\bar{X} = \left(x_1, x_2, x_3\right)$.

С учетом введенных обозначений математическая модель рассматриваемой задачи по критерию «минимум суммарных транспортных затрат» имеет вид

$$\min f(\bar{X}) = 1000x_1 + 1350x_2 + 1700x_3$$

$$x_1 + x_2 + x_3 = 2{,}0, \qquad (1)$$

$$40x_1 + 40x_2 + 20x_3 \le 60{,}0, \qquad (2)$$

$$0 \le x_1 \le 0{,}8, \qquad (3)$$

$$0 \le x_2 \le 0{,}9, \qquad (4)$$

$$0 \le x_3 \le 1{,}0. \qquad (5)$$

Условие (1) отражает потребность в материалах, (2) — ограничение по наличию ресурса «фонд рабочего времени экскаваторов» (мы не можем использовать больше того, что у нас в наличии). Условия (3)—(5) отражают тот факт, что добыча материалов идет в условиях ограниченности запасов материалов в соответствующих карьерах. Полученная задача — задача линейного программирования; решив ее симплекс-методом (см. ниже), найдем оптимальный план (решение):

$$\bar{X}^* = (x_1^*, x_2^*, x_3^*) : x_1^* = 0{,}8 \text{ (8000 м}^3\text{)};\ x_2^* = 0{,}2 \text{(2000 м}^3\text{)};$$

$$x_3^* = 1{,}0 \text{ (10 000 м}^3\text{)}.$$

Таким образом, из карьера 1 следует вывезти 8000 м^3 каменных материалов, из карьера 2 — 2000 м^3, из карьера 3 — 10 000 м^3. Это управленческое решение будет связано с минимальными транспортными затратами:

$$\min f(\bar{X}) = f(\bar{X}^*) = 1000 * 0{,}8 + 1350 * 0{,}2 + 1700 * 1{,}0 =$$
$$= 2770 \text{ (автомобиле-смен)}.$$

2.3. Математический аппарат линейного программирования

Изучение и понимание современных экономико-математических методов предполагает достаточно серьезную математическую подготовку экономистов. Для освоения задач и методов в пределах данной главы необходимы знания основных понятий и элементов высшей математики, матричной и векторной алгебры. Некоторые необходимые сведения из этих разделов математики приведены ниже.

Матрицы и определители

Рассмотрим $m \times n$ действительных чисел, записанных в виде прямоугольной таблицы из m строк и n столбцов:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ . & . & . & . & . \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}.$$

Данная таблица чисел называется *числовой матрицей* (в дальнейшем — просто *матрицей*). Числа a_{ij}, которые входят в матрицу, называются ее *элементами*. Индексы i и j элемента a_{ij} указывают соответственно номера строки и столбца, в которых расположен элемент a_{ij}. Матрицу, содержащую одну строку (или один столбец), называют также вектор-строкой (или вектор-столбцом).

Матрица, все элементы которой равны нулю, называется *нулевой*. Две матрицы называются *равными*, если числа строк и столбцов одной из них равны соответственно числам строк и столбцов другой и элементы этих матриц, расположенные на соответствующих местах, равны.

Матрицей, *транспонированной* к матрице A, называется матрица вида

$$A' = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ . & . & . & . \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix},$$

т.е. строками матрицы A' являются столбцы, а столбцами — строки матрицы A.

Если число строк равно числу столбцов ($m = n$), матрицу называют *квадратной матрицей* порядка n.

Элементы a_{11}, a_{22}, a_{33}, ..., a_{nn} образуют так называемую главную диагональ квадратной матрицы; элементы a_{1n}, a_{2n-1}, ..., a_{n1} — побочную диагональ квадратной матрицы.

Рассмотрим некоторые действия над матрицами.

1. Произведением матрицы A на число λ (или, что то же самое, числа λ на матрицу A) называется матрица

$$A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ . & . & . & . \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix},$$

получающаяся из A путем умножения каждого ее элемента на число λ.

2. Под суммой двух матриц

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & . & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ и } B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ . & . & . & . \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

понимается матрица

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ . & . & . & . \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix},$$

элементы которой равны суммам соответствующих элементов матриц A и B. При этом подразумевается, что число строк (столбцов) матрицы A равно числу строк (столбцов) матрицы B. Подобным же образом определяется и разность ($A - B$) матриц A и B.

Линейные операции над матрицами подчиняются обычным законам арифметики, например:

$$A + B = B + A, \; A + 0 = A$$

(все элементы матрицы 0 — нули),

$$\lambda(A + B) = \lambda A + \lambda B, \; 0 \cdot A = 0(\lambda = 0).$$

3. Произведением матрицы A из m строк и n столбцов на матрицу B из n строк и k столбцов называется матрица $C = AB$, имеющая m строк и k столбцов, элемент C_{ij} которой, расположенный в i-й строке и j-м столбце, равен сумме произведений элементов i-й строки матрицы A на соответствующие элементы j-го столбца матрицы B, т.е. находится по формуле скалярного произведения i-й вектор-строки матрицы A на j-й вектор-столбец матрицы B:

$$C_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj}.$$

В случае квадратных матриц можно составить как произведение AB, так и произведение BA. В общем случае $AB \neq BA$, т.е. переместительный закон для матриц не выполняется.

Для произведения матриц остаются в силе следующие законы арифметики.

1. Распределительный закон $(A + B)C = AC + BC$, $C(A + B) = CA + CB$.
2. Сочетательный закон $(AB)C = A(BC)$.

Среди квадратных матриц особую роль играет матрица

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ . & . & . & . \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

все элементы которой, расположенные на главной диагонали, равны единице, а остальные — нулю. Можно проверить, что для любой матрицы A:

$AE = EA = A$. Матрица E называется *единичной.*

Матрица B называется *обратной* для матрицы A, если $AB = BA = E$. Матрица B, обратная матрице A, обозначается через A^{-1} (функция = МОБР() **Мастера функций** MS Excel).

С каждой квадратной матрицей определенным образом связано некоторое число, называемое ее *определителем* (функция =МОПРЕД() **Мастера функций** MS Excel).

Для вычисления определителя любого порядка необходимо знание его свойств и теоремы о разложении определителя.

Приведем основные свойства определителей.

1. При транспонировании матрицы ее определитель не меняется. Это свойство свидетельствует о полном равноправии строк и столбцов определителя. Следовательно,

если некоторое утверждение справедливо относительно столбцов определителя, то аналогичное утверждение справедливо и для его строк.

2. Если все элементы какого-либо столбца (строки) определителя равны нулю, то и сам определитель равен нулю.

3. При перестановке двух любых столбцов (строк) определителя его знак меняется на противоположный, а абсолютная величина остается неизменной.

4. Определитель с двумя одинаковыми столбцами (строками) равен нулю.

5. Если j-й столбец (строка) A_j определителя D является линейной комбинацией

$$A_j = \lambda B + \mu C$$

двух произвольных столбцов (строк) B и C, то и сам определитель оказывается линейной комбинацией

$$D = D_j\left(\lambda B + \mu C\right) = \lambda D_j\left(B\right) + \mu D_j\left(C\right)$$

определителей $D_j(B)$ и $D_j(C)$.

Здесь $D_j(B)$ и $D_j(C)$ — определитель D, в котором столбец (строка) j заменен соответственно на столбец (строку) B и C. Остальные столбцы (строки) сохранены без изменения.

6. При умножении любого столбца (строки) определителя на произвольное число λ сам определитель умножается на это же число.

7. Если какой-либо столбец (строка) определителя является линейной комбинацией других его столбцов (строк), то определитель равен нулю.

8. Определитель не изменится, если к элементам любого его столбца (строки) прибавить соответствующие элементы другого столбца (строки), предварительно умноженные на одно и то же число.

Рассмотрим определитель n-го порядка:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nj} & \ldots & a_{nn} \end{vmatrix}.$$

Выделим в нем некоторый элемент, например a_{ij}. Вычеркнем в определителе i-ю строку и j-й столбец, в которых расположен выделенный элемент a_{ij}. В результате останется определитель $(n - 1)$-го порядка. Этот оставшийся определитель называется *минором* элемента a_{ij} в определителе D и обозначается M_{ij}.

Величину $A_{ij} = (-1)^{i+j} M_{ij}$ называют *алгебраическим дополнением* элемента a_{ij} в определителе D (или в соответствующей квадратной матрице).

Теорема о разложении определителя. Определитель матрицы A равен сумме произведений всех элементов некоторого столбца (строки) на их алгебраические дополнения:

$$D = \left|A\right| = \sum_{j=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij}.$$

Рассмотрим примеры вычисления определителей (предполагается знание правил вычисления определителей второго порядка).

1. Вычислить определитель

$$D = \begin{vmatrix} 2 & 3 & 1 \\ 4 & 2 & -1 \\ 3 & 5 & 2 \end{vmatrix}.$$

Разложим определитель D по элементам второго столбца: $D = 3A_{12} + 2A_{22} + 5A_{32}$. Переходя к минорам, имеем

$$D = 3(-1)^{1+2}\begin{vmatrix} 4 & -1 \\ 3 & 2 \end{vmatrix} + 2(-1)^{2+2}\begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} + 5(-1)^{3+2}\begin{vmatrix} 2 & 1 \\ 4 & -1 \end{vmatrix} =$$
$$= -3 \cdot 11 + 2 \cdot 1 - 5 \cdot (-6) = -1.$$

2. Вычислить определитель четвертого порядка

$$D = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}.$$

Используя свойства определителей, получим единичную первую строку и разложим по ней определитель D;

аналогично поступим с первым столбцом преобразованного определителя:

$$D=\begin{vmatrix}1&2&3&4\\2&3&4&1\\3&4&1&2\\4&1&2&3\end{vmatrix}=\begin{vmatrix}1&0&0&0\\2&-1&-2&-7\\3&-2&-8&-10\\4&-7&-10&-13\end{vmatrix}=(-1)^{1+1}\begin{vmatrix}-1&-2&-7\\-2&-8&-10\\-7&-10&-13\end{vmatrix}=$$

$$=-\begin{vmatrix}1&2&7\\2&8&10\\7&10&13\end{vmatrix}=-\begin{vmatrix}1&2&7\\0&4&-4\\0&-4&-36\end{vmatrix}=-(-1)^{1+1}\begin{vmatrix}4&-4\\-4&-36\end{vmatrix}=-4\cdot 4\begin{vmatrix}-1&1\\1&9\end{vmatrix}=$$

$$=-16(-9-1)=16\cdot 10=160.$$

Решение систем линейных уравнений

Рассмотрим систему из n линейных уравнений с n неизвестными (такие системы линейных уравнений называются *определенными*):

$$\begin{aligned}&a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1,\\&a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2,\\&.\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\\&a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n.\end{aligned}\tag{2.15}$$

Определитель Δ, составленный из коэффициентов при неизвестных, называют определителем системы (2.15):

$$\Delta=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\\cdot&\cdot&\cdot&\cdot\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}.$$

Решить систему уравнений (2.15) можно различными методами, в частности методом Крамера. В основе решения системы уравнений (2.15) методом Крамера лежит следующая теорема.

Теорема Крамера. Если определитель Δ системы (2.15) отличен от нуля, то система совместна и имеет единственное решение, которое можно найти по формуле

$$x_j=\frac{\Delta_j}{\Delta},\quad j=\overline{1,n}\,.$$

В этой формуле Δ_j является определителем, полученным из определителя системы Δ путем замены столбца j столбцом свободных членов.

Систему n линейных уравнений с n неизвестными (2.15) можно записать в матричном виде: $AX = B$, где A — квадратная матрица порядка n, составленная из коэффициентов при неизвестных; X — вектор-столбец из неизвестных; B — вектор-столбец свободных членов:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & . & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Если A — невырожденная матрица, т.е. ее определитель $|A| \neq 0$, то можно определить A^{-1}. С учетом этого имеют место матричные соотношения:

$$A^{-1} \cdot A \cdot X = A^{-1} \cdot B, \quad E \cdot X = A^{-1} \cdot B, \quad X = A^{-1} \cdot B. \qquad (2.16)$$

Обратная матрица может быть определена на базе следующей теоремы.

Теорема 2.1. Если определитель матрицы A не равен нулю, то матрица A имеет обратную матрицу A^{-1}, которая находится по формуле

$$A^{-1} = \frac{1}{\Delta} \tilde{A},$$

где $\tilde{A}$ — матрица, *присоединенная* к матрице A.

Матрица $\tilde{A}$ составляется из алгебраических дополнений к элементам транспонированной матрицы:

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ . & . & . & . \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}.$$

Таким образом, соотношение (2.16) лежит в основе решения системы уравнений (2.15) методом обратной матрицы (функция = МУМНОЖ (МОБР(A), B) **Мастера функций** MS Excel).

Рассмотрим систему m линейных уравнений с n неизвестными (при $m < n$ такие системы называются *неопределенными*):

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ &a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{aligned} \qquad (2.17)$$

или в векторной записи:

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n = B,$$

где $A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$, $A_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}$, ..., $A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ — соответствующие вектор-столбцы.

Запишем *расширенную* матрицу этой системы в виде

$$\begin{array}{cc} & \begin{array}{ccccc} A_1 & A_2 & \cdots & A_n & B \end{array} \\ \hat{A} = & \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ . & . & . & . & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right). \end{array}$$

Элементарными преобразованиями системы (2.17) (или матрицы $\hat{A}$) называются следующие преобразования:

- перестановка любых двух уравнений;
- умножение обеих частей одного из уравнений на любое отличное от нуля число;
- прибавление к обеим частям одного уравнения соответствующих частей другого, умноженных на любое число, отличное от нуля;
- вычеркивание нулевой строки (уравнения с нулевыми коэффициентами и свободными членом, равным 0).

Можно показать, что элементарные преобразования переводят данную систему уравнений в эквивалентную систему. Две системы линейных уравнений называются *эквивалентными*, или *равносильными*, если каждое решение первой системы (если они существуют) является решением второй, и наоборот. Соответствующие расширенные матрицы также называются эквивалентными.

При практическом решении системы линейных уравнений *методом Жордана — Гаусса* последовательно над строками матрицы $\hat{A}$ выполняют элементарные преобразования, так что некоторое неизвестное исключается из всех уравнений, кроме одного, т.е. в составе расширенной матрицы формируется единичная подматрица.

В процессе решения могут встретиться следующие случаи.

1. Будет получена матрица $\hat{A}'$, эквивалентная матрице $\hat{A}$, в левой части некоторой строки ее стоят нули, а в правой — число, отличное от нуля, что соответствует уравнению

$$0x_1+0x_2+\cdots+0x_n=b_1', (b_1'\neq 0).$$

Это признак несовместности системы (2.17), т.е. система не имеет решений.

2. В результате преобразований получилась матрица $\hat{A}'$ вида

$$\hat{A}'=\begin{pmatrix} 1 & 0 & \cdots & 0 & b_1' \\ 0 & 1 & \cdots & 0 & b_2' \\ . & . & . & . & . \\ 0 & 0 & \cdots & 1 & b_n' \end{pmatrix}.$$

В этом случае система (2.17) совместна, определенная и имеет единственное решение: $x_1=b_1', x_2=b_2', \ldots, x_n=b_n'$.

3. На некотором этапе получилась расширенная матрица вида

$$\hat{A}'=\left(\begin{array}{ccccccc|c} 1 & 0 & \cdots & 0 & a_{1r+1}' & \cdots & a_{1n}' & b_1' \\ 0 & 1 & \cdots & 0 & a_{2r+1}' & \cdots & a_{2n}' & b_2' \\ . & . & . & . & . & . & . & . \\ 0 & 0 & \cdots & 1 & a_{rr+1}' & \cdots & a_{rn}' & b_r' \end{array}\right).$$

Система совместна и имеет бесчисленное множество решений. *Общее решение* системы можно записать в виде

$$\begin{aligned} x_1 &= b_1'-a_{1r+1}'x_{r+1}-\cdots-a_{1n}'x_n \\ x_2 &= b_2'-a_{2r+1}'x_{r+1}-\cdots-a_{2n}'x_n \\ &. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ x_r &= b_r'-a_{rr+1}'x_{r+1}-\cdots-a_{rn}'x_n. \end{aligned}$$

Придавая каждой из стоящих в правых частях равенств переменных $x_{r+1}, x_{r+2}, \ldots, x_n$ произвольные значения, будем получать *частные решения* системы.

Неизвестные $x_1, x_2, \ldots, x_r$ называются *базисными*, или *основными*, они соответствуют линейно-независимым векторам $A_1, \ldots, A_r$.

Таким образом, любые r переменных называются базисными (основными), если определитель матрицы коэффициентов при них отличен от нуля, а остальные $(n - r)$ переменных называются *свободными*, или *неосновными*. *Базисным решением* системы уравнений называется частное решение, в котором неосновные переменные имеют нулевые значения. Каждому разбиению на основные и неосновные переменные соответствует одно базисное решение, а количество способов разбиения не превышает величины

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Если все компоненты базисного решения неотрицательны, то такое решение называется *опорным*.

Пример 2.3. Исследовать систему уравнений методом Жордана — Гаусса.

$$\begin{aligned}
&x_1 - 2x_2 + 3x_3 - 4x_4 + 2x_5 = 4\\
&x_2 - x_3 + x_4 + 4x_5 = -3\\
&x_1 + 3x_2 - 3x_4 = 1\\
&x_1 + x_2 + x_3 - 3x_4 + 3x_5 = 1.
\end{aligned}$$

Решение. Запишем расширенную матрицу системы уравнений и последовательно преобразуем ее элементарными преобразованиями

$$\left(\begin{array}{ccccc|c}
1 & -2 & 3 & -4 & 2 & 4\\
0 & 1 & -1 & 1 & 4 & -3\\
1 & 3 & 0 & -3 & 0 & 1\\
1 & 1 & 1 & -3 & 3 & 1
\end{array}\right) \to
\left(\begin{array}{ccccc|c}
1 & -2 & 3 & -4 & 2 & 4\\
0 & 1 & -1 & 1 & 4 & -3\\
0 & 5 & -3 & 1 & -2 & -3\\
0 & 3 & -2 & 1 & 1 & -3
\end{array}\right) \to$$

$$\to \left(\begin{array}{ccccc|c}
1 & 0 & 1 & -2 & 10 & -2\\
0 & 1 & -1 & 1 & 4 & -3\\
0 & 0 & 2 & -4 & -22 & 12\\
0 & 0 & 1 & -2 & -11 & 6
\end{array}\right) \to
\left(\begin{array}{ccccc|c}
1 & 0 & 1 & -2 & 10 & -2\\
0 & 1 & -1 & 1 & 4 & -3\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & -2 & -11 & 6
\end{array}\right) \to$$

$$\to \left(\begin{array}{ccccc|c}
1 & 0 & 0 & 0 & 21 & -8\\
0 & 1 & 0 & -1 & -7 & 3\\
0 & 0 & 1 & -2 & -11 & 6
\end{array}\right), \quad
\begin{aligned}
&x_1 + 21x_5 = -8\\
&x_2 - x_4 - 7x_5 = 3\\
&x_3 - 2x_4 - 11x_5 = 6.
\end{aligned}$$

Таким образом, система совместна, имеет бесчисленное множество решений. Общее решение записывается в виде

$$x_1 = -8 - 21x_5,$$
$$x_2 = 3 + x_4 + 7x_5,$$
$$x_3 = 6 + 2x_4 + 11x_5.$$

Любое частное решение получается из общего путем придания конкретных значений свободным переменным x_4 и x_5. Например, (−8; 4; 8; 1; 0) — частное решение. Одно из базисных решений получаем при $x_4 = 0$ и $x_5 = 0$, т.е. (−8; 3; 6; 0; 0).

Число базисных решений не превосходит $C_5^3 = 10$. Перейдем к другому базисному решению, взяв в расширенной матрице в качестве базисных решений векторы A_1, A_2, A_4; при этом переменные x_1, x_2, x_4 будут базисными, а x_3, x_5 — свободными. Переход от одного базиса к другому осуществим методом Жордана — Гаусса, т.е. используя элементарные преобразования:

$$\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & -8 \\ 0 & 1 & 0 & -1 & -7 & 3 \\ 0 & 0 & 1 & -2 & -11 & 6 \end{array}\right) \to \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 21 & -8 \\ 0 & 1 & 0 & -1 & -7 & 3 \\ 0 & 0 & -0{,}5 & 1 & 5{,}5 & -3 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 21 & -8 \\ 0 & 1 & -0{,}5 & 0 & -1{,}5 & 0 \\ 0 & 0 & -0{,}5 & 1 & 5{,}5 & -3 \end{array}\right).$$

Таким образом, получено еще одно базисное решение: (−8; 0; 0; –3; 0) и т.д.

Заметим, что оба полученных базисных решения не являются опорными решениями; последнее решение является также *вырожденным* (базисная переменная x равна 0).

Линейные векторные пространства

Определение 2.1. Упорядоченная система из n действительных чисел $a_1, a_2, \ldots, a_n$ называется n-мерным вектором и обозначается $\overline{a} = A = (a_1, a_2, \ldots, a_n)$. Числа $a_j (j = 1, 2, \ldots, n)$ называются компонентами вектора $\overline{a} = A$.

Определение 2.2. Совокупность всевозможных n-мерных векторов с введенными на ней операциями сложения и умножения на число называется n-мерным векторным пространством.

В матрице из m строк и n столбцов строки являются n-мерными векторами, столбцы — m-мерными векторами.

Вектор $\overline{a}=(a_1, a_2, \ldots, a_n)$ и вектор $\overline{b}=(b_1, b_2, \ldots, b_n)$ равны, если совпадают их компоненты, стоящие на одинаковых местах, т.е. если $a_j=b_j$ при $j=1, 2, \ldots, n$.

Суммой векторов $\overline{a}$ и $\overline{b}$ называется вектор $\overline{a}+\overline{b}=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$. Роль нуля играет нулевой вектор $\overline{0}=(0, 0, \ldots, 0)$.

Противоположным вектору $\overline{a}$ называется вектор $-\overline{a}=(-a_1, -a_2, \ldots, -a_n)$; очевидно, что $\overline{a}+(-\overline{a})=\overline{0}$.

Разность векторов $\overline{a}-\overline{b}=\overline{a}+(-\overline{b})$.

Произведением вектора $\overline{a}$ на число λ называется вектор $\lambda\overline{a}=(\lambda a_1, \lambda a_2, \ldots, \lambda a_n)$. Из этого определения вытекают следующие важные свойства:

$$\lambda(\overline{a}\pm\overline{b})=\lambda\overline{a}\pm\lambda\overline{b},$$
$$(k\pm\lambda)\overline{a}=k\overline{a}\pm\lambda\overline{a},$$
$$k(\lambda\overline{a})=(k\lambda)\overline{a}.$$

Следствиями этих свойств являются следующие: $0\cdot\overline{a}=0$, $(-1)\overline{a}=-\overline{a}$, $\lambda\cdot\overline{0}=\overline{0}$. Скалярным произведением двух векторов $\overline{a}$ и $\overline{b}$ (A и B) называется действительное число, равное сумме произведений соответствующих компонент этих векторов:

$$AB=a_1b_1+a_2b_2+\ldots+a_nb_n.$$

Например, левая часть линейного уравнения $a_1x_1+a_2x_2+\ldots+a_nx_n=b$ может быть представлена в виде скалярного произведения векторов $A\cdot X$, где $A=(a_1, a_2, \ldots, a_n)$, $X=(x_1, x_2, \ldots, x_n)$.

Вектор B называется *линейной комбинацией* векторов $A_1, A_2, \ldots, A_n$, если существуют такие числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, при которых выполняется соотношение $B=\lambda_1A_1+\lambda_2A_2+\ldots+\lambda_nA_n$. Система векторов $A_1, A_2, \ldots, A_r (r\geq 2)$ называется *линейно зависимой*, если хотя бы один из векторов системы является линейной комбинацией остальных, и *линейно независимой* — в противном случае. Можно сформулировать следующие равносильные сказанному определения.

Система векторов $A_1, A_2, \ldots, A_r$ — линейно зависимая, если существуют числа $\lambda_1, \lambda_2, \ldots, \lambda_r$, не все равные нулю, при которых имеет место равенство $\lambda_1A_1+\lambda_2A_2+\ldots+\lambda_rA_r=0$.

Если последнее соотношение возможно лишь в случае, когда все $\lambda_j=0$ $(j=\overline{1,r})$, то система векторов называется линейно независимой.

Например, система векторов $A_1 = (2, 4, 3)$, $A_2 = (2, 3, 1)$, $A_3 = (5, 3, 2)$, $A_4 = (1, 7, 3)$ линейно зависима: $A_1 + 2A_2 - A_3 - A_4 = 0$.

Рангом системы векторов

$$\begin{gathered} A_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \\ A_2 = (a_{21}, a_{22}, \ldots, a_{2n}), \\ \ldots\ldots\ldots\ldots \\ A_m = (a_{m1}, a_{m2}, \ldots, a_{mn}) \end{gathered}$$

называется максимальное число линейно независимых векторов этой системы. Ранг системы векторов равен *рангу матрицы* A, составленной из компонент векторов этой системы, т.е. наивысшему порядку минора матрицы A, отличного от нуля.

Пример 2.4. Определить, является ли система векторов $A_1 = (5, 4, 3, 2)$, $A_2 = (3, 3, 2, 2)$, $A_3 = (8, 1, 3, -4)$ линейно зависимой; если она линейно-зависима, то найти ее максимальную линейно независимую подсистему.

Решение. Составим матрицу из компонент векторов и найдем ее ранг.

Имеем $$A = \begin{pmatrix} 5 & 4 & 3 & 2 \\ 3 & 3 & 2 & 2 \\ 8 & 1 & 3 & -4 \end{pmatrix}.$$

Минор второго порядка $\begin{vmatrix} 5 & 4 \\ 3 & 3 \end{vmatrix} = 3 \neq 0.$

Рассмотрим два минора третьего порядка, которые его окаймляют:

$$\begin{vmatrix} 5 & 4 & 3 \\ 3 & 3 & 2 \\ 8 & 1 & 3 \end{vmatrix} = 118 - 118 = 0, \begin{vmatrix} 5 & 4 & 2 \\ 3 & 3 & 2 \\ 8 & 1 & -4 \end{vmatrix} = 2(59 - 59) = 0.$$

Ранг матрицы A равен 2, поэтому система векторов является зависимой. В матрицах, составленных из компонент любых двух векторов данной системы, содержатся миноры второго порядка, отличные от нуля, например,

$$\begin{vmatrix} 5 & 4 \\ 3 & 3 \end{vmatrix} = 3 \neq 0, \begin{vmatrix} 3 & 3 \\ 8 & 1 \end{vmatrix} = -21 \neq 0, \begin{vmatrix} 5 & 4 \\ 8 & 1 \end{vmatrix} = -27 \neq 0.$$

Поэтому максимальная линейно независимая подсистема состоит из двух любых векторов, а третий вектор является их линейной комбинацией.

Базисом n-мерного векторного пространства называется любая совокупность n линейно независимых векторов этого же пространства.

Теорема 2.2. Любой вектор n-мерного векторного пространства можно представить как линейную комбинацию векторов базиса, притом единственным образом.

Один из базисов n-мерного векторного пространства образует система единичных векторов

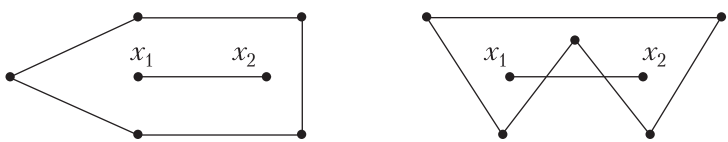
$$\begin{aligned} &E_1 = (1, 0, \ldots, 0) \\ &E_2 = (0, 1, \ldots, 0) \\ &. \quad . \quad . \quad . \quad . \quad . \\ &E_n = (0, 0, \ldots, 1). \end{aligned}$$

Компоненты любого n-мерного вектора можно считать координатами этого вектора в единичном базисе.

Пусть задано n-мерное линейное пространство E^n.

Определение 2.3. Множество X называется выпуклым, если вместе с любыми точками x_1 и x_2 множеству принадлежат точки (отрезок) $\lambda x_2 + (1-\lambda)x_1$ при всех $0 \le \lambda \le 1$.

Множество на рис. 2.1, *а* — выпуклое, на рис. 2.1, *б* — невыпуклое.



а) *б)*

Рис. 2.1

Определение 2.4. Функция $f(\bar{X})$, заданная на выпуклом множестве $X \subset E^n$, называется выпуклой, если для любых двух точек x_1 и x_2 из X и любого числа $0 \le \lambda \le 1$ выполняется соотношение

$$f\left[\lambda x_2 + (1-\lambda)x_1\right] \le \lambda f(x_2) + (1-\lambda) f(x_1).$$

Определение 2.5. Функция $f(\bar{X})$, заданная на выпуклом множестве $X \subset E^n$, называется вогнутой, если для любых двух точек x_1 и x_2 из X и любого числа $0 \le \lambda \le 1$ выполняется соотношение

$$f\left[\lambda x_2 + (1-\lambda)x_1\right] \ge \lambda f(x_2) + (1-\lambda)f(x_1).$$

Если приведенные неравенства считать строгими и они выполняются при $0 < \lambda < 1$, то функция $f(\bar{X})$ — строго выпуклая (вогнутая).

Можно показать, что если $f(\bar{X})$ — выпуклая функция, то функция $-f(\bar{X})$ — вогнутая, и наоборот.

На рис. 2.2, *а* функция $f(\bar{X})$ — выпуклая, на рис. 2.2, *б* — вогнутая.

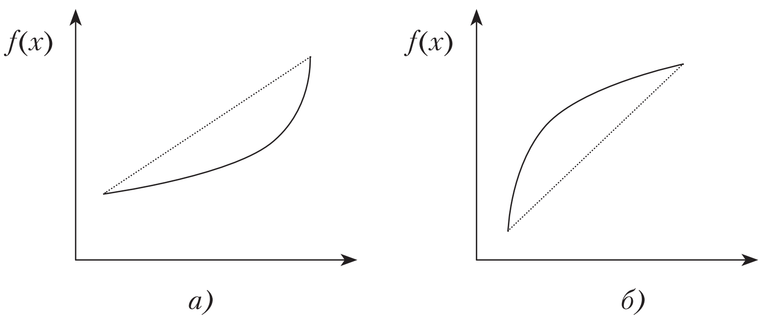


Рис. 2.2

Справедливы следующие утверждения относительно выпуклых множеств и функций.

1. Пересечение выпуклых множеств есть выпуклое множество.

2. Сумма вогнутых (выпуклых) функций есть вогнутая (выпуклая) функция.

3. Если $f(\bar{X})$ — выпуклая функция при $\bar{X} \ge 0$, то множество всех точек, удовлетворяющих условиям $f(\bar{X}) \le b$, $\bar{X} \ge 0$, выпукло (если оно не пустое; b — это постоянная).

4. Пусть $f(\bar{X})$ — выпуклая (вогнутая) функция, заданная на замкнутом выпуклом множестве $X \subset E^n$, тогда любой локальный минимум (максимум) $f(\bar{X})$ на X является и глобальным.

Приведем необходимое и достаточное условие выпуклости функции многих переменных. Пусть функция $f(\bar{X}) = f(x_1, x_2, \ldots, x_n)$ имеет все частные производные второго порядка, образующие матрицу

$$Q(\bar{X}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \ldots\ldots\ldots\ldots & \ldots & \ldots & \ldots\ldots\ldots\ldots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

Эта функция является выпуклой в области X тогда и только тогда, когда матрица Q для любой точки из этой области является неотрицательно (положительно) определенной. Напомним, что квадратная матрица $Q = (q_{i,j})_{n \times x}$ называется *неотрицательно (положительно) определенной*, если все определители

$$\Delta_1 = q_{11},\ \Delta_2 = \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix}, \ldots, \Delta_n = \begin{vmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ . & . & . & . \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{vmatrix},$$

т.е. все главные миноры матрицы, неотрицательны (положительны).

Пример 2.5. Показать, что функция $f(\bar{X}) = 2x_1^3 + x_2 - 6$ является выпуклой при $x_1 \geq 0$.

Составим матрицу из частных производных второго порядка для $f(\bar{X}) : Q(\bar{X}) = \begin{pmatrix} 12x_1 & 0 \\ 0 & 0 \end{pmatrix}$.

Найдем главные миноры $\Delta_1 = 12x_1,\ \Delta_2 = 0.$ Так как $\Delta_1 \geq 0,\ \Delta_2 = 0$ при $x_1 \geq 0,$ то функция является выпуклой.

Дадим определение глобального и локального максимумов. Функция $f(\bar{x})$ достигает на замкнутом (т.е. включающем свою границу) множестве X глобальный максимум в точке $\bar{x}^*$, если для любой точки, принадлежащей $X (\bar{x} \in X),$ выполняется условие $f(\bar{x}) \leq f(\bar{x}^*).$

Функция $f(\overline{x})$ достигает на замкнутом множестве X локального максимума в точке $\overline{x}^0$, если существует некоторой окрестность этой точки, для каждой точки которой выполняется условие $f(\overline{x}) \le f(\overline{x}^0)$.

Определения локального и глобального минимума формулируются аналогично.

На рис. 2.3 x_3^0 — точка локального минимума; x_1^0 — глобального минимума; α, x_2^0 — точки локального максимума; β — точка глобального максимума.

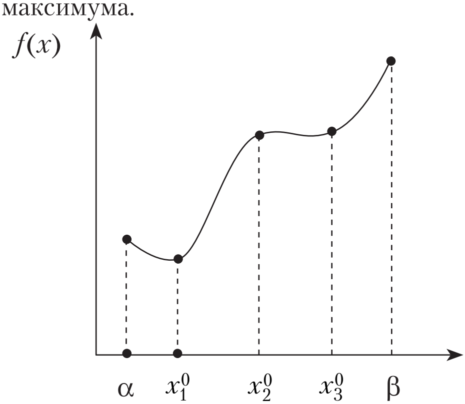


Рис. 2.3

Необходимые условия экстремума (максимума, минимума). Если в точке $\overline{x}^0 \in X$ функция $f(\overline{x}) = f(x_1, x_2, \ldots, x_n)$ имеет экстремум, то частные производные первого порядка равны нулю в этой точке:

$$\frac{\partial f(\overline{x}^0)}{\partial x_j} = 0, \; j \equiv \overline{1,n}.$$

Достаточные условия существования экстремума здесь не формулируются. О самом существовании точек глобального минимума и максимума говорит следующая теорема.

Теорема Вейерштрасса. Если функция $f(\overline{x})$ определена и непрерывна в ограниченной замкнутой области X, то она достигает в ней своих точных верхней и нижней границ (глобальный максимум и глобальный минимум).

Приведенные утверждения относительно выпуклых множеств и функций, условий существования экстремума позволяют делать выводы о свойствах тех или иных задач оптимального программирования, что является основой разработки и применения математических методов их решения. Например, симплекс-метод решения задачи линейного программирования использует, в частности, «свойство выпуклости» этой задачи: не существует локального экстремума, отличного от глобального.

2.4. Геометрическая интерпретация задачи

Рассмотрим задачу ЛП в стандартной форме записи:

$$\max(\min) f(\bar{X}) = f(x_1, x_2, ..., x_m) = c_1x_1 + c_2x_2 + ... + c_nx_n \quad (2.18)$$

при ограничениях

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \le b_1, \\ &a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \le b_2, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \quad (2.19) \\ &a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \le b_m, \\ &x_j \ge 0, j = 1, 2, ..., n. \quad (2.20) \end{aligned}$$

Рассмотрим эту задачу на плоскости, т.е. при $n = 2$. Пусть система неравенств (2.19), (2.20) совместна (имеет хотя бы одно решение):

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 \le b_1, \\ &a_{21}x_1 + a_{21}x_2 \le b_2, \\ &\dots\dots\dots\dots\dots \\ &a_{m1}x_1 + a_{m2}x_2 \le b_m, \\ &x_{1,2} \ge 0. \end{aligned}$$

Каждое неравенство этой системы геометрически определяет полуплоскость с граничной прямой $a_{i1}x_1 + a_{i2}x_2 = b_i, i = 1, 2, ..., m.$ Условия неотрицательности определяют полуплоскости, соответственно, с граничными прямыми $x_1 = 0$, $x_2 = 0$. Система совместна, поэтому полуплоскости, как выпуклые множества, пересекаясь, образуют общую часть, которая является выпуклым множеством и представляет собой совокупность точек, координаты каждой из которых являются решением данной системы. Совокупность этих точек называют *многоугольником решений*.

Он может быть точкой, отрезком, лучом, многоугольником, неограниченной многоугольной областью.

Если в системе ограничений (2.19)—(2.20) $n = 3$, то каждое неравенство геометрически представляет полупространство трехмерного пространства, граничная плоскость которого $a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 = b_i$, $i = 1, 2, ..., m$. Условия неотрицательности определяют полупространства с граничными плоскостями, соответственно $x_j = 0$, $j = 1, 2, 3$. Если система ограничений совместна, то эти полупространства, как выпуклые множества, пересекаясь, образуют в трехмерном пространстве общую часть, которая называется *многогранником решений*.

Пусть в системе ограничений (2.19)—(2.20) $n > 3$, тогда каждое неравенство определяет полупространство n-мерного пространства с граничной гиперплоскостью $a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n = b_i$, $i = 1, 2, ..., m$, а условия неотрицательности — полупространства с граничными гиперплоскостями $x_j = 0$, $j = 1, 2, ..., n$.

Если система ограничений совместна, то, по аналогии с трехмерным пространством, она образует общую часть n-мерного пространства, называемую *многогранником решений*, так как координаты каждой его точки являются решением.

Таким образом, геометрически задача линейного программирования (2.18)—(2.20) представляет собой отыскание такой точки многогранника решений, координаты которой доставляют линейной функции цели максимальное (минимальное) значение, причем допустимыми решениями являются все точки многогранника решений.

Если в ЗЛП ограничения заданы в виде неравенств с двумя переменными, то она может быть решена графически. Графический метод решения ЗЛП состоит из следующих этапов.

Этап 1. Сначала на координатной плоскости $x_1 0 x_2$ строится допустимая многоугольная область (область допустимых решений, область определения задачи), соответствующая ограничениям. Далее строится вектор-градиент линейной функции $f(\bar{X})$ в какой-нибудь точке $\bar{x}_0$, принадлежащей допустимой области:

$$\operatorname{grad} f(\bar{X}) = \nabla = \left(\frac{\partial f}{\partial x_1} = c_1, \frac{\partial f}{\partial x_2} = c_2\right). \qquad (2.21)$$

Этап 2. Прямая $c_1x_1 + c_2x_2 = f(\bar{x}_0)$, называемая *линией уровня* и перпендикулярная вектору-градиенту, передвигается в направлении этого вектора в случае максимизации $f(\bar{X})$ до тех пор, пока не покинет пределов области допустимых решений (ОДР). Предельная точка (или точки) области при этом движении и является точкой максимума $f(\bar{X})$.

Этап 3. Для нахождения координат точки максимума достаточно совместно решить два уравнения прямых, получаемых из соответствующих ограничений и дающих в пересечении точку максимума. Значение $f(\bar{X})$, найденное в получаемой точке, является максимальным.

В случае минимизации $f(\bar{X})$ прямую $c_1x_1 + c_2x_2 = f(\bar{x}_0)$ надо перемещать в направлении, противоположном вектору-градиенту. Ясно, что если прямая при своем движении не покидает ОДР, то соответствующий максимум или минимум $f(\bar{X})$ не существует (область определения задачи — незамкнутый многоугольник).

Пример 2.6. Решить графическим методом следующую ЗЛП:

$$\max f(\bar{X}) = 30x_1 + 60x_2$$
$$x_1 + 3x_2 \le 21$$
$$3x_1 + 2x_2 \le 21$$
$$3x_1 + x_2 \le 18$$
$$x_{1,2} \ge 0$$

Решение. Прямые ограничения означают, что область решений будет лежать в первой четверти декартовой (прямоугольной) системы координат; отметим штриховкой эту область на рис. 2.4.

Этап 1. Определим множество решений первого неравенства. Оно состоит из решения уравнения и строгого неравенства. Решением уравнения служат точки прямой $x_1 + 3x_2 - 21 = 0$. Построим прямую по двум точкам (0; 7) и (21; 0), которые легко получить в результате последовательного обнуления одной из переменных. На рис. 2.4 обозначим ее цифрой I.

Множество решений строгого неравенства — одна из полуплоскостей, на которые делит плоскость построенная прямая. Какая из них является искомой, можно выяснить при помощи одной контрольной точки. Если в произвольно взятой точке, не принадлежащей прямой, неравенство выполняется, то оно выполняется и во всех точках той полуплоскости, которой принадлежит контрольная точка, и не выполняется во всех точках другой полупло-

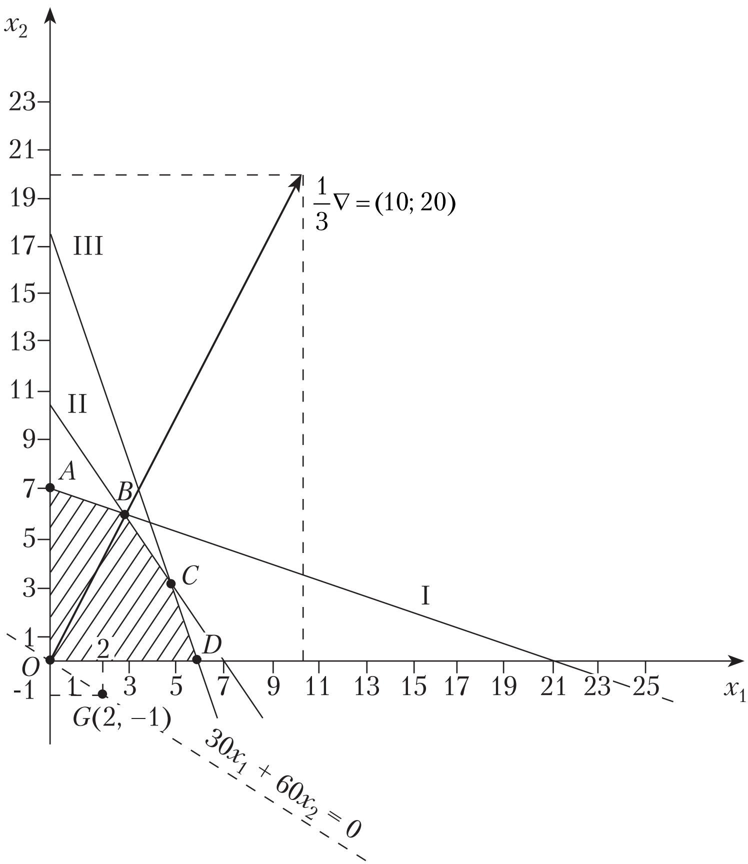


Рис. 2.4. **Решение ЗЛП графическим методом**

скости. В качестве такой точки удобно брать начало координат. Подставим координаты (0; 0) в неравенство, получим $-21 < 0$, т.е. оно выполняется. Следовательно, областью решения неравенства служит нижняя полуплоскость.

Аналогичным образом построим области решения двух других неравенств.

$3x_1 + 2x_2 - 21 = 0$; $x_1 = 0$, $x_2 = 10{,}5$; $x_2 = 0$, $x_1 = 7$ (на рис. 2.4 прямая II);

$3x_1 + 2x_2 - 21 < 0$ при $x_1 = x_2 = 0$, $-21 < 0$ выполняется, берется нижняя полуплоскость.

$3x_1 + x_2 - 18 = 0$; $x_1 = 0$, $x_2 = 18$; $x_2 = 0$, $x_1 = 6$ (на рисунке прямая III);

$3x_1 + x_2 - 18 < 0$ при $x_1 = x_2 = 0$, $-18 < 0$ выполняется, берется нижняя полуплоскость.

Заштрихуем общую область для всех неравенств, обозначим вершины многоугольника латинскими буквами и определим их координаты, решая систему уравнений двух пересекающихся соответствующих прямых. Например, определим координаты точки C, являющейся точкой пересечения второй и третьей прямой:

$$\begin{cases} 3x_1+2x_2+21, x_2=3, x_1=5, \\ 3x_1 + x_2=18. \end{cases}$$

Вычислим значение целевой функции в этой точке:

$f(\bar{X}) = 30x_1 + 60\, x_2 = 150 + 180 = 330.$

Аналогично поступим для других точек, являющихся вершинами замкнутого выпуклого многоугольника $OABCD$, представляющего собой область допустимых решений рассматриваемой ЗЛП. Координаты этих вершин имеют следующие значения:

$O(0; 0)$, $A(0; 7)$, $B(3; 6)$, $C(5; 3)$, $D(6; 0)$.

Этап 2. Приравняем целевую функцию постоянной величине a: $30x_1 + 60\, x_2 = a$.

Это уравнение является множеством точек, в котором целевая функция принимает значение, равное a. Меняя значение a, получим семейство параллельных прямых, каждая из которых называется *линией уровня*. Пусть $a = 0$, вычислим координаты двух точек, удовлетворяющих соответствующему уравнению $30x_1 + 60x_2 = 0$. В качестве одной из этих точек удобно взять точку O (0; 0), а так как при $x_1 = 2$ будет $x_2 = -1$, то в качестве второй точки возьмем точку $G(2; -1)$.

Через эти две точки проведем линию уровня $f(\bar{X}) = 30x_1 + 60x_2 = 0$ (пунктирная прямая на рис. 2.4).

Следует отметить, что во многих случаях удобно брать значение a равным не нулю, а целому числу, делящемуся без остатка на коэффициенты при неизвестных в целевой функции задачи. Например, в рассматриваемой задаче можно было бы брать уравнение линии уровня в виде $30x_1 + 60\, x_2 = 120$; тогда легко бы определялись две точки пересечения линии уровня с координатными осями.

Этап 3. Для определения направления движения к оптимуму построим вектор-градиент ∇, координаты которого (в соответствии с (2.21)) являются частными производными функции $f(\bar{X})$, т.е. $\nabla = (c_1, c_2) = (30; 60)$. Чтобы построить этот вектор, нужно соединить точку (30; 60) с началом координат. При максимизации целевой функции необходимо двигаться в направлении вектора-градиента, а при минимизации — в противоположном направлении. Для удобства можно строить вектор, пропорциональный вектору ∇. Так, на рис. 2.4 изображен вектор $(1/3)\nabla = (10; 20)$.

В нашем случае движение линии уровня (геометрически она перпендикулярна вектору-градиенту) будем осуществлять до ее пересечения с точкой B, далее она выходит из области допустимых решений. Следовательно, именно в этой точке достигается максимум целевой функции. Отсюда легко записать решение исходной ЗЛП: $\max f(\bar{X}) = 450$ и достигается при $x_1 = 3$, $x_2 = 6$.

Если поставить задачу минимизации функции $f(\bar{X}) = 30x_1 + 60x_2$ при тех же ограничениях, линию уровня необходимо смещать параллельно самой себе в направлении, противоположном вектору-градиенту ∇. Как это видно на рис. 2.4, минимум целевой функции достигается в точке $O(0; 0)$, следовательно, можно записать $\min f(\bar{X}) = 0$ и достигается при $x_1 = 0$, $x_2 = 0$.

При решении некоторых ЗЛП графическим методом может встретиться случай, когда линия уровня параллельна одной из сторон выпуклого многоугольника допустимых решений, причем эта сторона расположена в направлении смещения линии уровня при стремлении целевой функции к своему оптимуму. В этом случае оптимальное значение целевой функции достигается не в одной, а в двух угловых точках (вершинах) многоугольника решений и, следовательно, во всех точках отрезка, соединяющего эти вершины, т.е. задача будет иметь бесчисленное множество решений (первый особый случай).

Этот особый случай решения ЗЛП (случай неединственности решения) иллюстрируется, например, следующей задачей:

$$\max f(x_1, x_2) = (8x_1 + 10x_2)$$
$$5x_1 + x_2 \le 15$$
$$4x_1 + 5x_2 \le 40$$
$$x_2 \ge 3$$
$$x_1 \ge 0$$

Графическое решение этой задачи показано на рис. 2.5.

Линия уровня $8x_1 + 10x_2 = a$ параллельна одной из линий по границе области допустимых решений ($4x_1 + 5x_2 = 40$).

Это означает, что задача имеет бесконечное множество оптимальных решений (его задают координаты точек отрезка BC), среди которых два оптимальных решения — в угловых точках $B(0; 8)$ и $C(5/3; 20/3)$.

Точки отрезка BC задаются уравнением $x_2 = 8 - 0{,}8x_1$, где $0 \le x_1 \le 5/3$. При этом максимальное значение целевой функции равно 80.

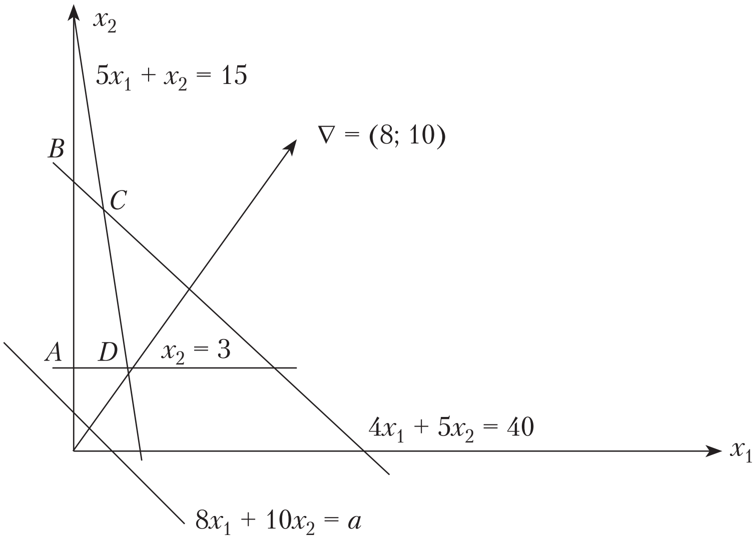


Рис. 2.5. **Случай неединственности решения ЗЛП**

Задача линейного программирования не будет иметь решений в случае, когда область допустимых решений есть пустое множество, т.е. система ограничений ЗЛП содержит противоречивые неравенства и на координатной плоскости нет ни одной точки, удовлетворяющей этим ограничениям (второй особый случай).

Очевидно также, если область допустимых решений задачи является незамкнутым выпуклым многоугольником в направлении оптимизации целевой функции, то целевая функция будет неограниченной и ЗЛП не будет иметь решений; в этом случае условно можно записать, что, например, $\max f(\bar{X}) = +\infty$ (третий особый случай).

Указанные выше два случая отсутствия решений в ЗЛП иллюстрируются рис. 2.6 и 2.7, на которых графически представлены, соответственно, задачи:

$$\begin{array}{c} \max f(x_1, x_2) = (3x_1 + 5x_2) \\ x_1 + x_2 \geq 2 \\ 4x_1 + 2x_2 \leq 2 \\ x_{1,2} \geq 0 \end{array} \qquad \begin{array}{c} \max f(x_1, x_2) = (3x_1 + 2x_2) \\ x_1 - x_2 \leq 1 \\ 2x_1 + x_2 \geq 1 \\ x_{1,2} \geq 0 \end{array}$$

Как видно из рисунков, первая из приведенных ЗЛП не имеет решения, поскольку ее множество допустимых ре-

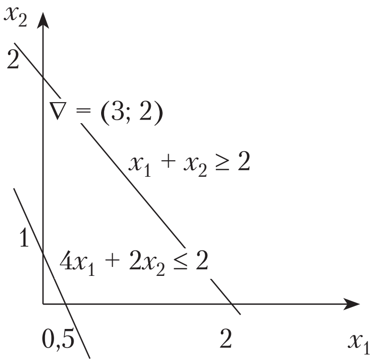


Рис. 2.6. **Противоречивость системы**

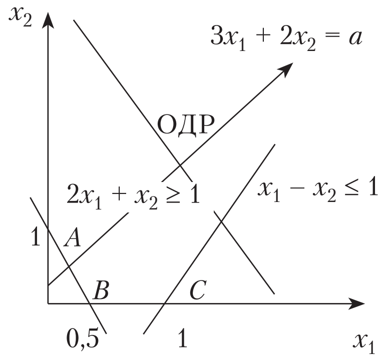


Рис. 2.7. **Неограниченность ЦФ ограничений**

шений — пустое множество, вторая — поскольку не существует конечного максимума на неограниченном множестве допустимых решений.

2.5. Симплексный метод решения задачи

Среди универсальных методов решения задач линейного программирования наиболее распространенным является *симплексный метод* (или *симплекс-метод*), разработанный американским ученым Дж. Данцигом. Суть этого метода заключается в том, что вначале получают допустимый вариант, удовлетворяющий всем ограничениям, но необязательно оптимальный (так называемое *начальное опорное решение*); оптимальность достигается последовательным улучшением исходного варианта за определенное число этапов (итераций). Нахождение начального опорного решения и переход к следующему опорному решению проводятся на основе применения рассмотренного выше метода Жордана — Гаусса для системы линейных уравнений канонической формы, в которой должна быть предварительно записана исходная ЗЛП; направление перехода от одного опорного решения к другому выбирается при этом на основе критерия оптимальности (целевой функции) исходной задачи. Симплекс-метод основан на следующих свойствах ЗЛП:

1) не существует локального экстремума, отличного от глобального. Другими словами, если экстремум есть, то он единственный;

2) множество всех допустимых решений (планов) задачи линейного программирования выпукло;

3) целевая функция ЗЛП достигает своего максимального (минимального) значения в угловой точке многогранника решений (в его вершине). Если целевая функция принимает свое оптимальное значение более чем в одной угловой точке, то она достигает того же значения в любой точке, являющейся выпуклой линейной комбинацией этих точек;

4) каждой угловой точке многогранника решений отвечает опорный план ЗЛП.

Рассмотрим две разновидности симплексного метода: симплекс-метод с естественным базисом и симплекс-метод с искусственным базисом (или M-метод).

Симплекс-метода с естественным базисом. Для применения этого метода ЗЛП должна быть сформулирована в канонической форме (2.12)—(2.14), причем матрица системы уравнений должна содержать единичную подматрицу размером $m \times m$. В этом случае очевиден начальный опорный план (неотрицательное базисное решение).

Для определенности предположим, что первые m векторов матрицы системы уравнений составляют единичную матрицу. Тогда первоначальный опорный план очевиден — $(b_1, b_2, ..., b_m, 0, ..., 0)$.

Проверка на оптимальность опорного плана проходит с помощью критерия (признака) оптимальности, переход к другому опорному плану — с помощью преобразований Жордана — Гаусса и с использованием признака оптимальности.

Полученный опорный план снова проверяется на оптимальность и т.д. Процесс заканчивается за конечное число шагов, причем на последнем шаге либо выявляется неразрешимость задачи (конечного оптимума нет), либо получается оптимальный опорный план и соответствующее ему оптимальное значение целевой функции.

Признак оптимальности заключается в следующих двух теоремах.

Теорема 2.3. Если для некоторого вектора, не входящего в базис, выполняется условие

$$\Delta_j = Z_j - c_j < 0, \text{ где } Z_j = \sum_{i=1}^{m} c_i \cdot a_{ij}, \; j = 1, 2, ..., n, \tag{2.22}$$

то можно получить новый опорный план, для которого значение целевой функции будет больше исходного; при этом могут быть два случая:

а) если все компоненты вектора, подлежащего вводу в базис, неположительны, то ЗЛП не имеет решения (целевая функция неограничена);

б) если имеется хотя бы одна положительная компонента у вектора, подлежащего вводу в базис, то можно получить новый опорный план.

Теорема 2.4. Если для всех векторов выполняется условие

$$\Delta_j = Z_j - c_j \geq 0,$$

то полученный опорный план является оптимальным.

На основании признака оптимальности в базис вводится вектор A_k, давший минимальную отрицательную величину симплекс-разности:

$$\Delta_k = \min\,(Z_j - c_j),\ j = 1, 2, ..., n;\ (Z_j - c_j) < 0.$$

Чтобы выполнялось условие неотрицательности значений опорного плана, выводится из базиса вектор A_r, который дает минимальное положительное оценочное отношение

$$Q = \min\,(b_i\,/\,a_{ik}) = b_r\,/\,a_{rk},\ a_{ik} > 0, i = 1, 2, ..., m. \qquad (2.23)$$

Если наименьшее значение Q достигается для нескольких базисных векторов, то, чтобы исключить возможность зацикливания (повторения базиса), можно применить следующий способ.

Вычисляются частные, полученные от деления всех элементов строк, давших одинаковое минимальное значение Q, на свои направляющие элементы. Полученные частные сопоставляются по столбцам слева направо, при этом учитываются и нулевые, и отрицательные значения. В процессе просмотра отбрасываются строки, в которых имеются большие отношения, и из базиса выводится вектор, соответствующий строке, в которой раньше обнаружится меньшее частное.

Для использования приведенной выше процедуры симплекс-метода к минимизации линейной функции $f(\bar{X})$ следует искать максимум функции $f_1(\bar{X}) = -f(\bar{X})$, затем полученный максимум взять с противоположным знаком. Это и будет искомый минимум исходной ЗЛП.

Следует заметить, что зацикливание в симплекс-методе может иметь место также и в случае так называемого вырождения. При m ограничениях ЗЛП значения m базисных переменных, как правило, положительны, а значения остальных $n - m$ свободных переменных всегда равны нулю. Однако возможно равенство нулю одной или нескольких базисных переменных; такой план называется *вырожденным*.

Вычисляемые по формуле (2.22) величины Δ_j называются *симплексными разностями*, или *оценками*, соответствующих переменных.

Так как в основе симплекс-метода лежит метод Жордана — Гаусса для системы линейных уравнений канонической формы ЗЛП, то при расчетах рекомендуется визуально проверять выполнение вытекающих из этого признаков: столбцы базисных векторов в симплексных таблицах имеют вид соответствующих столбцов единичной матрицы, а оценки базисных переменных всегда равны нулю.

Строка A_r называется *направляющей*, столбец A_k и элемент a_{rk} — *направляющими* (последний называют также *разрешающим* элементом и в сиплекс-таблицах выделяют рамкой).

Элементы вводимой строки, соответствующей направляющей строке, в новой симплекс-таблице вычисляются по формулам

$$a_{rj}^1 = a_{rj} / a_{rk}, \quad j = 1, 2, ..., n, \tag{2.24}$$

а элементы любой другой i-й строки пересчитываются по формулам

$$a_{ij}^1 = (a_{ij} \cdot a_{rk} - a_{rj} \cdot a_{ik}) / a_{rk}, \; i = 1, 2, ..., m; \; j = 1, 2, ..., n; \; i \neq r. \tag{2.25}$$

В соответствии с (2.24) и (2.25) значения базисных переменных нового опорного плана (показатели графы «план») рассчитываются по формулам

$$b_r^1 = b_r / a_{rk} \text{ для } i = r; \; b_i^1 = (b_i \cdot a_{rk} - b_r \cdot a_{ik}) / a_{rk}, \quad i = 1, 2, ..., m; \; i \neq r.$$

Таким образом, пересчет симплексных таблиц по формулам (2.24) и (2.25) проводится с помощью двух вычислительных процедур.

Процедура 1. Элементы вновь вводимой строки получаются путем деления элементов выводимой строки предыдущей таблицы на разрешающий (выделенный рамкой) элемент.

Процедура 2. Элементы других строк рассчитываются на базе преобразований Жордана — Гаусса, при этом целесообразно использовать так называемое правило прямоугольного треугольника. Суть этого правила заключается в том, что для расчета некоторого элемента новой таблицы (кроме элементов вновь вводимой строки) нужно выделить три числа:

1) соответствующий элемент (т.е. стоящий на том же месте) предыдущей таблицы;

2) элемент этой же строки предыдущей таблицы, стоящий в направляющем столбце;

3) элемент вновь введенной строки новой таблицы, стоящий в том же столбце, что и определяемый элемент. По своему расположению в таблицах эти три числа образуют прямоугольный треугольник, и для определения искомого элемента нужно из первого числа (стоящего в вершине прямого угла) вычесть произведение двух других чисел (стоящих в вершинах острых углов).

Процесс решения продолжают либо до получения оптимального плана, либо до установления неограниченности целевой функции. Если среди симплекс-разностей (оценок) оптимального плана нулевые только оценки, соответствующие базисным векторам, то это говорит о единственности оптимального плана. Если же нулевая оценка соответствует вектору, не входящему, то в общем случае это означает, что оптимальный план не единственный.

Рассмотрим алгоритм симплекс-метода на конкретной задаче.

Пример 2.7. Для производства продукции типа $П_1$ и $П_2$ предприятие использует два вида сырья: $С_1$ и $С_2$. Данные об условиях производства продукции приведены в табл. 2.1.

Таблица 2.1

Данные об условиях производства

Продукция / Сырье	Нормы расхода сырья, кг/ед.		Объемы запасов сырья, кг
	$П_1$	$П_2$	
$С_1$	1	3	300
$С_2$	1	1	150
Прибыль, у.е./ед. прод.	2	3	

Составить план производства по критерию «максимум прибыли».

Решение. Введем необходимые обозначения. Обозначим объем производства продукции $П_1$ через x_1, продукции $П_2$ через x_2. Таким образом, формально (математически) план производства (производственная программа) — это вектор $\bar{X} = (x_1, x_2)$. С учетом введенных обозначений математическая модель задачи по критерию «максимум прибыли» имеет вид

$$\max f(\bar{X}) = (2x_1 + 3x_2)$$

при ограничениях

$$\begin{gathered} x_1 + 3x_2 \le 300 \\ x_1 + x_2 \le 150 \\ x_1 \ge 0, x_2 \ge 0. \end{gathered}$$

Приведем эту ЗЛП к каноническому виду, введя дополнительные переменные x_3 и x_4:

$$\max f(\bar{X}) = 2x_1 + 3x_2 + 0x_3 + 0x_4$$

$$\begin{array}{ccccccccc} A_1 & & A_2 & & A_3 & & A_4 & & B \\ \begin{pmatrix} 1 \\ 1 \end{pmatrix} x_1 & + & \begin{pmatrix} 3 \\ 1 \end{pmatrix} x_2 & + & \begin{pmatrix} 1 \\ 0 \end{pmatrix} x_3 & + & \begin{pmatrix} 0 \\ 1 \end{pmatrix} x_4 & = & \begin{pmatrix} 300 \\ 150 \end{pmatrix}. \end{array}$$

или

$$\begin{gathered} x_1 + 3x_2 + x_3 = 300 \\ x_1 + x_2 + x_4 = 150 \\ x_{1,2,3,4} \ge 0 \end{gathered}$$

КЗЛП имеет необходимое число единичных столбцов, т.е. обладает очевидным начальным опорным планом (0, 0, 300, 150). Решение осуществляется симплекс-методом с естественным базисом с оформлением расчетов в симплекс-таблицах (табл. 2.2).

В исходной симплекс-таблице с номером 0 строка оценок Δ_j, $j = 1, 2, 3, 4$ определяется по приведенной выше формуле (2.22):

$$\Delta_1 = Z_1 - c_1 = 0 \cdot 1 + 0 \cdot 1 - 2 = -2; \Delta_2 = Z_2 - c_2 = \\ 0 \cdot 3 + 0 \cdot 1 - 3 = -3; \Delta_3 = \Delta_4 = 0.$$

Исходный опорный план (0, 0, 300, 150) не является оптимальным, так как среди оценок Δ_j имеются отрицательные. Переход к новому опорному плану осуществим, введя в базис вектор A_2, имеющий минимальную отрицательную оценку. Определяем вектор, выходящий из базиса:

$$Q = \min (300/3; 150/1) = 100,$$

т.е. вектор A_3 следует вывести из базиса. Разрешающим элементом является $a_{12} = 3$ (выделен рамкой). Переход к следующей симплекс-таблице осуществляем с помощью преобразований Жордана — Гаусса, при этом рекомендуется пользоваться двумя описанными выше вычислительными процедурами симплексного метода, включая правило прямоугольного треугольника.

Таблица 2.2

Решение ЗЛП в симплекс-таблицах

Номер симплекс-таблицы	Базис	c_j / c_i	План B	2	3	0	0	Q
				A_1	A_2	A_3	A_4	
0	A_3	0	300	1	$\boxed{3}$	1	0	100
	A_4	0	150	1	1	0	1	150
	Δ_j	—	0	−2	−3	0	0	—
I	A_2	3	100	1/3	1	1/3	0	300
	A_4	0	50	$\boxed{2/3}$	0	−1/3	1	75
	Δ_j	—	300	−1	0	1	0	—
II	A_2	3	75	0	1	1/2	−1/2	
	A_1	2	75	1	0	−1/2	3/2	
	Δ_j	—	375	0	0	1/2	3/2	—

Второй опорный план (0, 100, 0, 50) не оптимальный; переход к следующему опорному плану осуществим, вводя в базис вектор A_1 и выводя вектор A_4.

В симплекс-таблице II получен оптимальный опорный план, поскольку все симплекс-разности (оценки) $\Delta_j \geq 0, j = 1, 2, 3, 4$. Оптимальные значения переменных равны: $x_1^* = 75$, $x_2^* = 75$ (основные переменные), $x_3^* = 0$, $x_4^* = 0$ (дополнительные переменные). Максимальное значение целевой функции равно 375.

Таким образом, в рассмотренной задаче об оптимальном использовании ограниченных ресурсов, оптимальная производственная программа состоит в выпуске 75 ед. продукции первого вида и 75 ед. продукции второго вида. С этой программой связана максимальная прибыль от реализации готовой продукции — 375 у.е.

Симплекс-метод с искусственным базисом (*M*-метод)

Применяется в тех случаях, когда затруднительно найти первоначальный опорный план исходной задачи ЛП, записанной в канонической форме.

M-метод заключается в применении правил симплекс-метода к так называемой *M*-задаче. Она получается из исходной добавлением к левой части системы уравнений в канонической форме исходной ЗЛП таких искусственных единичных векторов с соответствующими неотрицательными искусственными переменными, чтобы вновь полученная матрица содержала систему единичных линейно-независимых векторов. В линейную форму исходной задачи добавляется в случае ее максимизации слагаемое, представляющее собой произведение числа $(-M)$ на сумму искусственных переменных, где M — достаточно большое положительное число.

В полученной задаче первоначальный опорный план очевиден. При применении к этой задаче симплекс-метода оценки Δ_j теперь будет зависеть «от буквы M». Для сравнения оценок нужно помнить, что M — достаточно большое положительное число, поэтому из базиса будут выводиться в первую очередь искусственные переменные.

В процессе решения M-задачи следует вычеркивать в симплекс-таблице искусственные векторы по мере их выхода из базиса. Если все искусственные векторы вышли из базиса, то получаем исходную задачу. Если оптимальное решение M-задачи содержит искусственные векторы или M-задача неразрешима, то исходная задача также неразрешима.

Путем преобразований число вводимых переменных, составляющих искусственный базис, может быть уменьшено до одной.

Пример 2.8. Найти максимум целевой функции: $\max f(\bar{X}) = 3x_1 + 2x_2 + x_3$ при условиях

$$\begin{gathered} 2x_1 + x_2 = 8, \\ x_1 + x_2 + x_3 = 6, \\ x_{1,2,3} \geq 0. \end{gathered}$$

Решение. Матрица условий содержит только один единичный вектор, добавим еще один искусственный вектор (искусственную неотрицательную переменную y_1 в первое ограничение):

$$P_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Получим следующую M-задачу: найти максимум целевой функции $\max f(\bar{X}) = 3x_1 + 2x_2 + x_3 - My_1$ при условиях

$$\begin{gathered} 2x_1 + x_2 + y_1 = 8, \\ x_1 + x_2 + x_3 = 6, \\ x_{1,2,3} \geq 0,\ y_1 \geq 0. \end{gathered}$$

M-задачу решаем симплекс-методом. Начальный опорный план (0, 0, 6, 8), решение проводим в симплекс-таблицах (табл. 2.3).

В начальной таблице наименьшее Δ_j соответствует вектору A_1 — он вводится в базис, а искусственный вектор P_1 из базиса выводится, так как ему отвечает наименьшее Q. Столбец, соответствующий P_1, из дальнейших симплексных таблиц вычеркивается.

Таблица 2.3

Решение *M*-задачи в симплекс-таблицах

Номер симплекс-таблицы	Базис	c_j / c_i	План B	3	2	1	$-M$	Q
				A_1	A_2	A_3	P_1	
0	P_1	$-M$	8	$\boxed{2}$	1	0	1	4
	A_3	1	6	1	1	1	0	6
	Δ_j	—	$-8M+6$	$-2M-2$	$-M-1$	0	0	—
I	A_1	3	4	1	0,5	0		—
	A_3	1	2	0	0,5	1		—
	Δ_j	—	14	0	0	0		—

Полученный новый опорный план является опорным планом исходной задачи. Для него все $\Delta_j \geq 0$, поэтому он является и оптимальным. Таким образом, получен оптимальный план исходной задачи (4, 0, 2) и максимальное значение целевой функции $\max f(\overline{X}) = 14$.

Пример 2.9. Решить ЗЛП:

$$\begin{gathered}\min f(\overline{X}) = 10x_1 - 5x_2 \\ 2x_1 - x_2 \geq 3 \\ x_1 + x_2 \geq 2 \\ x_1 + 2x_2 \geq -1 \\ x_{1,2} \geq 0\end{gathered}$$

Решение. Приведем ЗЛП к каноническому виду, перейдя к задаче «на максимум»:

$$\begin{gathered}\max f_1(\overline{X}) = -10x_1 + 5x_2 \\ 2x_1 - x_2 - x_3 = 3, \\ x_1 + x_2 - x_4 = 2, \\ -x_1 - 2x_2 + x_5 = 1, \\ x_j \geq 0, j = 1, 2, 3, 4, 5.\end{gathered}$$

Для нахождения опорного плана переходим к M-задаче:

$$\max g(\bar{X}, \bar{Y}) = -10x_1 + 5x_2 + 0x_3 + 0x_4 + 0x_5 - M(y_1 + y_2)$$
$$2x_1 - x_2 - x_3 + y_1 = 3$$
$$x_1 + x_2 - x_4 + y_2 = 2$$
$$-x_1 - 2x_2 + x_5 = 1$$
$$x_{1,2,3,4,5} \geq 0,\ y_1, y_2 \geq 0.$$

Дальнейшее решение проводим в симплекс-таблицах (табл. 2.4).

Таблица 2.4

Решение ЗЛП M-методом

Номер симплекс-таблицы	Базис	c_j / c_i	B	-10	5	0	0	0	$-M$	$-M$	Q
				A_1	A_2	A_3	A_4	A_5	P_1	P_2	
	P_1	$-M$	3	$\boxed{2}$	-1	-1	0	0	1	0	3/2
0	P_2	$-M$	2	1	1	0	-1	0	0	1	2
	A_5	0	1	-1	-2	0	0	1	0	0	—
—	Δ_j	—	$-5M$	$-3M+$ $+10$	-5	M	M	0	0	0	—
	A_1	-10	3/2	1	$-1/2$	$-1/2$	0	0		0	—
I	P_2	$-M$	1/2	0	$\boxed{3/2}$	1/2	-1	0		1	1/3
	A_5	0	5/2	0	$-5/2$	$-1/2$	0	1		0	—
—	Δ_j	—	$-M/2-$ -15	0	$-3M/2$	$-M/2+$ $+5$	M	0		0	—
	A_1	-10	5/3	1	0	$-1/3$	$-1/3$	0			—
II	A_2	5	1/3	0	1	1/3	$-2/3$	0			—
	A_5	0	10/3	0	0	0	$-5/3$	1			—
—	Δ_j	—	-15	0	0	5	0	0			—

В симплекс-таблице II получен опорный план исходной ЗЛП; поскольку все симплекс-разности $\Delta_j \geq 0$, $j = 1, 2, 3, 4, 5$, то этот план является и оптимальным, т.е. $x_1^* = 5/3$, $x_2^* = 1/3$ (исходные переменные), $x_3^* = 0$, $x_4^* = 0$, $x_5^* = 10/3$ (дополнительные переменные), при этом $\min f(\bar{X}) = -\max f_1(\bar{X}) = -(-15) = 15$.

При рассмотрении графического метода выделялись три особых случая решения ЗЛП. В симплекс-методе эти случаи определяются следующим образом.

1. Если найден оптимальный план и оценки всех свободных переменных строго больше нуля, то оптимальный план является *единственным*; если оценки некоторых свободных переменных в оптимальном плане равны нулю, то этот план будет *неединственным*, так как ввод этих переменных в базис не нарушает критерия оптимальности и не меняет оптимальное значение целевой функции. В соответствии с этим оптимальный план в табл. 2.2 является единственным, а в табл. 2.3 и 2.4 — неединственным (первый особый случай).

2. Если в процессе решения ЗЛП *M*-методом искусственные переменные не выводятся из базиса, это является свидетельством того, что область определения исходной ЗЛП является пустым множеством: в этом случае ЗЛП не имеет решения *ввиду противоречивости системы ограничений* (второй особый случай).

3. Если в направляющем столбце все элементы a_{ik} неположительны (см. 2.23), то это свидетельствует о незамкнутости области определения ЗЛП; в этом случае ЗЛП *не имеет решения ввиду неограниченности целевой функции* (третий особый случай).

Для автоматизации решения задач линейного программирования могут быть использованы стандартные офисные средства Microsoft Excel — надстройка **Поиск решения**, использующая симплексный метод (линейная оптимизация с помощью надстройки **Поиск решения** подробно рассмотрена, например, в литературе[1]).

Однако для корректного и эффективного использования программных средств необходимо знать основы линейного программирования, изложенные выше в данной главе.

Вопросы и задания

1. В чем суть принципа оптимальности в планировании и управлении?
2. Сформулируйте общую постановку задачи линейного программирования. Каковы особенности канонической формы записи этой задачи?
3. Дайте общую характеристику метода Жордана — Гаусса исследования систем линейных уравнений.

---

[1] *Мур Дж., Уэдерфорд Л. и др.* Экономическое моделирование в Microsoft Excel. М. : Издательский дом «Вильямс», 2004.

4. В чем заключается геометрическая интерпретация задачи линейного программирования?

5. Каковы основные этапы графического метода решения задач линейного программирования?

6. В чем суть симплекс-метода? На каких свойствах задач линейного программирования он основан?

7. Сформулируйте последовательность этапов практической реализации алгоритмов симплекс-метода при решении задач линейного программирования.

8. Когда возникает необходимость использования симплекс-метода с искусственным базисом (M-метода)? В чем суть этой модификации симплекс-метода?

Упражнения

1. Исследовать методом Жордана — Гаусса систему линейных уравнений; в случае совместности системы найти общее решение, некоторое частное небазисное решение, все базисные решения, указав при этом опорные решения:

$$\text{а) } \begin{aligned} &4x_1 - x_2 + 2x_3 - 3x_4 = 2, \\ &2x_1 + 3x_2 - x_3 + x_4 = 5, \\ &x_1 - 4x_2 + 3x_3 - 4x_4 = -3; \end{aligned} \qquad \text{б) } \begin{aligned} &x_1 + x_2 - x_3 + 2x_4 = 2, \\ &-x_1 + x_2 - 3x_3 - x_4 = 1, \\ &3x_1 - x_2 + 5x_3 + 4x_4 = 3; \end{aligned}$$

$$\text{в) } \begin{aligned} &x_1 + 2x_2 + x_3 + x_4 - 3x_5 = 4, \\ &3x_1 + 4x_2 - 2x_3 - x_4 + x_5 = 6, \\ &-x_1 + 4x_3 + 3x_4 - 7x_5 = 2, \\ &2x_1 + 6x_2 - x_3 + x_5 = 6. \end{aligned}$$

2. Решить графическим методом следующие задачи линейного программирования:

$$\text{а) } \begin{gathered} \max f(\overline{X}) = x_1 + 3x_2 \\ -x_1 + x_2 \le 3 \\ x_1 + x_2 \le 7 \\ 3x_1 + x_2 \le 15 \\ x_{1,2} \ge 0 \end{gathered} \qquad \text{б) } \begin{gathered} \min f(\overline{X}) = -6x_1 + 9x_2 \\ x_1 + 3x_2 \ge 9 \\ -2x_1 + x_2 \le 5 \\ 2x_1 - 3x_2 \le 0 \\ x_{1,2} \ge 0 \end{gathered}$$

$$\text{в) } \begin{gathered} \min f(\overline{X}) = 2x_1 - x_2 \\ 2x_1 + 3x_2 \ge 12 \\ x_1 - x_2 \le 1 \\ 0 \le x_1 \le 5 \\ x_2 \ge 0 \end{gathered}$$

3. Решить симплексным методом следующие задачи линейного программирования:

а) $\max f(\bar{X}) = -x_1 + x_2$

$$\begin{gathered} 2x_1 + x_2 \le 15 \\ x_1 + x_2 \le 3 \\ x_2 \le 6 \\ x_{1,2} \ge 0 \end{gathered}$$

б) $\min f(\bar{X}) = -2x_1 - 3x_2$

$$\begin{gathered} 3x_1 + 3x_2 \le 15 \\ x_1 + 3x_2 \le 9 \\ x_1 \le 4 \\ x_{1,2} \ge 0 \end{gathered}$$

в) $\min f(\bar{X}) = 2x_1 - 3x_2$

$$\begin{gathered} 2x_1 - x_2 - x_3 \ge 3 \\ x_1 - x_2 + x_3 \ge 2 \\ 0 \le x_1 \le 5 \\ x_{1,2,3} \ge 0 \end{gathered}$$

4. Для выпуска четырех видов продукции требуются затраты сырья, рабочего времени и оборудования. Исходные данные приведены в таблице.

Тип ресурса	Нормы затрат ресурсов на единицу продукции				Наличие ресурсов
	1	2	3	4	
Сырье	3	5	2	4	60
Рабочее время	22	14	18	30	400
Оборудование	10	14	8	16	128
Прибыль на ед. продукции	30	25	8	16	—

Сформулировать экономико-математическую модель задачи на максимум прибыли и найти оптимальный план выпуска продукции.

Глава 3
ОПТИМИЗАЦИОННЫЕ ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МОДЕЛИ

— Теория двойственности в анализе оптимальных решений экономических задач

— Транспортная задача

— Целочисленное программирование

— Задачи многокритериальной оптимизации

— Нелинейное и динамическое программирование; понятие об имитационном моделировании

— Модели сетевого планирования и управления

— В результате изучения материала этой главы студенты должны:

***знать*:**

- основные понятия теории двойственности линейного программирования;
- методы формулировки и решения транспортных задач;
- особенности методов целочисленного, нелинейного, динамического программирования и имитационного моделирования;
- принципы и методы использования сетевых моделей в экономике;

***уметь*:**

- использовать при анализе оптимальных решений теоремы двойственности;
- решать закрытые и открытые транспортные задачи;
- применять методы целочисленного и многокритериального программирования при решении практических задач;
- строить сетевые модели и оптимизировать их;

***владеть*:**

- понятийным аппаратом теории двойственности;
- практическими навыками постановки и решения транспортных задач, задач целочисленного программирования и многокритериальной оптимизации, задач сетевого планирования и управления.

3.1. Теория двойственности в анализе оптимальных решений экономических задач

Рассмотрим основные понятия и выводы специального раздела линейного программирования — теория двойственности. В главе 2 показано, что любую задачу линейного программирования можно записать следующим образом:

$$f(\bar{X}) = \sum_{j=1}^{n} c_i \cdot x_j \to \max; \tag{3.1}$$

$$\sum_{j=1}^{n} a_{ij} \cdot x_j \le b_j,\ i = \overline{1,m}; \tag{3.2}$$

$$x_j \ge 0,\ j = \overline{1,n}. \tag{3.3}$$

В этой главе для большей наглядности используются записи типа $f(\bar{X}) \to \max(\min)$, эквивалентные записям $\max(\min) f(\bar{X})$.

С каждой задачей линейного программирования тесно связана другая линейная задача, называемая *двойственной*; первоначальная задача называется *исходной* или прямой.

Связь исходной и двойственной задач заключается, в частности, в том, что решение одной из них может быть получено непосредственно из решения другой.

Хорошо разработанный математический аппарат линейного программирования позволяет не только получать с помощью эффективных вычислительных процедур оптимальный план, но и делать ряд экономически содержательных выводов, основанных на свойствах задачи, двойственной к исходной ЗЛП. Переменные двойственной задачи y_i называют *объективно обусловленными оценками* или *двойственными оценками*. Модель двойственной задачи имеет вид

$$g(\bar{Y}) = \sum_{i=1}^{m} b_i \cdot y_i \to \min; \tag{3.4}$$

$$\sum_{i=1}^{m} a_{i,j} \cdot y_i \ge c_j,\ j = \overline{1,n}; \tag{3.5}$$

$$y_i \ge 0,\ i = \overline{1,m}. \tag{3.6}$$

Каждая из задач двойственной пары фактически является самостоятельной задачей линейного программирования и может быть решена независимо от другой. Однако при определении симплексным методом оптимального плана одной из задач находится решение и другой задачи.

Двойственная задача, по отношению к исходной, составляется согласно следующим правилам:

1) целевая функция исходной задачи (3.1)—(3.3) формулируется на максимум, а целевая функция двойственной задачи (3.4)—(3.6) — на минимум, при этом в задаче на мак-

симум все неравенства в функциональных ограничениях имеют вид «≤», а в задаче на минимум — вид «≥»;

2) Матрица

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ . & . & . & . & . \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix},$$

составленная из коэффициентов при неизвестных в системе ограничений (3.2) исходной задачи, и аналогичная матрица

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ . & . & . & . \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}$$

в двойственной задаче получаются друг из друга транспонированием;

3) число переменных в двойственной задаче равно числу функциональных ограничений (3.2) исходной задачи, а число ограничений в системе (3.5) двойственной задачи — числу переменных в исходной задаче;

4) коэффициентами при неизвестных в целевой функции (3.4) двойственной задачи являются свободные члены в системе (3.2) ограничений исходной задачи, а правыми частями в ограничениях (3.5) двойственной задачи — коэффициенты при неизвестных в целевой функции (3.1) исходной задачи;

5) каждому ограничению одной задачи соответствует переменная другой задачи: номер переменной совпадает с номером ограничения; при этом ограничению, записанному в виде неравенства «≤», соответствует переменная, связанная условием неотрицательности. Если функциональное ограничение исходной задачи является равенством, то соответствующая переменная двойственной задачи может принимать как положительные, так и отрицательные значения.

Математические модели пары двойственных задач могут быть симметричными и несимметричными. В *несимметричных* двойственных задачах система ограничений исходной задачи задается в виде равенств, а двойственной — в виде неравенств, причем в последней переменные могут быть и отрицательными. В *симметричных* задачах система ограничений как исходной, так и двойственной задачи задается неравенствами, причем на двойственные переменные налагается условие неотрицательности. В дальнейшем мы будем рассматривать только симметричные взаимодвойственные задачи линейного программирования.

Итак, согласно теории линейного программирования каждой ЗЛП вида (3.1)—(3.3) соответствует двойственная ей ЗЛП: (3.4)—(3.6). Основные утверждения о взаимодвойственных задачах содержатся в двух следующих теоремах.

Первая теорема двойственности. Для взаимодвойственных ЗЛП имеет место один из взаимоисключающих случаев.

1. В прямой и двойственной задачах имеются оптимальные решения, при этом значения целевых функций на оптимальных решениях совпадают: $\max f(\bar{X}) = \min g(\bar{Y})$.

2. В прямой задаче допустимое множество не пусто, а целевая функция на этом множестве не ограничена сверху. При этом у двойственной задачи будет пустое допустимое множество.

3. В двойственной задаче допустимое множество не пусто, а целевая функция на этом множестве не ограничена снизу. При этом у прямой задачи допустимое множество оказывается пустым.

4. Обе из рассматриваемых задач имеют пустые допустимые множества.

Вторая теорема двойственности (теорема о дополняющей нежесткости). Пусть $\bar{X} = (x_1, x_2, ..., x_n)$ — допустимое решение прямой задачи (3.1)—(3.3), а $\bar{Y} = (y_1, y_2, ..., y_m)$ — допустимое решение двойственной задачи (3.4)—(3.6). Для того чтобы они были оптимальными решениями соответствующих взаимодвойственных задач (3.1)—(3.3) и (3.4)—(3.6), необходимо и достаточно, чтобы выполнялись следующие соотношения:

$$y_i\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right) = 0,\ i = \overline{1,m}; \tag{3.7}$$

$$x_j\left(\sum_{i=1}^{m} a_{ij}y_i - c_j\right) = 0,\ j = \overline{1,n}. \tag{3.8}$$

Условия (3.7) и (3.8) позволяют, зная оптимальное решение одной из взаимодвойственных задач, найти оптимальное решение другой задачи.

Рассмотрим еще одну теорему, выводы которой будут использованы в дальнейшем.

Теорема об оценках (третья теорема двойственности). Значения переменных y_i в оптимальном решении двойственной задачи представляют собой оценки влияния свободных членов b_i системы ограничений — неравенств прямой задачи на величину целевой функции этой задачи:

$$\Delta f(\bar{X}) = \Delta b_i y_i. \tag{3.9}$$

Решая ЗЛП симплексным методом, мы одновременно решаем двойственную ЗЛП. Значения переменных двойственной задачи y_i в оптимальном плане называют, как выше уже отмечено, объективно обусловленными, или двойственными оценками.

Рассмотрим экономическую интерпретацию двойственной задачи на следующем примере.

Пример 3.1 (задача оптимального использования ресурсов). Пусть для выпуска четырех видов продукции P_1, P_2, P_3, P_4 на предприятии используют три вида сырья S_1, S_2 и S_3. Объемы выделенного сырья, нормы расхода сырья и прибыль на единицу продукции при изготовлении каждого вида продукции приведены в табл. 3.1. Требуется определить план выпуска продукции, обеспечивающий наибольшую прибыль.

Решение. Составим экономико-математическую модель задачи оптимального использования ресурсов на максимум прибыли. В качестве неизвестных примем объем выпуска продукции j-го вида x_j (j = 1, 2, 3, 4).

Таблица 3.1

Исходные данные

Вид сырья	Запасы сырья	Вид продукции			
		P_1	P_2	P_3	P_4
S_1	35	4	2	2	3
S_2	30	1	1	2	3
S_3	40	3	1	2	1
Прибыль		14	10	14	11

Модель задачи по критерию «максимум прибыли»:

$$f(\bar{X}) = 14x_1 + 10x_2 + 14x_3 + 11x_4 \to \max$$
$$4x_1 + 2x_2 + 2x_3 + 3x_4 \le 35$$
$$x_1 + x_2 + 2x_3 + 3x_4 \le 30$$

$$3x_1 + x_2 + 2x_3 + x_4 \le 40$$
$$x_j > 0,\ j = 1, 2, 3, 4.$$

Теперь сформулируем двойственную задачу. Пусть некая организация решила закупить все ресурсы рассматриваемого предприятия. При этом необходимо установить оптимальную цену на приобретаемые ресурсы y_1, y_2, y_3 исходя из следующих объективных условий:

1) покупающая организация старается минимизировать общую стоимость ресурсов;

2) за каждый вид ресурсов надо уплатить не менее той суммы, которую хозяйство может выручить при переработке сырья в готовую продукцию.

Согласно первому условию общая стоимость сырья выразится величиной $g\{\bar{Y}) = 35y_1 + 30y_2 + 40y_3 \to \min$. Согласно второму требованию вводятся ограничения: на единицу первого вида продукции расходуются четыре единицы первого ресурса ценой y_1, одна единица второго ресурса ценой y_2 и три единицы третьего ресурса ценой y_3. Стоимость всех ресурсов, расходуемых на производство единицы первого вида продукции, равна $4y_1 + y_2 + 3y_3$ и должна составлять не менее 14, т.е. $4y_1 + y_2 + 3y_3 \ge 14$.

В результате аналогичных рассуждений относительно производства второго, третьего и четвертого видов продукции получаем систему неравенств:

$$4y_1 + y_2 + 3y_3 \ge 14,$$
$$2y_1 + y_2 + y_3 \ge 10,$$
$$2y_1 + 2y_2 + 2y_3 \ge 14,$$
$$3y_1 + 3y_2 + y_3 \ge 11.$$

По экономическому смыслу цены неотрицательные:

$$y_1 \ge 0,\ y_2 \ge 0,\ y_3 \ge 0.$$

Получили симметричную пару взаимодвойственных задач. В результате решения данной задачи симплексным методом получен оптимальный план $\bar{X} = (0; 5; 12{,}5; 0)$; $\bar{Y} = (3; 4; 0)$. На рис. 3.1 приведен результат решения задачи с использованием надстройки **Поиск решения** Excel. Жирным шрифтом выделены оптимальные значения $\bar{X}$ и $\bar{Y}$.

Экономический смысл первой теоремы двойственности следующий. План производства X и набор оценок ресурсов Y оказываются оптимальными тогда и только тогда, когда прибыль от реализации продукции, определенная при известных заранее ценах продукции $c_1, c_2, ..., c_n$, равна затратам на ресурсы по «внутренним» (определяемым только из решения задачи) ценам ресурсов y_1, y_2,..., y_m. Для всех же других планов $\bar{X}$ и $\bar{Y}$ обеих задач прибыль от продукции всегда меньше (или равна) стоимости за-

Ячейки переменных

Ячейка	Имя	Окончательное значение	Приведенная стоимость	Целевая функция Коэффициент	Допустимое увеличение	Допустимое уменьшение
B1		**0**	–2	14	2	1E+30
C1		**5**	0	10	4	0.666666667
D1		**12.5**	0	14	2	4
E1		**0**	–10	11	10	1E+30

Ограничения

Ячейка	Имя	Окончательное значение	Теневая цена	Ограничение Правая сторона	Допустимое увеличение	Допустимое уменьшение
F3		35	**3**	35	25	5
F4		30	**4**	30	5	12.5
F5		30	**0**	40	1E+30	10

Рис. 3.1. **Отчет по устойчивости Microsoft Excel**

траченных ресурсов: $f(\bar{X}) \le g(\bar{Y})$, т.е. ценность всей выпущенной продукции не превосходит суммарной оценки имеющихся ресурсов. Значит, величина $g(\bar{Y}) - f(\bar{X})$ характеризует производственные потери в зависимости от рассматриваемой производственной программы и выбранных оценок ресурсов.

Из первой теоремы двойственности следует, что при оптимальных производственной программе и векторе оценок ресурсов производственные потери равны нулю.

Экономический смысл первой теоремы двойственности можно интерпретировать и так: предприятию безразлично, производить ли *продукцию* по оптимальному плану $\bar{X}$ и получить максимальную прибыль либо продать ресурсы по оптимальным ценам $\bar{Y}$ и возместить от продажи равные ей минимальные затраты на ресурсы.

Из второй теоремы двойственности в данном случае следуют такие требования на оптимальную производственную программу $\bar{X} = (x_1, x_2, ..., x_n)$ и оптимальный вектор оценок $\bar{Y} = (y_1, y_2, ..., y_m)$:

$$\text{если } y_i > 0, \text{ то } \sum_{j=1}^{n} a_{i,j} x_j = b_i, \ i = \overline{1,m}; \qquad (3.10)$$

$$\text{если } \sum_{j=1}^{n} a_{i,j} x_j < b_i, \text{ то } y_i = 0, \ i = \overline{1,m};\,;$$

$$\text{если } x_j > 0, \text{ то } \sum_{i=1}^{m} a_{i,j} y_i = c_j, \ j = \overline{1,n}; \qquad (3.11)$$

$$\text{если } \sum_{i=1}^{m} a_{i,j} y_i > c_j, \text{ то } x_j = 0, \ j = \overline{1,n}.$$

Условия (3.10) можно интерпретировать так: если оценка y_i единицы ресурса i-го вида положительна, то при оптимальной производственной программе этот ресурс используется полностью; если же ресурс используется не полностью, то его оценка равна нулю.

Из условия (3.11) следует, что если j-й вид продукции вошел в оптимальный план, то он в оптимальных оценках не убыточен; если же j-й вид продукции убыточен, то он не войдет в план, не будет выпускаться.

Кроме нахождения оптимального решения должно быть обеспечено получение дополнительной информации о возможных изменениях решения при изменении параметров системы. Эту часть исследования обычно называют анализом модели на чувствительность. Он необходим, например, в тех случаях, когда некоторые характеристики исследуемой системы не поддаются точной оценке.

Экономико-математический анализ решений осуществляется в двух основных направлениях: в виде вариантных расчетов по моделям с сопоставлением различных вариантов плана и в виде анализа каждого из полученных решений с помощью двойственных оценок. Вариантные расчеты могут осуществляться при неизменной структуре самой модели (постоянном составе неизвестных, способов производства, ограничений задачи и одинаковом критерии оптимизации), но с изменением численной величины конкретных показателей модели. Вариантные расчеты могут проводиться также при варьировании элементов самой модели: изменении критерия оптимизации, добавлении новых ограничений на ресурсы или на способы производства их использования, расширения множества вариантов и т.д.

Одно из эффективных средств экономико-математического анализа — использование объективно обусловленных оценок оптимального плана. Такого рода анализ базируется на свойствах двойственных оценок. Выше мы установили общие математические свойства двойственных оценок для задач на оптимум любой экономической природы. Однако экономическая интерпретация этих оценок может быть совершенно различной для разных задач.

Перейдем к рассмотрению конкретных экономических свойств оценок y_i оптимального плана. Сначала перечислим эти свойства, а затем проиллюстрируем их конкретными примерами.

Свойство 1. Оценки как мера дефицитности ресурсов и продукции.

Свойство 2. Оценки как мера влияния ограничений на функционал.

Свойство 3. Оценки как инструмент определения эффективности отдельных вариантов.

Свойство 4. Оценки как инструмент балансирования суммарных затрат и результатов.

Пример 3.2 (задача о планировании выпуска тканей). Пусть некоторая фабрика выпускает три вида тканей, причем суточное плановое задание составляет не менее 90 м тканей первого вида, 70 м — второго и 60 м — третьего. Суточные ресурсы следующие: 780 единиц производственного оборудования, 850 единиц сырья и 790 единиц электроэнергии, расход которых на один метр ткани представлен в табл. 3.2.

Таблица 3.2

Нормы расхода ресурсов

Ресурсы	Ткани		
	I	II	III
Оборудование	2	3	4
Сырье	1	4	5
Электроэнергия	3	4	2

Цена за один метр ткани вида I равна 80 денежным единицам, II — 70 денежным единицам, III — 60 денежным единицам.

Необходимо определить, сколько метров ткани каждого вида следует выпустить, чтобы общая стоимость выпускаемой продукции была максимальной.

Решение. Составим модель задачи. Введем следующие обозначения. Неизвестными в задаче являются объемы выпуска ткани каждого вида:

x_1 — количество метров ткани вида I;

x_2 — количество метров ткани вида II;

x_3 — количество метров ткани вида III.

$$f(\bar{X}) = \sum_{j=1}^{n} c_j x_j \to \max;$$

$$\sum_{j=1}^{n} a_{i,j} x_j \le b_i,\ i = \overline{1,m};$$

$$x_j \ge T_j,\ j = \overline{1,n};$$

$$x_j \ge 0.$$

С учетом имеющихся данных модель примет вид:

$$f(\bar{X}) = 80x_1 + 70x_2 + 60x_2 \to \max$$

$$\left.\begin{array}{l} 2x_1 + 3x_2 + 4x_3 \le 780 \\ x_1 + 4x_2 + 5x_3 \le 850 \\ 3x_1 + 4x_2 + 2x_3 \le 790 \end{array}\right\} \text{Ограничение по ресурсам}$$

$$\left.\begin{array}{l} x_1 \ge 90 \\ x_2 \ge 70 \\ x_3 \ge 60 \end{array}\right\} \text{Плановое задание}$$

В результате решения задачи симплексным методом получен следующий оптимальный план: максимум общей стоимости продукции $f(\bar{X})$ = 19 075 ден. ед. достигается при

x_1 = 112,5 м — оптимальный объем выпуска ткани вида I;

x_2 = 70 м — оптимальный объем выпуска ткани вида II;

x_3 = 86,25 м — оптимальный объем выпуска ткани вида III.

Решение двойственной задачи получим с использованием теорем двойственности. Введем обозначения:

y_1 — двойственная оценка ресурса «оборудование»;

y_2 — двойственная оценка ресурса «сырье»;

y_3 — двойственная оценка ресурса «электроэнергия»;

y_4 — двойственная оценка ткани вида I;

y_5 — двойственная оценка ткани вида II;

y_6 — двойственная оценка ткани вида III.

Модель двойственной задачи имеет вид:

$$g(\bar{Y}) = 780y_1 + 850y_2 + 790y_3 + 90y_4 + 70y_5 + 60y_6 \to \min$$
$$2y_1 + y_2 + 3y_3 + y_4 \ge 80,$$
$$3y_1 + 4y_2 + 4y_3 + y_5 \ge 70,$$
$$4y_1 + 5y_2 + 2y_3 + y_6 \ge 60,$$
$$y_{1,2,3} \ge 0,\ y_{4,5,6} \le 0.$$

Из соотношений второй теоремы двойственности вытекают следующие условия:

для каждого ресурса:

$$\text{если } \sum_{j=1}^{n} a_{ij}x_j < b_i, \text{ то } y_i = 0;$$

$$\text{если } y_i > 0, \text{ то } \sum_{j=1}^{n} a_{ij}x_j = b_i.$$

для задания по выпуску продукции:

$$\text{если } x_j > T_j, \text{ то } y_{m+j} = 0;$$

$$\text{если } y_{m+j} < 0, \text{ то } x_j = T_j. \qquad (3.12)$$

Для нашего примера в этих соотношениях $m = 3$ (количество типов ресурсов).

Подставим значения $x_1 = 112{,}5$, $x_2 = 70$ и $x_3 = 86{,}25$ в ограничения прямой задачи:

$2 \cdot 112{,}5 + 3 - 70 + 4 \cdot 86{,}25 = 780,$
$112{,}5 + 4 - 70 + 5 \cdot 86{,}25 = 823{,}75 < 850,$
$3 - 112{,}5 + 4 - 70 + 2 - 86{,}25 = 790,$
$112{,}5 > 90,$
$70 = 70,$
$86{,}25 > 60.$

Суточные ресурсы по оборудованию и электроэнергии использованы полностью. Сырье используется не полностью, имеется остаток в размере 850 − 823,75 = 26,25 (кг). План выпуска по тканям вида I и III перевыполнен.

Из второй теоремы двойственности вытекает, что y_2, y_4 и y_6 равны нулю. Остается найти значения y_1, y_3 и y_5. Так как x_1, x_2 и x_3 — больше нуля, то все три ограничения двойственной задачи выполняются как равенства:

$$\begin{gathered} 2y_1 + y_2 + 3y_3 + y_4 = 80, \\ 3y_1 + 4y_2 + 4y_3 + y_5 = 70, \\ 4y_1 + 5y_2 + 2y_3 + y_6 = 60. \end{gathered}$$

Учитывая, что $y_2 = y_4 = y_6 = 0$, имеем:

$$\begin{gathered} 2y_1 + 3y_3 = 80, \\ 3y_1 + 4y_3 + y_5 = 70, \\ 4y_1 + 2y_3 = 60, \end{gathered}$$

откуда $y_1 = 2{,}5$; $y_3 = 25$; $y_5 = -37{,}5$.

Подставив значения неизвестных в целевую функцию двойственной задачи, проверим, выполняется ли условие $f(\bar{X}) = g(\bar{Y})$ для оптимального плана:

$$g(\bar{Y}) = 780 \cdot 2{,}5 + 850 \cdot 0 + 790 \cdot 25 + 90 \cdot 0 - 70 \cdot 37{,}5 + 60 \cdot 0 = 19\,075.$$

Условие первой теоремы двойственности выполняется, следовательно, рассмотренный план выпуска тканей и соответствующая ему система оценок ресурсов и продукции оптимальны.

Экономическое истолкование оценок есть интерпретация их общих экономико-математических свойств применительно к конкретному содержанию задачи. По условию (3.10) не использованный полностью в оптимальном плане ресурс получает нулевую оценку. Нулевая оценка ресурса свидетельствует о его недефицитности. Ресурс недефицитен не из-за его неограниченных запасов (они ограничены величиной b_i), а из-за невозможности его полного использования в оптимальном плане. Так как суммарный расход недефицитного ресурса меньше его общего количества, то план производства им не лимитируется. Данный ресурс не препятствует и дальше максимизировать целевую функцию $f(\bar{X})$.

Ограничивают целевую функцию дефицитные ресурсы, в данном примере — оборудование и электроэнергия. Они полностью использованы в оптимальном плане. По условию (3.10) оценка таких ресурсов положительна ($y_1 = 2{,}5$; $y_3 = 25$). Рассмотрим теперь понятие дефицитности продукции. По условию (3.12) нулевую оценку ($y_4 = 0$, $y_6 = 0$) получает продукция, задания по выпуску которой в оптимальном плане перевыполняются. Очевидно, перевыполнение плана целесообразно по выгодной продукции (ткани I и III видов), т.е. такой, производство которой способствует достижению максимума критерия оптимальности. Размеры производства такой выгодной продукции определяются не величиной задания на выпуск (T_j) (в оптимальном плане они перекрыты), а ограниченностью дефицитных ресурсов. Эту продукцию выпускают как можно больше, пока хватит ресурсов. Выпуск выгодной продукции лимитируется не только фактом ограниченности дефицитных ресурсов, но и тем, что часть дефицитных ресурсов требуется выделить на обеспечение выпуска невыгодной продукции в соответствии с плановыми заданиями. По условию (3.12) отрицательную оценку ($y_5 = -37{,}5$) получает продукция, задания, по выпуску которой, не перевыполняются. Так как по условию задачи ($X_j > T_j$) плановые задания должны быть обязательно выполнены, то продукция делится на выгодную (виды I и III ткани) и невыгодную (вид II ткани).

Если в ограничение двойственной задачи, относящееся к виду II ткани,

$$3\,y_1 + 4y_2 + 4\,y_3 + y_5 \geq 70$$

подставить полученные значения двойственных оценок, то получаем

$$3 - 2{,}5 + 4 \cdot 0 + 4 \cdot 25 - 37{,}5 = 70,$$
$$107{,}5 - 37{,}5 = 70,$$

т.е. стоимость ресурсов, затраченных на один метр ткани вида II, составляет 107,5 денежной единицы, и это на 37,5 денежной единицы больше цены одного метра ткани этого вида. Таким образом, вид II ткани убыточен для фабрики: на каждом выпущенном метре ткани этого вида фабрика теряет 37,5 денежной единицы.

В соответствии с критерием оптимальности плана, в зависимости от того перевыполняется план выпуска или нет, выпуск ткани вида II поглощает часть дефицитных ресурсов, чем сдерживает рост выпуска выгодной продукции, а тем самым и рост целевой функции.

Оценка ресурса показывает, насколько изменится критерий оптимальности при изменении количества данного ресурса на единицу. Для недефицитного ресурса оценка равна нулю, поэтому изменение его величины не повлияет на критерий оптимальности. Дефицитность ресурса измеряется вкладом единицы ресурса в изменение целевой функции.

Влияние ограничений по выпуску продукции на критерий оптимальности противоположно влиянию ограничений по ресурсам. Если продукция невыгодна (вид II ткани, $y_5 = -37{,}5$), то увеличение плановых заданий по ее выпуску ведет к уменьшению выпуска выгодной продукции и ухудшает план. Наоборот, уменьшение плановых заданий по невыгодной продукции позволяет снизить ее выпуск, перебросить сэкономленные ресурсы на дополнительный сверхплановый выпуск выгодных видов продукции, что увеличивает значение целевой функции. Изменение плановых заданий по выгодной продукции не изменяет значения целевой функции.

Перейдем к анализу модели на чувствительность.

Пример 3.3. На основании информации, приведенной в табл. 3.3, составить план производства, максимизирующий объем прибыли.

Таблица 3.3

Ресурсы	Затраты ресурсов на единицу продукции		Наличие ресурсов
	A	*A*	
Труд	2	4	2000
Сырье	4	1	1400
Оборудование	2	1	800
Прибыль на единицу продукции	40	60	

Решение. Экономико-математическая модель задачи будет иметь вид:

$f(\bar{X}) = 40x_1 + 60x_2 \to \max;$
$2x_1 + 4x_2 \le 2000,$
$4x_1 + x_2 \le 1400,$
$2x_1 + x_2 \le 800,$
$x_{1,2} \ge 0.$

В результате решения задачи симплексным методом был получен следующий оптимальный план:

$\bar{X} = (200; 400; 0; 200; 0),$
$f(\bar{X}) = 40x_1 + 60x_2 = 40 \cdot 200 + 60 \cdot 400 = 32\,000,$
$\bar{Y} = (40/3; 0; 20/3),$
$g(\bar{Y}) = 2000y_1 + 1400y_2 + 800y_3 = 2000 \cdot 40/3 + 800 \cdot 20/3 = 32\,000.$

Соответствующий отчет об устойчивости решения представлен на рис. 3.2.

Ячейки переменных

Ячейка	**Имя**	**Окончательное значение**	**Приведенная стоимость**	**Целевая функция Коэффициент**	**Допустимое увеличение**	**Допустимое уменьшение**
\$A\$2		**200**	0	40	80	10
\$B\$2		**400**	0	60	20	40

Ограничения

Ячейка	**Имя**	**Окончательное значение**	**Теневая цена**	**Ограничение Правая сторона**	**Допустимое увеличение**	**Допустимое уменьшение**
\$C\$4		2000	**13.33333333**	2000	**1200**	**600**
\$C\$5		1200	**0**	1400	**1E+30**	**200**
\$C\$6		800	**6.666666667**	800	**85.71428571**	**300**

Рис. 3.2. **Отчет об устойчивости**

После того как оптимальное решение получено, выявляется его чувствительность к определенным изменениям исходной модели. В нашей задаче, например, может представить интерес то, как повлияет на оптимальное решение изменение запасов сырья и изменение прибыли от единицы продукции. В связи с этим логично выяснить.

1. Увеличение объемов какого вида ресурсов наиболее выгодно?

2. Насколько можно увеличить запас сырья для улучшения полученного оптимального значения целевой функции?

3. Каков диапазон изменения того или иного коэффициента целевой функции, при котором не происходит изменения оптимального решения?

4. Целесообразность включения в план новых изделий.

Постараемся последовательно ответить на все поставленные вопросы.

1. Ценность ресурсов. В примере 3.3 объективно обусловленные оценки ресурса «труд» равны 40/3 ($y_1 = 40/3$), «сырье» — 0 ($y_2 = 0$), «оборудование» — 20/3 ($y_3 = 20/3$).

Дефицитный ресурс, полностью используемый в оптимальном плане ($\sum a_{i,j} x_j = b_i$), имеет положительную оценку ($y_i > 0$); недефицитный, не полностью используемый ресурс (для которого $\sum a_{i,j} x_j < b_i$), имеет нулевую оценку ($y_i = 0$). В примере «сырье» не является дефицитным ресурсом:

$$4x_1 + x_2 \le 1400,$$
$$4 \cdot 200 + 400 = 1200 < 1400 = b_2,$$
$$y_2 = 0;$$

а «труд» и «оборудование» — дефицитные ресурсы:

$$2x_1 + 4x_2 < 2000,$$
$$2 \cdot 200 + 4 \cdot 400 = 2000 = b_1,\ y_1 = 40/3;$$
$$2x_1 + x_2 \le 800,$$
$$2 \cdot 200 + 400 = 800 = b_3,\ y_3 = 20/3.$$

Чем выше величина оценки y_i, тем острее дефицитность i-го ресурса. В примере «труд» более дефицитен, чем «оборудование»: 40/3 > 20/3. Наиболее выгодно увеличение объемов ресурса труда.

Заметим, что ценность различных видов сырья нельзя отождествлять с действительными ценами, по которым осуществляется его закупка. В данном случае речь идет о некоторой мере, имеющей экономическую природу, которая характеризует ценность сырья только относительно полученного оптимального решения.

2. Чувствительность решения к изменению запасов сырья. Предположим, что запас сырья ресурса «труд» изменился на 12 единиц, т.е. теперь он составляет 2000 + 12 = 2012 единиц.

Из теоремы об оценках $\Delta f(\bar{X}) = y_i \cdot \Delta b_i$ известно, что колебание величины b_i приводит к увеличению или уменьшению $f(\bar{X})$. Оно определяется величиной y_i в случае, когда при изменении величин b_i значения переменных y_i, в оптимальном плане соответствующей двойственной задачи, остаются неизменными. Поэтому необходимо найти такие интервалы изменения каждого из свободных членов системы ограничений исходной ЗЛП, в которых оптимальный план двойственной задачи не менялся бы.

Для двойственных оценок оптимального плана весьма существенное значение имеет их предельный характер.

Точной мерой влияния ограничений на функционал оценки являются лишь при малом приращении ограничения.

Известно, что оценки не меняют своей величины, если не меняется набор векторов, входящих в базис оптимального плана, тогда как интенсивность этих векторов (значения неизвестных) в плане могут меняться.

Рассмотрим модель исходной задачи (3.1)—(3.3) в матричной форме:

$$f(\bar{X}) = C \cdot X \to \max,$$
$$A \cdot X \le B,$$
$$X \ge 0,$$

где $X = (x_1, x_2, ..., x_n)$ — вектор неизвестных;
$C = (c_1, c_2, ..., c_n)$ — вектор коэффициентов при неизвестных в целевой функции;
$B = (b_1, b_2, ..., b_m)$ — вектор свободных членов ограничений исходной задачи;
$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & .... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn} \end{pmatrix}$ — матрица коэффициентов в системе ограничений.

Приведем задачу к канонической форме, введем m дополнительных переменных. Задача примет вид:

$$f(\bar{X}) = C \cdot X \to \max,$$
$$A \cdot X = B,$$
$$X \ge 0,$$

где вектор неизвестных переменных X будет теперь иметь размерность $n + m$. Размерность матрицы A также изменится и будет равна $m \cdot (n + m)$.

Пусть известен оптимальный план. Разобьем вектор X на два подвектора: $X^* > 0$ и $X^0 = 0$. В первый включены неизвестные, вошедшие в базис оптимального решения (т.е. ненулевые в оптимальном плане). Соответственно матрицу A разобьем на две подматрицы: A^* (размерность $m \times m$) и A^0 (размерность $m \times n$). Первую из них сформируют т.е. столбцы матрицы A, которые соответствуют ненулевым неизвестным в оптимальном плане. Тогда $A^*X^* + A^0X^0 = B$. Так как $A^0X^0 = 0$, то $A^*X^* = B$. Умножив обе части последнего равенства на матрицу, обратную матрице A, получим $A^{*-1}A^*X^* = A^{*-1}B$. Так как $A^{*-1}A^* = E$, где E — единичная матрица, то $X^* = A^{*-1}B$. Обозначим A^{*-1} через D, тогда $X^* = DB$.

Матрица D характеризует влияние ресурсов на величину выпуска продукции X. Изменим размер выделяемых ресурсов, т.е. дадим приращение ΔB вектору B. Тогда

$X + \Delta X = D(B + \Delta B) = DB + D\Delta B.$

С учетом $X = DB$ можно записать:

$\Delta X = D\Delta B.$

Это соотношение определяет величину структурных сдвигов в выпуске продукции при изменении ограничений исходной задачи. Из соотношений второй теоремы двойственности видно, что двойственные оценки (переменные двойственные задачи) тесным образом связаны с оптимальным планом простой задачи. Всякое изменение исходных данных прямой задачи может оказать влияние как на ее оптимальный план ($\Delta X = D\Delta B$), так и на систему оптимальных двойственных оценок. Поэтому, чтобы проводить экономический анализ с использованием двойственных оценок, нужно знать их интервал устойчивости.

Второе свойство двойственных оценок означает, что изменение значений величины b_i приводит к увеличению или уменьшению $f(\bar{X})$. Это изменение, как выше уже отмечено, определяется величиной y_i и может быть определено лишь тогда, когда при изменении величин b_i значения переменных y_i в оптимальном плане соответствующей двойственной задачи остаются неизменными. Поэтому необходимо определить такие интервалы изменения каждого из свободных членов системы линейных уравнений $AX = B$, в которых оптимальный план двойственной задачи не меняется. Это имеет место тогда, когда среди компонент вектора $X = DB$ нет отрицательных.

Исходя из этого, получаем следующие оценки нижних и верхних пределов устойчивости двойственных оценок при изменении каждого ограничения в отдельности. Пределы уменьшения (нижняя граница) определяются по тем x_k ($k = 1, ..., m$), для которых соответствующие $d_{ki} > 0$:

$$\Delta b_i^{(-)} = \min\{x_k/d_{ki}\} \text{ для } d_{ki} > 0. \quad (3.13)$$

Пределы увеличения (верхняя граница) определяются по тем x_k, для которых $d_{ki} < 0$:

$$\Delta b_i^{(+)} = |\max\{x_k/d_{ki}\}| \text{ для } d_{ki} < 0. \quad (3.14)$$

Ослабление какого-либо i-го ограничения приводит к тому, что с определенного момента оказывается возможным изменить структуру (набор векторов) в базисе плана, что ведет к скачкообразному уменьшению величины оценки. Так продолжается до тех пор, пока i-й ресурс вообще перестанет быть дефицитным и его оценка обратится в нуль.

Определим интервалы устойчивости двойственных оценок в примере 3.3. Матрица A имеет вид

$$A=\begin{pmatrix}2 & 4\\ 4 & 1\\ 2 & 1\end{pmatrix}.$$

После приведения задачи к канонической форме матрица A примет следующий вид:

$$A=\begin{pmatrix}2 & 4 & 1 & 0 & 0\\ 4 & 1 & 0 & 1 & 0\\ 2 & 1 & 0 & 0 & 1\end{pmatrix}.$$

С ненулевыми значениями в оптимальный план вошли $x^*{}_1 = 200$, $x^*{}_2 = 400$ и $x^*{}_4 = 200$, следовательно, матрица A^* будет составлена из первого, второго и четвертого столбцов матрицы A:

$$A^*=\begin{pmatrix}2 & 4 & 0\\ 4 & 1 & 1\\ 2 & 1 & 0\end{pmatrix}.$$

Для вычисления интервалов устойчивости необходимо найти матрицу $D = A^{*-1}$ (правила вычисления обратной матрицы приведены в главе 2.

$$D=\begin{pmatrix}-1/6 & 0 & 2/3\\ 1/3 & 0 & -1/3\\ 1/3 & 1 & -7/3\end{pmatrix}=\begin{pmatrix}-0{,}166 & 0 & 0{,}333\\ 0{,}333 & 0 & -0{,}333\\ 0{,}333 & 1 & -2{,}333\end{pmatrix}.$$

При вычислении интервалов устойчивости по формулам (3.13) и (3.14) примем $x^*{}_1 = 200 = x_{k\text{-}1}$, $x^*{}_2 = 400 = x_k{=}_2$, $x^*{}_4 = 200 = x_k{=}_3$.

Интервалы устойчивости первого ресурса — «труд»:

$\Delta b_1{}^{(-)} = \min\{x_2/d_{21};\ x_3/d_{31}\} = \min\{400/0,\ 3333;\ 200/0,\ 3333\} = \min\{1200;\ 600\} = 600$;

$\Delta b_1{}^{(+)} = |x_1/d_{11}| = |200/-0{,}16667| = 1200$;

$b_1 = \{b_1 - \Delta b_1{}^{(-)};\ b_1 + \Delta b_1{}^{(+)}\} = \{2000 - 600;\ 2000 + 1200\} = \{1400;\ 3200\}$.

При изменении запасов ресурса «труд» в пределах от 1400 до 3200 единиц двойственная оценка его не изменится.

Интервалы устойчивости второго ресурса — «сырье»:

этот ресурс в оптимальном плане используется не полностью и поэтому не имеет верхней границы интервалов устойчивости;

нижняя граница определяется следующим образом:

$\Delta b_2^{(-)} = x_3/d_{31} = 200/1 = 200$;

$b_2 = \{b_2 - \Delta b_2^{(-)};\ b_2\} = \{1400 - 200;\ 1400\} = \{1200;\ 1400\}$.

Интервалы устойчивости третьего ресурса — «оборудование»:

$\Delta b_3^{(-)} = \{x_1/d_{13}\} = 200/0{,}6667 = 300$;

$\Delta b_3^{(+)} = |\max\{x_2/d_{23};\ x_3/d_{33}\}| = |\max\{400/-0{,}3333;\ 200/-2{,}3333\}| = |\max\{-1200;\ -85{,}7144\}| = |-85{,}7144| = 85{,}7144$;

$b_3 = \{b_3 - \Delta b_3^{(-)};\ b_3 + \Delta b_3^{(+)}\} = \{800 - 300;\ 800 + 85{,}7144\} = \{500;\ 885{,}7144\}$.

В нашем примере определим величину изменения объема прибыли от реализации продукции при увеличении ресурса «труд» на 12 единиц. Эти изменения находятся в интервалах устойчивости двойственных оценок, поэтому можно воспользоваться теоремой об оценках:

$\Delta f(\bar{X}) = 12 \cdot 40/3 = 160$.

Объем прибыли увеличится на 160 единиц.

Такой же ответ мы получили бы, если бы решили симплексным методом задачу с новыми ограничениями по ресурсу «труд». Новый оптимальный план:

$x_1^* = 198$,

$x_2^* = 404$,

$f(\bar{X}) = 32\ 160$,

$\Delta f(\bar{X}) = 32\ 160 - 32\ 000 = 160$.

Структурных сдвигов в программе не произошло, но значения переменных в плане изменились: продукции вида А может быть выпущено на 2 единицы меньше, а продукции вида Б — на 4 больше. Значение целевой функции при новых ограничениях увеличится на 160 единиц.

3. Чувствительность решения к изменению коэффициентов целевой функции. Так как любые изменения коэффициентов целевой функции оказывают влияние на оптимальность полученного ранее решения, то наша цель — найти такие диапазоны изменения коэффициентов в целевой функции (рассматривая каждый из коэффициентов отдельно), при которых оптимальные значения переменных остаются неизменными. Допустимые диапазоны изменения коэффициентов в целевой функции определятся из соотношений:

$\Delta c_i^{(-)} = \min\{y_k/d_{ik}\}$ для $d_{ik} > 0$;

$\Delta c_i^{(+)} = |\max\{y_k/d_{ik}\}|$ для $d_{ik} < 0$.

Используя эти соотношения в рассматриваемой задаче, получим для первого коэффициента целевой функции:

$\Delta c_1^{(-)} = y_3/d_{31} = 20/3 : 2/3 = 10$,

$\Delta c_1^{(+)} = |y_1 d_{11}| = |40/3 : (-1/6)| = 80$,

$c_1 = \{c_1 - c_1^{(-)};\ c_1 + c_1^{(+)}\} = \{40 - 1\ 0;\ 40 + 80\} = \{30;\ 120\}$;

для второго коэффициента:

$\Delta c_2^{(-)} = y_1/d_{21} = 40/3 : 1/3 = 40$,

$\Delta c_2^{(+)} = |y_3/d_{23}| = |20/3 : (-1/3)| = 20$,

$c_2 = \{c_2 - \Delta c_2^{(-)}; c_2 + c_2^{(+)}\} = \{60 - 40; 60 + 20\} = \{20; 80\}$.

Таким образом, найденный оптимальный план выпуска продукции не будет меняться при изменении прибыли на единицу продукции А в диапазоне от 30 до 120 и прибыли на единицу второй продукции Б в диапазоне от 20 до 80.

4. Целесообразность включения в план новых изделий. Пусть в рассматриваемой нами задаче предприятию были предложены на выбор три новых изделия, за счет которых можно было бы расширить номенклатуру выпускаемой продукции при тех же запасах ресурсов. Нормы затрат ресурсов и прибыль от реализации единицы продукции для этих изделий представлены в табл. 3.4. Определим из предложенных видов изделий выгодные для предприятия с экономической точки зрения.

Таблица 3.4

Ресурсы	Объективно обусловленые оценки ресурсов	Затраты ресурсов на одно изделие		
		В	Г	Д
Труд	40/3	6	4	2
Сырье	0	2	1	3
Оборудование	20/3	3	1	2
Прибыль на одно изделие		80	70	45

Эту задачу можно решить на основании свойства 3 двойственных оценок: в оптимальный план задачи на получение максимума прибыли может быть включен лишь тот вариант, для которого прибыль, недополученная из-за отвлечения дефицитных ресурсов, т.е. величина $\sum_{i=1}^{m} a_{ij} y_i$, покрывается полученной прибылью c_j.

Таким образом, характеристикой того или иного варианта служит разность $\Delta_j = \sum_{i=1}^{m} a_{ij} y_i - c_j$, при этом если $\begin{Bmatrix} A_j < 0, \\ \Delta_j \le 0 \end{Bmatrix}$, то вариант выгоден; если $\Delta_j > 0$, то невыгоден.

Для решения задачи воспользуемся соотношением и рассчитаем характеристики новых изделий.

Для изделия В:

$\Delta В = 6 \cdot 40/3 + 0 \cdot 2 + 20/3 \cdot 3 - 80 = 20$.

Поскольку $\Delta В = 100 - 80 = 20 > 0$, то делаем вывод, что изделие В невыгодно для включения в план, так как затраты на его изготовление не покрываются получаемой прибылью.

Аналогично для изделия Г:

$\Delta Г = 4 - 40/3 + 20/3 - 70 = 160/3 - 20/3 - 70 = 180/3 - 70 =$
$= 60 - 70 = -10 < 0$ — выгодно;

для изделия Д:

$\Delta Д = 2 \cdot 40/3 + 20/3 \cdot 2 - 45 = 80/3 + 40/3 - 45 = 40 - 45 = -5 < 0$ — выгодно.

В рассмотренных выше задачах детально изучены три первых свойства двойственных оценок и использование этих свойств при анализе оптимальных решений экономических задач: оценки как меры дефицитности ресурсов, оценки как меры влияния ограничений на функционал, оценки как инструмент определения эффективности отдельных вариантов.

Свойство 4 — оценки как инструмент балансирования суммарных затрат и результатов — вытекает из первой теоремы двойственности, в которой устанавливается связь между функционалами прямой и двойственной задач: $\max f(\bar{X}) = \min g(\bar{Y})$. В конкретных задачах такого рода соотношения «затраты — результаты», т.е. равновесие затрат и результатов в точке оптимума, могут иметь различное экономическое содержание.

В рассматриваемых нами задачах экономический смысл равенства функционалов прямой и двойственной задач состоит в том, что максимум прибыли может быть обеспечен лишь при минимуме недополученной прибыли от использования дефицитных ресурсов.

3.2. Транспортная задача

Как показано выше, многие прикладные модели в экономике сводятся к задачам линейного программирования. Практически все задачи линейного программирования можно решить, используя ту или иную модификацию симплексного метода. Однако существуют более эффективные вычислительные процедуры решения некоторых типов задач линейного программирования, основанные на специфике ограничений этих задач. Рассмотрим так называемую *транспортную задачу по критерию стоимости*, которую можно сформулировать следующим образом.

В m пунктах отправления $A_1, A_2, ..., A_m$, которые в дальнейшем будем называть поставщиками, сосредоточено определенное количество единиц некоторого однородного продукта, которое обозначим a_i $(i = 1, 2, ..., m)$. Данный продукт потребляется в n пунктах $B_1, B_2, ..., B_n$, которые будем называть потребителями; объем потребления обозначим

b_j (j = 1, 2, ..., n). Известны расходы на перевозку единицы продукта из пункта A_i в пункт B_j, которые равны c_{ij} и приведены в матрице транспортных расходов $C = (c_{ij})$.

Требуется составить такой план прикрепления потребителей к поставщикам, т.е. план перевозок, при котором весь продукт вывозится из пунктов A_i в пункты B_j в соответствии с потребностью и общая величина транспортных издержек будет минимальной.

Обозначим количество продукта, перевозимого из пункта A_i в пункт B_j, через x_{ij}. Совокупность всех переменных x_{ij} для краткости обозначим $\bar{X}$, тогда целевая функция задачи будет иметь вид

$$f(\bar{X}) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min, \tag{3.15}$$

а ограничения выглядят следующим образом:

$$\sum_{i=1}^{m} x_{ij} = b_j,\ j = \overline{1,n}. \tag{3.16}$$

$$\begin{aligned} &\sum_{j=1}^{n} x_{ij} = a_i,\ i = \overline{1,m}, \\ &x_{ij} \ge 0. \end{aligned} \tag{3.17}$$

Условия (3.16) означают полное удовлетворение спроса во всех пунктах потребления; условия (3.17) определяют полный вывоз продукции от всех поставщиков.

Необходимым и достаточным условием разрешимости задачи (3.15)—(3.17) является *условие баланса*:

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j. \tag{3.18}$$

Транспортная задача, в которой имеет место равенство (3.18), называется *закрытой* и может быть решена как задача линейного программирования с помощью симплексного метода. Однако благодаря особенностям переменных задачи и системы ограничений разработаны специальные, менее громоздкие методы ее решения. Наиболее применяемым методом является *метод потенциалов*, при котором каждой i-й строке (i-му поставщику) устанавливается потенциал u_i, который можно интерпретировать как цену продукта в пункте поставщика, а каждому столбцу j (j-му потребителю) устанавливается потенциал v_j, который можно принять условно за цену продукта в пункте потребителя.

В простейшем случае цена продукта в пункте потребителя равна его цене в пункте поставщика плюс транспортные расходы на его доставку, т.е.

$$v_j = u_i + c_{ij}. \tag{3.19}$$

Алгоритм метода потенциалов для закрытой транспортной задачи детально описан в ряде учебных пособий (см., например, [6]). *Первым этапом* этого алгоритма является составление начального распределения (начального плана перевозок); для реализации этого начального этапа имеется в свою очередь ряд методов: северо-западного угла, наименьших стоимостей, аппроксимаций Фогеля и др. *Вторым этапом* служат построение системы потенциалов на основе равенства (3.19) и проверка начального плана на оптимальность; в случае его неоптимальности переходят к *третьему этапу*, содержание которого заключается в реализации так называемых циклов перераспределения (корректировка плана прикрепления потребителей к поставщикам), после чего переходят опять ко второму этапу. Совокупность процедур третьего и второго этапов образует одну итерацию; эти итерации повторяются, пока план перевозок не окажется оптимальным по критерию (3.15).

Если баланс (3.18) не выполняется, то ограничения (3.16) или (3.17) имеют вид неравенств типа «меньше или равно»; транспортная задача в таком случае называется *открытой*. Для решения открытой транспортной задачи методом потенциалов ее сводят к закрытой задаче путем ввода или фиктивного потребителя, если в неравенства превращаются условия (3.17), или фиктивного поставщика — в случае превращения в неравенства ограничений (3.16) (см. ниже пример 3.5).

Рассмотрим этапы реализации метода потенциалов для закрытой транспортной задачи более подробно. Прежде всего следует отметить, что при условии баланса (3.18) ранг системы линейных уравнений (3.16), (3.17) равен $m + n - 1$; таким образом из общего числа $m \times n$ неизвестных базисных неизвестных будет $m + n - 1$. Вследствие этого при любом допустимом базисном распределении в матрице перевозок (таблице поставок), представленной в общем виде в табл. 3.5, будет занято ровно $m + n - 1$ клеток, которые будем называть *базисными* в отличие от остальных, *свободных*, клеток; занятые клетки будем отмечать диагональной чертой.

Таблица 3.5

Мощности поставщиков	Мощности потребителей			
	b_1	b_2	...	b_n
a_1	c_{11} / x_{11}	c_{12} / x_{12}	...	c_{1n} / x_{1n}
a_2	c_{21} / x_{21}	c_{22} / x_{22}	...	c_{2n} / x_{2n}
...	...	...	...	...
a_m	c_{m1} / x_{m1}	c_{m2} / x_{m2}	...	c_{mn} / x_{mn}

Этап 1. **Первоначальное закрепление потребителей за поставщиками.** Рассмотрим два метода получения начального распределения (начального опорного плана): метод северо-западного угла и метод наименьших стоимостей. При каждом из этих методов при заполнении некоторой клетки, кроме последней, вычеркивается или только строка матрицы перевозок, или только столбец; лишь при заполнении последней клетки вычеркиваются и строка, и столбец. Такой подход будет гарантировать, что базисных клеток будет ровно $m + n - 1$. Если при заполнении некоторой (не последней) клетки одновременно удовлетворяются мощности и поставщика, и потребителя, то вычеркивается, например, только строка, а в соответствующем столбце заполняется незанятая клетка так называемой нулевой поставкой, после чего вычеркивается и столбец. Для идентификации клетки обычно в скобках указываются номера ее строки и столбца.

В методе северо-западного угла всегда в первую очередь заполняется клетка (из числа невычеркнутых), стоящая в верхнем левом (северо-западном) углу матрицы перевозок. Пример составления начального распределения данным методом показан в табл. 3.6: заполняется клетка (1; 1) и вычеркивается первый столбец, заполняется клетка (1; 2) и вычеркивается первая строка; заполняется клетка (2; 2) и вычеркивается второй столбец; заполняется клетка (2; 3) и вычеркивается вторая строка; заполняется клетка (3; 3) и вычеркивается третий столбец; наконец, заполняется клетка (3; 4) и вычеркиваются последние строка и столбец. Число запятых клеток равно $m + n - 1 = 3 + 4 - 1 = 6$. Суммарные затраты па реализацию данного плана перевозок составят

$$f(\bar{X}) = 4\cdot 30 + 5\cdot 30 + 3\cdot 70 + 6\cdot 30 + 7\cdot 10 + 4\cdot 110 = 1170.$$

Таблица 3.6

Мощности поставщиков	Мощности потребителей			
	30	100	40	110
60	4 / 30	5 / 30	2	3
100	1	3 / 70	6 / 30	2
120	6	2	7 / 10	4 / 110

Недостатком данного метода является то, что он не учитывает значения элементов c_{ij} матрицы транспортных расходов, в результате чего полученное этим методом начальное распределение (начальный опорный план перевозок) может быть достаточно далеко от оптимального.

В различных модификациях метода наименьших стоимостей заполнение клеток матрицы перевозок проводится с учетом значений величин c_{ij}. Так, в модификации «двойного предпочтения» отмечают клетки с наименьшими стоимостями перевозок сначала по каждой строке, а затем по каждому столбцу. Клетки, имеющие две отметки, заполняют в первую очередь, затем заполняют клетки с одной отметкой, а данные о нераспределенном грузе записывают в неотмеченные клетки с наименьшими стоимостями. При этом из двух клеток с одинаковой стоимостью перевозок предпочтение отдается клетке, через которую осуществляется больший объем перевозок. Вычеркивание строк и столбцов при заполнении клеток проводится по описанным выше правилам. Пример начального распределения методом наименьших стоимостей дня тех же исходных данных, что и ранее, представлен в табл. 3.7.

Порядок заполнения клеток: (2; 1), (3; 2), (1; 3), (2; 4), (1; 4), (3; 4). Суммарные затраты на перевозки, представленные в табл. 3.7, составляют

$$f(\bar{X}) = 1\cdot 30 + 2\cdot 100 + 2\cdot 40 + 2\cdot 70 + 3\cdot 20 + 4\cdot 20 = 590.$$

Таблица 3.7

Мощности поставщиков	Мощности потребителей			
	30	100	40	110
60	4	5	2 / 40	3 / 20
100	1 / 30	3	6	2 / 70
120	6	2 / 100	7	4 / 20

Следовательно, данный план перевозок значительно ближе к оптимальному, чем план, составленный по методу северо-западного угла.

Этап 2. **Проверка оптимальности полученного плана перевозок**. Введем специальные показатели u_i для каждой строки матрицы перевозок (каждого поставщика), где $i = \overline{1,m}$, и показатели v_j для каждого столбца (каждого потребителя), где $j = \overline{1,n}$. Эти показатели называются потенциалами поставщиков и потребителей, их удобно интерпретировать как цены продукта в соответствующих пунктах поставщиков и потребителей. Потенциалы подбираются таким образом, чтобы для заполненной клетки (i; j) выполнялось равенство (3.19). Совокупность уравнений вида (3.19), составленных для всех заполненных клеток (всех базисных неизвестных), образует систему $m + n - 1$ линейных уравнений с $m + n$ неизвестными u_i и v_j. Эта система всегда совместна, причем значение одного из неизвестных можно задавать произвольно (например, $u_i = 0$), тогда значения остальных неизвестных находятся из системы однозначно.

Рассмотрим процесс нахождения потенциалов для базисного начального распределения по методу северо-западного угла, представленного в табл. 3.6. Задав $u_i = 0$ и используя формулу (3.19) для заполненных клеток (1; 1) и (1; 2), находим $v_1 = 4$ и $v_2 = 5$. Зная v_2, по заполненной клетке (2; 2) находим $u_3 = 2$, а зная v_2, по заполненной клетке (2; 3) находим $v_3 = 8$. Зная v_3, по заполненной клетке (3; 3) находим $u_3 = 1$, а затем по заполненной клетке (3; 4) находим $v_4 = 5$. Результаты представлены в табл. 3.8, где потенциалы поставщиков приведены в последнем столбце, а потенциалы потребителей — в последней строке.

Смысл прямоугольного контура, проведенного пунктиром в табл. 3.8, и знаков при его вершинах пояснен далее при описании этапа 3 метода потенциалов.

Таблица 3.8

Мощности поставщиков	Мощности потребителей				u_i
	30	100	40	110	
60	4 / 30	5 / 30 −	2 +	3	0
100	1	3 / 70 +	6 / 30 −	2	2
120	6	2	7 / 10	4 / 110	1
v_j	4	5	8	5	

Аналогичные результаты для начального распределения по методу наименьших стоимостей, приведенного в табл. 3.7, представлены в табл. 3.9.

Таблица 3.9

Мощности поставщиков	Мощности потребителей				u_i
	30	100	40	110	
60	4	5	2 / 40	3 / 20	0
100	1 / 30	3	6	2 / 70	1
120	6	2 / 100	7	4 / 20	−1
v_j	2	1	2	3	

Чтобы оценить оптимальность распределения, для всех клеток (i, j) матрицы перевозок определяются их *оценки,* которые обозначим через Δ_{ij}, по формуле

$$\Delta_{ij} = (u_i + c_{ij}) - v_j. \qquad (3.20)$$

Используя ранее принятую интерпретацию, выражение $(u_i + c_{ij})$ можно трактовать как сумму цены продукта у поставщика и стоимости перевозки; эта сумма путем вычитания сравнивается с ценой продукта у соответствующего потребителя v_j. Очевидно, оценки заполненных клеток равны нулю (цена потребителя покрывает цену поставщика и стоимость перевозок). Таким образом, об оптимальности распределения можно судить по величинам оценок свободных клеток. Если оценка некоторой свободной клетки отрицательна, это можно интерпретировать так: цена, предлагаемая соответствующим потребителем, больше суммы цены поставщика и стоимости перевозки, т.е. если бы эта клетка была занята, то можно было бы получить дополнительный экономический эффект. Следовательно, условием оптимальности распределения служит условие неотрицательности оценок свободных клеток матрицы перевозок.

Оценки клеток по формуле (3.20) удобно представить в виде *матрицы оценок*. Для ранее рассматриваемого распределения, полученного методом северо-западного угла (см. табл. 3.8), матрица оценок клеток имеет вид

$$(\Delta_{ij}) = \begin{pmatrix} 0 & 0 & -6 & -2 \\ -1 & 0 & 0 & -1 \\ 3 & -2 & 0 & 0 \end{pmatrix}. \qquad (3.21)$$

Наличие большего числа отрицательных оценок свободных клеток свидетельствует о том, что данный план перевозок далек от оптимального (напомним, что суммарные затраты на перевозку по этому плану равны 1170).

Для распределения, полученного методом наименьших стоимостей (см. табл. 3.9), матрица оценок клеток имеет вид

$$(\Delta_{ij}) = \begin{pmatrix} 2 & 4 & 0 & 0 \\ 0 & 3 & 5 & 0 \\ 3 & 0 & 4 & 0 \end{pmatrix}.$$

Так как все оценки неотрицательны, то не имеется возможности улучшить данный план перевозок, т.е. он оптимален (суммарные затраты на перевозку по этому плану равны 590). Кроме того, следует отметить, что в данном случае оценки всех свободных клеток строго больше нуля, т.е. любой другой план, предусматривающий занятие хотя бы одной из этих клеток, будет менее оптимален. Это говорит о том, что найденный оптимальный план является *единственным*. Наличие нулевых оценок свободных клеток в оптимальном плане перевозок, наоборот, свидетельствует о *неединственности* оптимального плана.

Этап 3. **Улучшение неоптимального плана перевозок (циклы перераспределения)**. Чтобы улучшить неоптимальный план перевозок, выбирается клетка матрицы перевозок с отрицательной оценкой; если таких клеток несколько, то обычно (но необязательно) выбирается клетка с наибольшей по абсолютной величине отрицательной оценкой. Например, для распределения, представленного в табл. 3.8, такой клеткой может служить клетка (1; 3) (см. матрицу оценок (3.21)).

Для выбранной клетки строится замкнутая линия (*контур*), начальная вершина которой лежит в выбранной клетке, а все остальные вершины находятся в занятых клетках; при этом направления отдельных отрезков контура могут быть только горизонтальными и вертикальными. Вершиной контура, кроме первой, является занятая клетка, где отрезки контура образуют один прямой угол (нельзя рассматривать как вершины клетки, где горизонтальные и вертикальные отрезки контура пересекаются). Очевидно, число отрезков контура, как и его вершин, будет четным. В вершинах контура расставляются поочередно знаки «+» и «–», начиная со знака «+» в выбранной свободной клетке. Пример простого контура показан пунктиром в табл. 3.8, хотя вид контура может быть самым разнообразным (см., например, контур в табл. 3.11).

Величина перераспределяемой поставки определяется как наименьшая из величин поставок в вершинах контура со знаком «–», и на эту величину увеличиваются поставки в вершинах со знаком «+» и уменьшаются поставки в вершинах со знаком «–». Это правило гарантирует, что в вершинах контура не появится отрицательных поставок, начальная выбранная клетка окажется занятой, в то время как одна из занятых клеток при этом обязательно освободится. Если величина перераспределяемой поставки равна поставкам не в одной, а в нескольких вершинах контура со знаком «–» (это как раз имеет место в контуре перераспределения в табл. 3.8), то освобождается только одна клетка, обычно с наибольшей стоимостью перевозки, а все другие такие клетки остаются занятыми с нулевой поставкой.

Результат указанных операций для представленного в табл. 3.8 распределения поставок показан в табл. 3.10. Суммарные затраты на перевозки по этому плану составляют:

$$f(\bar{X}) = 4\cdot 30 + 5\cdot 0 + 2\cdot 30 + 3\cdot 100 + 7\cdot 10 + 4\cdot 110 = 990,$$

что значительно меньше предыдущей суммы затрат 1170, хотя план перевозок в табл. 3.10 еще не является оптимальным. Об этом свидетельствует наличие отрицательных значений в матрице оценок клеток этого плана (соответствующие потенциалы u_i и v_j найдены способом, изложенным при описании этапа 2):

$$(\Delta_{ij}) = \begin{pmatrix} 0 & 0 & 0 & 4 \\ -1 & 0 & 6 & 5 \\ -3 & -8 & 0 & 0 \end{pmatrix}.$$

Таблица 3.10

Мощности поставщиков	Мощности потребителей				u_i
	30	100	40	110	
60	4 / 30	5 / 0	2 / 30	3	0
100	1	3 / 100	6	2	2
120	6	2	7 / 10	4 / 110	–5
v_j	4	5	2	–1	

Транспортные задачи, в базисном плане перевозок которых имеют место занятые клетки с нулевой поставкой (или в первоначальном распределении, или в процессе итераций), называются *вырожденными*; пример такой задачи представлен в табл. 3.10. В случае вырожденной транспортной задачи существует опасность *зацикливания*, т.е. бесконечного повторения итераций (бесконечного перебора одних и тех же базисных комбинаций занятых клеток). Как правило, в практических задачах транспортного типа зацикливание не встречается, тем не менее следует знать, что существуют специальные правила, позволяющие выйти из цикла, если зацикливание все же произойдет. При отсутствии вырождения метод потенциалов конечен и приводит к оптимальному плану перевозок за конечное число шагов.

Пример 3.4. Решим методом потенциалов закрытую транспортную задачу, заданную в табл. 3.11, в которую уже внесено некоторое допустимое базисное распределение. Суммарные транспортные расходы составляют при этом плане перевозок $f(\overline{X}) = 3\cdot 5 + 2\cdot 25 + 1\cdot 20 + 2\cdot 25 + 1\cdot 15 + 4\cdot 20 = 230$. Потенциалы по формуле (3.19) находим следующим образом: задавая $u_1 = 0$, находим по клетке (1; 1) $v_i = 3$, по клетке (1; 2) $v_4 = 2$, а по клетке (1; 4) $v_4 = 1$; затем по клетке (2; 1) находим $u_2 = 1$ и по клетке (2; 3) $v_3 = 2$; наконец, по клетке (3; 3) находим $u_3 = -2$.

Таблица 3.11

Мощности поставщиков	Мощности потребителей				u_1
	30	25	35	20	
50	3 + / 5	2 − / 25	4	1 / 20	0
40	2 / − 25	3	1 + / 15	5	1
20	3	2 / +	4 / − 20	4	−2
v_j	3	2	2	1	

Матрица оценок клеток для этого плана рассчитывается по формуле (3.21):

$$(\Delta_{ij}) = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 5 \\ -2 & -2 & 0 & 1 \end{pmatrix}.$$

Наличие отрицательных оценок свидетельствует о том, что план неоптимален. Построим контур перераспределения, например, для клетки (3; 2) в табл. 3.11 он показан пунктиром и его вершинам присвоены соответствующие знаки.

Наименьшая поставка в вершине контура со знаком «−» равна 20, поэтому проведем перераспределение поставок, уменьшив поставки в клетках со знаком «−» на 20 и увеличив поставки в клетках со знаком «+» также на 20; при этом клетка (3; 2) заполняется, а клетка (3; 3) освобождается. Новый план представлен в табл. 3.12, соответствующие значения потенциалов показаны в последних столбце и строке.

Таблица 3.12

Мощности поставщиков	Мощности потребителей				u_i
	30	25	35	20	
50	3 / 25	2 / 5	4	1 / 20	0
40	2 / 5	3	1 / 35	5	1
20	3	2 / 20	4	4	0
v_j	3	2	2	1	

Матрица оценок клеток этого распределения не содержит отрицательных значений, следовательно, данный план перевозок является оптимальным, стоимость перевозок по этому плану равна

$$(\Delta_{ij}) = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 5 \\ 0 & 0 & 2 & 3 \end{pmatrix}.$$

$$f(\bar{X}) = 3\cdot 25 + 2\cdot 5 + 1\cdot 20 + 2\cdot 5 + 1\cdot 35 + 2\cdot 20 = 190.$$

Наличие нулевой оценки незанятой клетки (3;1) говорит о том, что оптимальный план не является единственным. Можно отметить также, что, применяя для начального распределения в этой транспортной задаче модификацию двойного предпочтения метода наименьших стоимостей, мы сразу же получили бы оптимальное распределение, представленное в табл. 3.12.

Пример 3.5. Рассмотрим пример решения открытой транспортной задачи, исходные данные которой представлены в табл. 3.13.

Таблица 3.13

Мощности поставщиков	Мощности потребителей			
	30	100	40	100
60	4	5	2	3
100	1	3	6	2
120	6	2	7	4

Проверим выполнение условия баланса (3.18):

$$\sum_{i=1}^{m} a_i = 60 + 100 + 120 = 280;$$

$$\sum_{j=1}^{n} b_j = 30 + 100 + 40 + 100 = 270.$$

Условие баланса не выполняется, следовательно, задача действительно является открытой. Сформулируем экономико-математическую модель этой задачи.

Пусть x_{ij} — объемы перевозок от i-го поставщика j-му потребителю. Целевая функция задачи будет иметь вид

$$\max f(\overline{X}) = 4 \cdot x_{11} + 5 \cdot x_{12} + 2 \cdot x_{13} + 3 \cdot x_{14} +$$
$$+ x_{21} + 3 \cdot x_{22} + 6 \cdot x_{23} + 2 \cdot x_{24} +$$
$$+ 6 \cdot x_{31} + 2 \cdot x_{32} + 7 \cdot x_{33} + 4 \cdot x_{34},$$

а ограничения выглядят следующим образом:

по поставщикам:

$$x_{11} + x_{12} + x_{13} + x_{14} \le 60$$
$$x_{21} + x_{22} + x_{23} + x_{24} \le 100$$
$$x_{31} + x_{32} + x_{33} + x_{34} \le 120;$$

по потребителям:

$$x_{11} + x_{21} + x_{31} = 30$$
$$x_{12} + x_{22} + x_{32} = 100$$
$$x_{13} + x_{23} + x_{33} = 40$$
$$x_{14} + x_{24} + x_{23} = 100;$$

прямые ограничения: $x_{ij} \ge 0;\ i = \overline{1,3};\ j = \overline{1,4}$.

Сведем эту задачу к закрытой, введя фиктивного потребителя с мощностью 280 – 270 = 10; стоимости перевозок единицы груза в соответствующих (фиктивных) клетках принимаются равными нулю, и эти клетки при составлении начального распределения перевозок заполняются в самую последнюю очередь. Начальный план перевозок по методу наименьших стоимостей представлен в табл. 3.14.

Матрица оценок клеток этого плана в соответствии с методом потенциалов имеет вид

$$(\Delta_{ij}) = \begin{pmatrix} 2 & 4 & 0 & 0 & 1 \\ 0 & 3 & 5 & 0 & 2 \\ 3 & 0 & 4 & 0 & 0 \end{pmatrix}.$$

Таблица 3.14

Мощности поставщиков	Мощности потребителей					u_i
	30	100	40	100	10	
60	4	5	2 / 40	3 / 20	0	0
100	1 / 30	3	6	2 / 70	0	1
120	6	2 / 100	7	4 / 10	0 / 10	−1
v_j	2	1	2	3	−1	

Анализ элементов матрицы оценок клеток свидетельствует о том, что план перевозок в табл. 3.14 является оптимальным и единственным. Стоимость перевозок по этому плану составит 550, при этом мощности потребителей будут удовлетворены полностью, а у третьего поставщика останутся невывезенными 10 единиц груза.

3.3. Целочисленное программирование

Целочисленным (иногда его называют также *дискретным*) программированием называется раздел математического программирования, изучающий экстремальные задачи, в которых на искомые переменные накладывается условие целочисленности, а область допустимых решений конечна. Изучение этого раздела в курсе «Экономико-математические методы и прикладные модели» вызывается тем, что огромное количество экономических задач носит дискретный, чаще всего целочисленный характер, что связано, как правило, с физической неделимостью многих элементов расчета: например, нельзя построить два с половиной завода, купить полтора автомобиля и т.д. В ряде случаев такие задачи решаются обычными методами, например симплексным методом, с последующим округлением до целых чисел. Однако такой подход оправдан, когда отдельная единица составляет очень малую часть всего объема (например, товарных запасов); в противном случае он может внести значительные искажения в действительно оптимальное решение. Поэтому разработаны специальные методы решения целочисленных задач, среди которых можно выделить два направления: методы отсечения (отсекающих плоскостей) и комбинаторные методы.

Метод отсекающих плоскостей состоит в построении дополнительных ограничений и применении двойственного симплексного метода. Представление о *комбинаторных методах* дает широко используемый на практике *метод ветвей и границ*.

Рассмотрим более подробно один из методов отсекающих плоскостей, известный как *метод Гомори*. Метод Гомори для линейных задач целочисленного программирования основан на понятии *конгруэнтности* действительных чисел. Любое действительное число можно представить в виде суммы его целой и дробной частей: $x = [x] + \{x\}$, где квадратные скобки означают целую часть, а фигурные — дробную. Например, 7,5 = [7,5] + {7,5} = 7 + 0,5. Неотрицательные числа (для простоты мы будем рассматривать только их) называются конгруэнтными, если равны их дробные части. Если обозначать конгруэнтность чисел символом «≡», то, например, 7,5 ≡ 0,5; 6,3 ≡ 2,3; в частности, все целые числа конгруэнтны нулю, поэтому условие целочисленности переменной x можно записать: $x_i \equiv 0$.

По методу Гомори первый этап решения целочисленных задач не отличается от обычного расчета по симплексному алгоритму. Если среди значений переменных в оптимальном плане есть дробные, то составляется дополнительное ограничение, отсекающее дробную часть решения, но оставляющее в силе все прочие условия, которым должен удовлетворять оптимальный план. Это дополнительное ограничение присоединяется к исходным ограничениям задачи, и вновь применяется процедура симплексного метода. Алгоритм Гомори позволяет прийти к оптимальному целочисленному решению за конечное число шагов.

Пример 3.6. Пусть для приобретения оборудования, размещаемого на производственной площади 38 м^2, фирма выделяет 20 млн руб. Имеются единицы оборудования двух типов: типа А стоимостью 5 млн руб., требующее производственную площадь 8 м^2 и имеющее производительность 7 тыс. ед. продукции за смену, и типа Б — стоимостью 2 млн руб., занимающее площадь 4 м^2 и дающее за смену 3 тыс. ед. продукции. Требуется рассчитать оптимальный вариант приобретения оборудования, обеспечивающий максимум производительности участка.

Решение. Сформулируем экономико-математическую модель задачи. Пусть x_1, x_2 — количество приобретаемых машин типа

А и типа Б соответственно. Тогда целевая функция задачи будет иметь вид

$$\max f(\bar{X}) = 7x_1 + 3x_2$$

при ограничениях:

$$\begin{gathered} 5x_1 + 2x_2 \le 20 \\ 8x_1 + 4x_2 \le 38 \\ x_{1,2} \ge 0;\ x_{1,2} \equiv 0 \end{gathered}$$

Сформулирована задача линейного целочисленного программирования.

Введем дополнительные переменные x_3, x_4, с помощью которых исходные неравенства преобразуются в равенства:

$$\begin{gathered} 5x_1 + 2x_2 + x_3 = 20; \\ 8x_1 + 4x_2 + x_4 = 38, \end{gathered}$$

из которых следует, что переменные x_3, x_4 могут принимать только неотрицательные целочисленные значения. Фрагмент симплексной таблицы на последней итерации (без учета целочисленности) приведен в табл. 3.15.

Таблица 3.15

Базисные переменные	План	x_1	x_2	x_3	x_4
x_1	1	1	0	1	–0,5
x_2	7,5	0	1	–2	1,26
$\Delta_j = Z_j - c_j$	29,5	0	0	1	0,25

Отсюда видно, что в оптимальном плане $x_1 = 1$; $x_2 = 7{,}5$ и максимум целевой функции равен

$$f(\bar{X}) = 7 \cdot 1 + 3 \cdot 7{,}5 = 29{,}5.$$

Для нецелочисленной переменной x_2 составляем из приведенной симплексной табл. 3.15 уравнение:

$$7{,}5 = x_2 - 2x_3 + 1{,}25$$

откуда

$$x_2 = 7{,}5 + 2x_3 - 1{,}25x_4.$$

Так как x_2 — целое число, то целой должна быть и правая часть последнего уравнения («≡» — символ конгруэнтности):

$$7{,}5 + 2x_3 - 1{,}25x_4 \equiv 0;$$

отсюда можно получить, что

$$0{,}25x_4 \equiv 0{,}5,$$

т.е. выражение $0{,}25x_4$ может быть равно 0,5, или 1,5, или 2,5 и т.д. Следовательно, появляется дополнительное ограничение: $0{,}25x_4 \geq 0{,}5$, которое путем ввода дополнительной неотрицательной целочисленной переменной x_5 преобразуется в равенство, так что система ограничений исходной задачи в каноническом виде содержит три уравнения:

$$5x_1 + 2x_2 + x_3 = 20;$$
$$8x_1 + 4x_2 + x_4 = 38;$$
$$0{,}25x_4 - x_5 = 0{,}5.$$

Повторив процесс решения симплексным методом для данной расширенной системы ограничений, получим новый оптимальный план, в котором переменные, входящие в базис, принимают целые значения: $x_1 = 2$; $x_2 = 5$; $x_4 = 2$. Таким образом, приобретение двух машин типа А и пяти машин типа Б обеспечивает максимум производительности участка, равный 29 тыс. ед. продукции в смену. Заметим, что если бы в качестве плана был выбран вариант, получаемый в результате округления первоначального решения задачи симплексным методом ($x_1 = 1$; $x_2 = 7$), то суммарная производительность была бы равна лишь 28 тыс. ед. продукции.

Рассмотрим далее ряд специальных оптимизационных задач, сводящихся к задачам линейного целочисленного программирования. Одной из таких задач является *задача о назначениях*, с помощью которой можно получить ответ на вопросы типа: как распределить рабочих по станкам, чтобы общая выработка была наибольшей или затраты на заработную плату наименьшими; как наилучшим образом распределить экипажи самолетов; как назначить людей на различные должности (отсюда и название задачи) и т.д.

Математически такие задачи относятся к тому же типу распределительных задач, что и рассмотренная в параграфе 3.2 транспортная задача, с той особенностью, что в них объемы наличных и требующихся ресурсов для выполнения каждой работы равны единице ($a_i = b_j = 1$), а все переменные x_{ij} либо равны единице, если i-й работник назначен на j-ю работу, либо равны нулю в других случаях. Исходные данные задачи о назначениях группируются в таблице, которая называется *матрицей оценок*, а результаты — в *матрице назначений*.

Задача о назначениях в общем виде может быть сформулирована следующим образом. Имеется n работников, которые могут выполнять n работ, причем использование i-го работника на j-й работе, например, приносит доход c_{ij}.

Требуется поручить каждому из работников выполнение одной вполне определенной работы, чтобы максимизировать в данном случае суммарный доход.

Введем переменные:

$$x_{ij} = \begin{cases} 1, \text{если } i\text{-й работник выполняет } j\text{-ю работу;} \\ 0 - \text{в противном случае.} \end{cases}$$

Задача состоит в том, чтобы найти распределение $\overline{X} = (x_{ij})$ работников по работам (т.е. найти матрицу назначений), которое максимизирует целевую функцию

$$f(\overline{X}) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \to \max \tag{3.22}$$

при ограничениях

$$\sum_{j=1}^{n} x_{ij} = 1,\ i = \overline{1, n}; \tag{3.23}$$

$$\sum_{i=1}^{n} x_{ij} = 1,\ j = \overline{1, n}, \tag{3.24}$$

причем x_{ij} равно либо 0, либо 1 (так называемые *булевы переменные*) для всех $i, j = \overline{1, n}$.

Ограничения (3.23) отражают условие того, что за каждым работником может быть закреплена только одна работа, а ограничения (3.24) означают, что для выполнения каждой работы может быть выделен только один работник.

Если в задаче о назначениях элементы матрицы оценок представляют собой, например, время выполнения каждым работником любой из работ, то целевая функция этой задачи будет минимизироваться. Следует заметить также, что при решении задачи о назначениях часто используются алгоритмы и методы решения транспортных задач, в частности метод потенциалов.

Другой задачей подобного рода является *задача о коммивояжере*, которая может быть сформулирована следующим образом. Имеется n городов, пронумерованных числами от 1 до n. Коммивояжер, выезжая из города 1, должен побывать в каждом городе ровно один раз и вернуться в исходный пункт. Пусть известны расстояния c_{ij} между городами ($i, j = \overline{1, n}; i \neq j$). Требуется найти самый короткий маршрут.

Введем переменные:

$$x_{ij} = \begin{cases} 1, \text{если в маршрут входит переезд} \\ \text{из города } i \text{ в город } j; \\ 0 - \text{в противном случае } (i, j = \overline{1,n};\ i \neq j). \end{cases}$$

Требование однократного въезда в города и выезда из них запишется в виде следующих ограничений:

$$\begin{aligned} &\sum_{i=1}^{n} x_{ij} = 1,\ j = \overline{1,n}, \\ &\sum_{j=1}^{n} x_{ij} = 1,\ i = \overline{1,n}. \end{aligned} \qquad (3.25)$$

Однако ограничения (3.25) полностью не определяют допустимые маршруты, так как не исключают возможности разрыва пути, т.е. появления нескольких не связанных между собой подмаршрутов для части городов. Поэтому следует ввести дополнительно n переменных u_i ($u_i (i = \overline{1,n})$, принимающих только целые неотрицательные значения, и записать для них специальные ограничения:

$$u_i - u_j + n \cdot x_{ij} \le n - 1;\ i, j = \overline{2,n};\ i \neq j. \qquad (3.26)$$

Общее число таких ограничений равно $(n - 1) \cdot (n - 2)$, и они, не исключая допустимый маршрут, исключают возможность существования подмаршрутов.

Таким образом, задача о коммивояжере состоит в минимизации целевой функции:

$$f(\bar{X}, \bar{U}) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} \to \min \qquad (3.27)$$

при условиях (3.25), (3.26), где переменные x_{ij}, u_i принимают только неотрицательные целые значения.

К задачам целочисленного программирования приводят также многие оптимальные задачи *теории расписаний*, в которой рассматриваются методы оптимизации оперативно-календарного планирования. В качестве примера таких задач можно привести задачу определения оптимальной очередности обработки изделий на различных станках или других рабочих местах, задачу составления программы «диспетчер» для управления работой ЭВМ в мультипрограммном режиме и др.

3.4. Задачи многокритериальной оптимизации

В практической деятельности часто встречаются задачи, заключающиеся в поиске лучшего (оптимального) решения при наличии различных несводимых друг к другу критериев оптимальности. Например, принятие решения о строительстве дороги в объезд города должно учитывать такие факторы, как выигрыш города в целом по соображениям экологии, проигрыш отдельных предприятий и фирм, например, из-за уменьшения проезжающих через город потенциальных покупателей и многие другие. Если такого рода задачи решаются методами математического программирования, то говорят о задачах *многокритериальной оптимизации*. Эти задачи могут носить как линейный, так и нелинейный характер. Поскольку методы решения таких задач излагаются ниже на примере линейных многокритериальных оптимизационных задач, это объясняет рассмотрение этой темы в данной главе учебного пособия.

Задачи многокритериальной оптимизации возникают в тех случаях, когда имеется несколько целей, которые не могут быть отражены одним критерием (например, стоимость и надежность). Требуется найти точку области допустимых решений, которая минимизирует или максимизирует все такие критерии. Если в подобного рода задачах речь идет не о разнородных критериях некоторой системы, а о сопоставлении однородных критериев разных ее подсистем (например, отрасли, группы населения и т.п.), то эти *задачи называются задачами векторной оптимизации*.

Обозначим i-й частный критерий через $Z_i(\bar{X})$, где $\bar{X}$ — допустимое решение, а область допустимых решений — через Q. Если учесть, что изменением знака функции всегда можно свести задачу минимизации к задаче максимизации, то кратко задачу многокритериальной оптимизации можно сформулировать следующим образом:

$$Z(\bar{X}) = \left\langle Z_1(\bar{X}), Z_2(\bar{X}), \ldots, Z_m(\bar{X}) \right\rangle \to \max \qquad (3.28)$$

$$\bar{X} \in Q. \qquad (3.29)$$

Некоторые частные критерии могут противоречить друг другу, другие действуют в одном направлении, третьи — индифферентны, безразличны друг к другу. Поэтому процесс решения многокритериальных задач неизбежно связан с экспертными оценками как самих критериев, так и взаимоотношений между ними. Известен ряд методов решения задач многокритериальной оптимизации:

— оптимизация одного признанного наиболее важным критерия, остальные критерии при этом играют роль дополнительных ограничений;

— упорядочение заданного множества критериев и последовательная оптимизация по каждому из них (этот подход рассмотрен ниже на примере *метода последовательных уступок*;

— сведение многих критериев к одному введением экспертных весовых коэффициентов для каждого из критериев таким образом, что более важный критерий получает более высокий вес.

Возвращаясь к задаче многокритериальной оптимизации в общей постановке (3.28), (3.29), отметим, что в идеальном случае можно вести поиск такого решения, которое принадлежит пересечению множеств оптимальных решений всех однокритериальных задач. Однако такое пересечение обычно оказывается пустым множеством, поэтому приходится рассматривать так называемое переговорное множество *эффективных решений* (*оптимальных по Парето*). Критерий оптимальности итальянского экономиста В. Парето применяется при решении таких задач, когда оптимизация означает улучшение одних показателей при условии, чтобы другие не ухудшались.

Определение 3.1. Вектор $\bar{X}^* \in Q$ называется эффективным (оптимальным по Парето) решением задачи (3.28), (3.29), если не существует такого вектора $\bar{X} \in Q$, что

$$Z_i(\bar{X}) \geq Z_i(\bar{X}^*), \quad i = \overline{1,m}, \tag{3.30}$$

причем хотя бы для одного значения i имеет место строгое неравенство.

Множество допустимых решений, для которых невозможно одновременно улучшить все частные показатели эффективности (т.е. улучшить хотя бы один из них, не ухудшая остальных), принято называть *областью Парето*, или *областью компромиссов*, а принадлежащие ей решения — *эффективными*, или *оптимальными по Парето*.

В общем случае эффективные решения не эквивалентны друг другу, так что про два оптимальных по Парето решения нельзя сказать, какое из них лучше. Поэтому при решении многокритериальных задач необходимо дополнительное изучение эффективных решений. Для этого можно было бы сформулировать некоторый критерий и оптимизировать его на множестве эффективных решений. Однако при этом возникают значительные трудности в связи с тем, что, как правило, область компромиссов не является выпуклой, и полученная задача в общем случае будет задачей невыпуклого программирования. Обычный подход заключается в стремлении «свернуть» частные критерии в один обобщенный скалярный критерий, оптимизация которого приводит к оптимальному решению задачи в целом. Формулировка подходящего обобщенного критерия в зависимости от конкретных условий как раз и является основным вопросом, который изучается в многокритериальной оптимизации.

В некоторых случаях вместо одного обобщенного критерия и решения одной соответствующей задачи скалярной оптимизации предлагается рассматривать последовательность обобщенных критериев и последовательность задач скалярной оптимизации. К сожалению, многие из описанных в литературе подобных процедур не всегда приводят к эффективным решениям.

Рассмотрим один из таких методов решения многокритериальных задач — *метод последовательных уступок*.

Метод последовательных уступок решения задач многокритериальной оптимизации применяется в случае, когда частные критерии могут быть упорядочены в порядке убывания их важности. Предположим, что все частные критерии максимизируются и пронумерованы в порядке убывания их важности. Находим максимальное значение Z_1^*, первого по важности критерия в области допустимых решений, путем решения однокритериальной задачи

$$\begin{aligned} &Z_1(\bar{X}) \to \max, \\ &\bar{X} \in Q. \end{aligned}$$

Затем, исходя из практических соображений и принятой точности, назначается величина допустимого отклонения $\delta_1 > 0$ (экономически оправданной уступки) критерия Z_1 и находится максимальное значение второго критерия Z_2^* при условии, что значение первого критерия не должно отклоняться от своего максимального значения более чем на величину допустимой уступки, т.е. решается задача:

$$Z_2(\overline{X}) \to \max,$$
$$Z_1(\overline{X}) \geq Z_1^* - \delta_1,$$
$$\overline{X} \in Q.$$

Снова назначается величина уступки $\delta_2 > 0$ по второму критерию, которая вместе с первой уступкой используется для нахождения условного максимума третьего частого критерия:

$$Z_3(\overline{X}) \to \max,$$
$$Z_1(\overline{X}) \geq Z_1^* - \delta_1,$$
$$Z_2(\overline{X}) \geq Z_2^* - \delta_2,$$
$$\overline{X} \in Q.$$

Аналогичные процедуры повторяются до тех пор, пока не будет выявлено максимальное значение последнего по важности критерия Z_m при условии, что значение каждого из первых $m - 1$ частных критериев отличается от соответствующего условного максимума не более чем на величину допустимой уступки по данному критерию. Полученное на последнем этапе решение считается оптимальным. Следует заметить, что этот метод не всегда приводит к эффективному решению.

Пример 3.7. Решение задачи многокритериальной оптимизации методом последовательных уступок.

Решение. Пусть задача трехкритериальной оптимизации имеет вид

$$Z_1 = -x_1 + 2x_2 \to \max \tag{3.31}$$

$$Z_2 = 2x_1 + x_2 \to \max \tag{3.32}$$

$$Z_1 = x_1 - 3x_2 \to \max \tag{3.33}$$

$$\left.\begin{array}{l} x_1 + x_2 \leq 6 \\ 1 \leq x_1 \leq 3 \\ 1 \leq x_2 \leq 4 \end{array}\right\}. \tag{3.34}$$

Заметим, что так как коэффициенты при одних и тех же переменных в данных частных критериях имеют разные знаки, то в заданной области допустимых решений невозможно одновременно улучшить все частные критерии, т.е. в рассматриваемом случае область компромиссов (область Парето) совпадает с областью допустимых решений (3.34).

Для определенности будем считать, что допустимые уступки по первым двум критериям заданы: $\delta_1 = 3$; $\delta_3 = 5/3$.

Максимизируем функцию Z_1 в области допустимых решений, т.е. решаем одну критериальную задачу (3.31), (3.34). Это несложно сделать рассмотренным в главе 2 графическим методом решения задач линейного программирования (рис. 3.3).

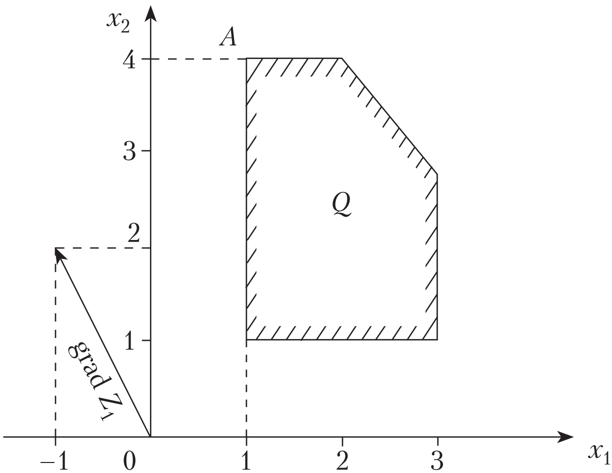


Рис. 3.3

Максимум функции Z_1 при условиях (3.34) достигается в точке A области Q с координатами (1; 4), так что в данном случае

$$x_1^* = 1; \quad x_2^* = 4; \quad \max Z_1 = Z_1^* = Z_1(A) = 7.$$

Переходим к максимизации функции Z_1 при условиях (3.34) и дополнительном ограничении, позволяющем учесть, что по критерию Z_1 нельзя уступать более чем на δ_1. Так как в нашем примере $Z_1^* - \delta_1 = 4$, то дополнительное ограничение будет иметь вид

$$-x_1 + 2x_2 \geq 4. \tag{3.35}$$

Задачу (3.32), (3.34), (3.35) также решаем графически (рис. 3.4).

Получаем, что максимум функции Z_2 при условиях (3.34), (3.35) достигается в точке B части Q_1 области Q, так что

$$x_1^{**} = 8/3; \quad x_2^{**} = 10/3; \quad \max Z_2 = Z_2^* = Z_2(B) = 26/3.$$

Теперь уступаем по критерию Z_2 на величину уступки $\delta_2 = 5/3$ и получаем второе дополнительное ограничение:

$$2x_1 + x_2 \geq 7. \tag{3.36}$$

Максимизируем функцию Z_3 при условиях (3.34), (3.35) и (3.36). Решение этой задачи представлено на рис 3.5.

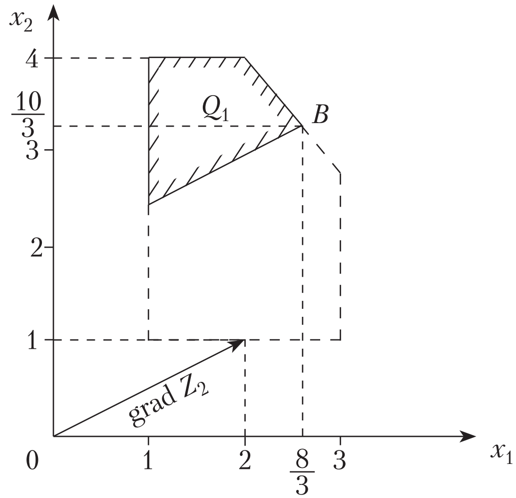


Рис. 3.4

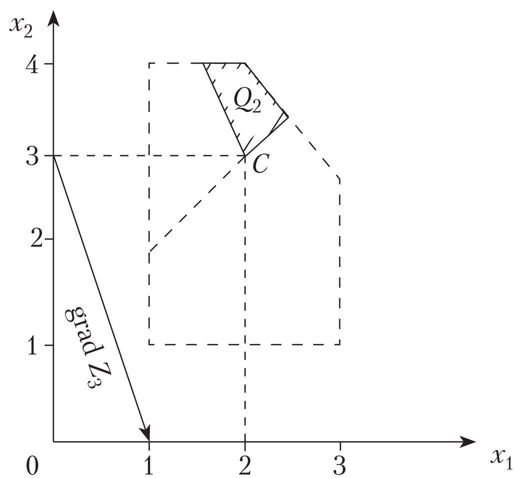


Рис. 3.5

Таким образом, получаем оптимальное решение рассматриваемой трехкритериальной задачи (точка C на рис. 3.5):

$$x_1 = 2;\ x_2 = 3.$$

Соответствующие значения частных критериев при этом составляют:

$$Z_1 = 4;\ Z_2 = 7;\ Z_3 = -7.$$

3.5. Нелинейное и динамическое программирование; понятие об имитационном моделировании

Задача нелинейного программирования формулируется так же, как и общая задача оптимального программирования, т.е. в виде (3.37)—(3.39), со следующими требованиями к целевой функции и допустимой области: целевая функция $f(\bar{X})=(x_1, x_2, \ldots, x_n)$ или (и) хотя бы одна из функций $g_i(x_1, x_2, \ldots, x_n), i=\overline{1,m}$ являются нелинейными:

$$\max(\min) f(x_1, x_2, \ldots, x_n), \tag{3.37}$$

$$g_i(x_1, x_2, \ldots, x_n)\{\leq,=,\geq\} b_i, \ \ i=\overline{1,m}, \tag{3.38}$$

$$x_i \geq 0, \ j=\overline{1,n}. \tag{3.39}$$

Напомним, что для задач линейного программирования характерно следующее.

- Множество допустимых решений выпукло. Это выпуклое множество имеет конечное число вершин, называемых обычно крайними (угловыми) точками.
- Множество всех точек $\left[\overline{x_1}, \ldots \overline{x_n}\right]$ n-мерного пространства, в которых целевая функция принимает заданное значение, есть гиперплоскость (линия) уровня. Кроме того, гиперплоскости, соответствующие разным значениям целевой функции, параллельны.
- Локальный max или min является также глобальным max или min целевой функции на множестве допустимых решений, т.е. не существует локального оптимума целевой функции, отличного от глобального оптимума.
- Если оптимальное значение целевой функции ограничено, то по крайней одна крайняя точка множества допустимых решений является оптимальным решением. Кроме того, начав с произвольной вершины множества допустимых решений, можно достичь оптимума за конечное число шагов, причем на каждом шаге совершается переход только в соседнюю вершину. Окончательно данная вершина является оптимальной тогда и только тогда, когда значение целевой функции в ней, по крайней мере, не меньше, чем значения целевой функции во всех примыкающих вершинах.

У произвольной задачи нелинейного программирования некоторые или все приведенные выше свойства ЗЛП отсутствуют. Вследствие этого задачи нелинейного программирования несравненно сложнее задач ЛП и для них не существует общего универсального метода решения (аналогичного симплексному методу).

Пример 3.8. Найти max и min функции $Z = f(x_1, x_2) = x_1^2 + x_2^2$ при ограничениях

$$\begin{cases} x_1 \cdot x_2 \le 4 \\ x_1 + x_2 \ge 5 \\ x_1 \le 7 \\ x_2 \le 6 \\ x_1 \ge 0, x_2 \ge 0. \end{cases}$$

Решение. Линии уровня целевой функции (Z = const) представляют собой окружности с центром в начале координат; область допустимых решений не является выпуклой и состоит из двух отдельных частей (на рис. 3.6 эти части заштрихованы, линии уровня показаны пунктиром). Минимальное значение функции Z = 17 достигается в точках A(1; 4) и L(4; 1). Функция Z имеет два локальных максимума: в точке D(2/3; 6), где функция $Z(D) = \dfrac{328}{9}$, и в точке M(7; 4/7), в которой функция $Z(M)$ = 2417/49.

Точка M — точка глобального максимума (рис. 3.6).

Особое место занимают задачи типа

$$\max(\min) Z = f(x_1, x_2, \ldots, x_n), \tag{3.40}$$

$$g_i(x_1, x_2, \ldots, x_n) = 0, \; i = \overline{1, m}, \tag{3.41}$$

для решения которых можно воспользоваться *классическим методом оптимизации Лагранжа*, или методом *разрешающих множителей*. При этом предполагают, что функции f и $g_i (i = \overline{1, m})$ непрерывны вместе со своими первыми частными производными.

Для решения задачи составляют *функцию Лагранжа*

$$F(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i g_i(x_1, x_2, \ldots, x_n),$$

определяют частные производные этой функции по переменным $x_j (j = \overline{1, n})$ и множителям Лагранжа $\lambda_i (i = \overline{1, m})$, приравнивают их нулю и получают систему уравнений:

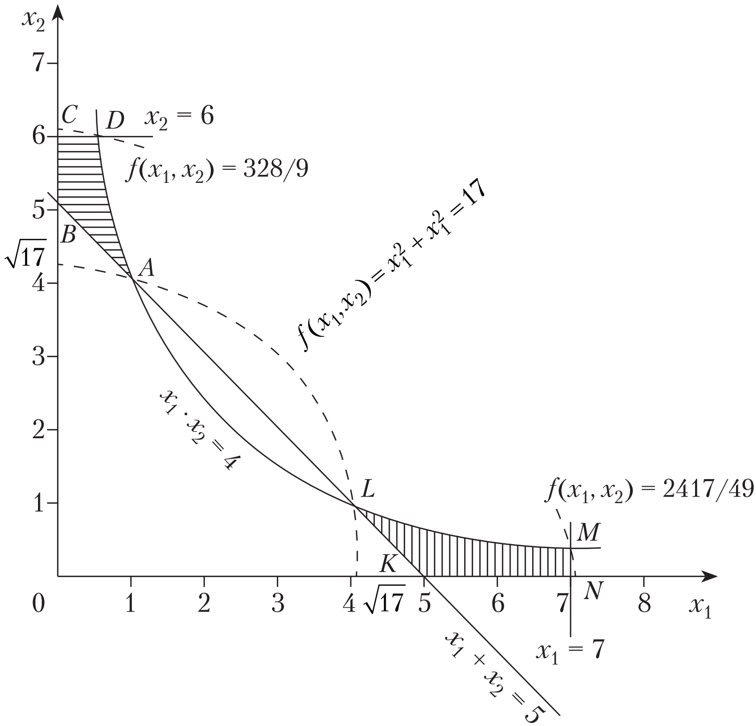


Рис. 3.6

$$\begin{cases} \dfrac{\partial F}{\partial x_j} = \dfrac{\partial f}{\partial x_j} + \sum\limits_{i=1}^{m} \lambda_i \dfrac{\partial g_i}{\partial x_j} = 0, \ j = \overline{1,n}, \\ \dfrac{\partial F}{\partial \lambda_i} = g_i(x_1, x_2, \ldots, x_n) = 0, \ i = \overline{1,m}. \end{cases} \tag{3.42}$$

В основе метода Лагранжа решения классической задачи оптимизации (3.40), (3.41) лежит утверждение, что если функция $Z = f(x_1, x_2, \ldots, x_n)$ в точке $\overline{X^0} = (x_1^0, x_2^0, \ldots, x_n^0)$ имеет экстремум, то существует такой вектор $(\lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0)$, что точка $(x_1^0, x_2^0, \ldots, x_n^0, \lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0)$ является решением системы (3.42). Следовательно, решая систему (3.42), получаем множество стационарных точек, в которых функция Z может иметь экстремальные значения. При этом, как правило, неизвестен способ определения точек глобального максимума или минимума. Однако если решения системы найдены, то для определения глобального максимума или минимума достаточно найти значения функции в соответствующих точках области определения.

Пример 3.9. Найти экстремум функции $Z = x_1x_2 + x_2x_3$ при ограничениях

$$x_1 + x_2 = 2,$$
$$x_2 + x_3 = 2.$$

Решение. Составляем функцию Лагранжа

$$F(x_1, x_2, x_3, \lambda_1, \lambda_2) = x_1x_2 + x_2x_3 + \lambda_1(x_1 + x_2 - 2) + \lambda_2(x_2 + x_3 - 2),$$

дифференцируем ее по переменным $x_1, x_2, x_3, \lambda_1, \lambda_2$ и полученные выражения приравниваем нулю:

$$\begin{cases} \lambda_1 + x_2 = 0, \\ x_1 + x_3 + \lambda_1 + \lambda_2 = 0, \\ x_2 + \lambda_2 = 0, \\ x_1 + x_2 - 2 = 0, \\ x_2 + x_3 - 2 = 0. \end{cases}$$

Из первого и третьего уравнений следует, что $\lambda_1 = \lambda_2 = -x_2$, поэтому

$$x_1 - 2x_2 + x_3 = 0,$$
$$x_1 + x_2 = 2,$$
$$x_2 + x_3 = 2,$$

откуда $x_1^0 = x_2^0 = x_3^0 = 1$ и $Z^0 = 2$. Поскольку, например, точка (0; 2; 0) принадлежит допустимой области и в ней Z = 0, то делаем вывод, что точка (1; 1; 1) — точка глобального максимума.

К классу задач нелинейного программирования, изученному наиболее основательно, относятся задачи с линейными ограничениями и нелинейной целевой функцией. В общем виде такая задача записывается следующим образом:

$$\text{найти } \max(\min) Z = f(x_1, x_2, \ldots, x_n),$$

$$\sum_{j=1}^{n} a_{ij}x_j \{\le, =, \ge\} b_i,\ i = \overline{1, m},$$

$$x_j \ge 0,\ j = \overline{1, n}.$$

Отметим, что даже для задач с линейными ограничениями вычислительные методы разработаны лишь в тех случаях, когда целевая функция имеет определенные свойства, например, функция Z — *сепарабельная*, т.е.

$$Z = f(x_1, x_2, \dots, x_n) = f_1(x_1) + f_2(x_2) + \dots + f_n(x_n).$$

Чтобы гарантировать возможность отыскания оптимального решения, на функции $f_j(x_j)$ должны быть наложены добавочные ограничения. Другим примером могут служить задачи, в которых целевая функция может быть записана как сумма линейной и квадратичной форм, так что

$$\begin{aligned} Z = f(x_1, x_2, \dots, x_n) = \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} x_i x_j = c_1 x_1 + c_2 x_2 + \dots \\ + c_n x_n + d_{11} x_1^2 + d_{12} x_1 x_2 + \dots \\ + d_{1n} x_1 x_n + \dots + d_{nn} x_n^2. \end{aligned}$$

Такие нелинейные задачи называются задачами квадратичного программирования. Чтобы быть уверенным, что оптимальное решение и в этом случае может быть найдено, на величины d_{ij} следует наложить некоторые ограничения.

Имеются достаточно эффективные методы решения задачи *выпуклого программирования*, т.е. задачи минимизации нелинейной, но гладкой выпуклой функции при ограничениях, заданных нелинейными неравенствами, определяющими выпуклое множество переменных:

$$\min f(x_1, \dots, x_n), \tag{3.43}$$

$$g_i(x_1, x_2, \dots, x_n) \le b_i,\ i = \overline{1,m}, \tag{3.44}$$

$$x_j \ge 0,\ j = \overline{1,n}. \tag{3.45}$$

В случае максимизации в таких задачах целевая функция должна быть вогнутой. Симплексный алгоритм для решения общей задачи ЛП представляет собой итеративную процедуру, с помощью которой точное оптимальное решение может быть получено за конечное число шагов. Для задач нелинейного программирования вычислительный метод, дающий точное оптимальное решение за конечное число шагов, удается построить не всегда. Здесь часто приходится соглашаться на использование методов, дающих только приближенное решение или требующих для сходимости бесконечного числа шагов.

Один из наиболее мощных методов решения задач нелинейного программирования состоит в преобразовании задачи каким-либо образом к виду, допускающему применение симплексного алгоритма. Природа «преобразования», с помощью которого нелинейная задача может быть приведена к форме, допускающей применение симплексного метода, очень сильно зависит от типа задачи. В некоторых случаях не требуется никакой предварительной аппроксимации, в других аппроксимация нужна. Однако эта аппроксимация может быть сделана сколь угодно точной (ценой увеличения объема вычислений).

Широко применяется *градиентный метод*. Он представляет собой итеративную процедуру, в которой переходят шаг за шагом от одного допустимого решения к другому так, что значение целевой функции улучшается. Однако в отличие от симплексного метода ЛП в нем не используется переход от одной вершины к другой. Вообще говоря, для сходимости к решению здесь требуется бесконечное число итераций.

Широкое применение нашли методы штрафных функций и барьеров. *Метод штрафных функций* аппроксимирует задачу с ограничениями задачей без ограничений с функцией, которая налагает штраф за выход из допустимой области.

Идея *метода барьеров* аналогична методу штрафных функций, однако аппроксимация здесь осуществляется «изнутри» допустимой области.

Весьма полезным вычислительным методом для решения некоторых типов задач нелинейного программирования является метод динамического программирования (ДП). При решении задачи этим методом процесс решения расчленяется на этапы, решаемые последовательно во времени и приводящие, в конечном счете, к искомому решению. Типичные особенности многоэтапных (многошаговых) задач, решаемых методом динамического программирования, состоят в следующем.

- Процесс перехода производственно-экономической системы из одного состояния в другое должен быть марковским (процессом с отсутствием последействия). Это значит, что если система находится в некотором состоянии $S^n \in S_n$, то дальнейшее развитие процесса зависит только от данного состояния и не зависит от того, каким путем система приведена в это состояние.
- Процесс длится определенное число шагов N. На каждом шаге осуществляется выбор одного управления u^n, под

воздействием которого система переходит из одного состояния S^n в другое $S^{n+1}: S^n \xrightarrow{u^n} S^{n+1}$. Поскольку процесс марковский, то $u^n = u^n(S^n)$ зависит только от текущего состояния.

- Каждый шаг (выбор очередного решения) связан с определенным эффектом, который зависит от текущего состояния и принятого решения: $\varphi_n(S^n, u^n)$.
- Общий эффект (доход) за N шагов слагается из доходов на отдельных шагах, т.е. критерий оптимальности должен быть аддитивным (или приводящимся к нему).

Требуется найти такое решение u^n для каждого шага (n = 1, 2, 3, ..., N), т.е. последовательность $(u^1, \ldots, u^N)$, чтобы получить максимальный эффект (доход) за N шагов.

Любая возможная допустимая последовательность решений $(u^1, \ldots, u^N)$ называется *стратегией управления*. Стратегия управления, доставляющая максимум критерию оптимальности, называется *оптимальной*.

В основе общей концепции метода ДП лежит *принцип оптимальности Беллмана*.

Оптимальная стратегия обладает таким свойством, что независимо от того, каким образом система оказалась в рассматриваемом конкретном состоянии, последующие решения должны составлять оптимальную стратегию, привязывающуюся к этому состоянию. Математически этот принцип записывается в виде *рекуррентного соотношения ДП* (РДП):

$$f_n(S^n) = \max\left\{\varphi_n(S^n, u^n) + f_{n-1}(S^{n-1}, u^n)\right\},$$

$$u^n \in u^n(S^n), S^n \in S_n,$$

где $u^n(S^n)$ — все допустимые управления при условии, что система находится в состоянии S^n;
$\varphi_n(S^n, u^n)$ — эффект от принятия решения u^n;
$f_n(S^n)$ — эффект за оставшиеся n шагов.

Благодаря принципу оптимальности удается при последующих переходах испытывать не все возможные варианты, а лишь оптимальные выходы. РДП позволяют заменить трудоемкое вычисление оптимума по N переменным в исходной задаче решением N задач, в каждой из которых оптимум находит лишь по одной переменной.

Имеется очень много практически важных задач, которые ставятся и решаются как задачи ДП (задачи о замене оборудования, о ранце, распределения ресурсов и т.д.)

В качестве примера построения РДП рассмотрим использование принципа оптимальности для реализации математической модели задачи оптимального распределения некоторого ресурса в объеме x:

$$\max\{\varphi_1(x_1)+\varphi_2(x_2)+\ldots+\varphi_N(x_N)\},$$
$$x_1+x_2+\ldots+x_N=x,$$
$$x_j\geq 0;\ j=\overline{1,N},$$

где x_j — количество ресурса, используемое j-м способом; $\varphi_j(x_j)$ — доход от применения способа j, $j=\overline{1,N}$.

Рекуррентные соотношения, с помощью которых находится решение этой задачи, имеют вид:

$$f_1(x)=\max\{\varphi_1(x_1)\},$$
$$0\leq x_1\leq x,$$
$$f_n(x)=\max\{\varphi_n(x_n)+f_{n-1}(x-x_n)\},\ n=\overline{2,N},$$
$$0\leq x_n\leq x.$$

Пример 3.10. Найти максимум функции

$$f(\bar{X})=3x_1^2-4x_2+3x_3^3$$

при ограничениях

$$4x_1+3x_2+2x_3\leq 8,$$
$$x_j\geq 0,\ x_j \text{ — целые};\ j=1,\ 2,\ 3.$$

Решение. Целевая функция задачи $f(\bar{X})$ является аддитивной, так как ее можно представить в виде суммы функций $f_j(x_j)$, каждая из которых зависит только от одной переменной x_j:

$$f(\bar{X})=f_1(x_1)+f_2(x_2)+f_3(x_3),$$

где $f_1(x_1)=3x_1^2;\ f_2(x_2)=-4x_2;\ f_3(x_3)=3x_3^3.$

Находим $f_1(\lambda)=\max 3x_1^2$. Поскольку на переменные x_j накладываются условия целочисленности и неотрицательности, то $0\leq x_1\leq\left[\dfrac{\lambda}{4}\right]$ (знак $[\ ]$ обозначает целую часть числа, т.е. наибольшее целое число, не превосходящее данное); таким образом, $x_1\in\{0,\ 1,\ 2\}$. Для каждого фиксированного значения λ вычисляем значение функции $f_1(\lambda)$ и выбираем среди них максимальное, при этом в соответствии с ограничениями задачи λ может принимать все целые значения от 0 до 8. Далее вычисляем:

$$f_2(\lambda)=\max\left\{-4x_2+f_1(\lambda-3x_2)\right\},$$

$$\text{для всех } 0\le x_2\le\left[\frac{\lambda}{3}\right];$$

$$f_3(8)=\max\left\{3x_3^3+f_2(8-2x_3)\right\}$$

$$\text{для всех } 0\le x_3\le\left[\frac{8}{2}\right].$$

Все вычисления приведены в табл. 3.16.

Таким образом, $\max f(\bar{X})=\max f_3(\lambda)=f_3^*(8)=192.$ Оптимальную стратегию находим следующим образом. Сначала устанавливаем, что $x_3^*=4$ (соответствует максимальному значению 192). Значение $x_2^*=0$ находим из соответствующих граф табл. 3.16 для $\lambda=8-2x_3^*=8-2\cdot 4=0$ $(f_2(\lambda)=f_2(0)=0$ при $x_2^*=0)$.

Далее находим значение $x_1^*=0$ для $\lambda=8-2x_3^*-3x_2^*=$ $=8-2\cdot 4=0$ $(f_1(\lambda)=f_1(0)=0$ при $x_1^*=0$). Таким образом, оптимальная стратегия имеет вид (0; 0; 4).

Таблица 3.16

Результаты расчетов в соответствии с РДП

λ	$f_1(\lambda)$			$f_2(\lambda)$			$f_3(\lambda)$				
	$x_1=0$	$x_1=1$	$x_1=2$	$x_2=0$	$x_2=1$	$x_2=2$	$x_3=0$	$x_3=1$	$x_3=2$	$x_3=3$	$x_3=4$
0	0			0							
1	0			0							
2	0			0							
3	0			0	−4						
4	0	3		0	−4						
5	0	3		0	−4						
6	0	3		0	−4	−8					
7	0	3		0	−1	−8					
8	0	3	12	0	−1	−8	12	6	27	81	192*

Рассмотрим далее ряд основных понятий, связанных с имитационным *моделированием*. Во всех рассматриваемых выше оптимизационных моделях, так или иначе, предполагалась возможность использования аналитических методов решения, однако для многих задач анализа и управления в экономике такой возможности не существует. Если изучаемые процессы имеют явно нелинейный характер и при этом осложнены разного рода вероятностными характеристиками, то о практически полезном аналитическом решении не может быть и речи. В этих случаях могут быть применены методы *машинной имитации*, т.е. методы экспериментального изучения социально-экономических систем с помощью ЭВМ. Машинная имитация применяется тогда, когда реальный экономический эксперимент по каким-либо причинам невозможен, и тогда имитация выступает в качестве замены реального эксперимента либо в качестве предварительного этапа, позволяющего принять более обоснованное решение о проведении такого эксперимента.

При машинной имитации формируется так называемая *имитационная система*, в которую входят *имитационная модель*, имитирующая исследуемый процесс, и набор алгоритмов и программ, предназначенных как для обеспечения диалога человека и ЭВМ (*внутреннее математическое обеспечение*), так и для решения задач типа ввода и вывода информации, формирования базы данных и т.д. (*внешнее математическое обеспечение*). Имитационная модель при этом сама является своего рода программой для ЭВМ. Практическое применение этой модели заключается в наблюдении за результатами весьма многовариантных расчетов по такой программе при различных задаваемых значениях вводимых экзогенных переменных. В процессе анализа этих результатов могут быть сделаны выводы о поведении системы без ее построения, если эта система только проектируется, без вмешательства в ее функционирование, если это действующая система, и без ее разрушения, если целью эксперимента является определение пределов воздействия на систему. Таким образом, могут быть достигнуты цели экономико-математического моделирования в тех случаях, когда аналитическое решение невозможно.

Процесс последовательной разработки имитационной модели начинается с создания простой модели, которая затем постепенно усложняется в соответствии с предъявляемыми решаемой проблемой требованиями. В каждом цикле имитационного моделирования можно выделить следующие этапы.

1. Формирование проблемы: описание исследуемой проблемы и определение целей исследования.
2. Разработка модели: логико-математическое описание моделируемой системы в соответствии с формулировкой проблемы.
3. Подготовка данных: идентификация, спецификация и сбор данных.
4. Трансляция модели: перевод модели со специальных имитационных языков, используемых на этапе 2 (СИМУЛА, СЛАМ и др.), на язык, приемлемый для используемой ЭВМ.

5. Верификация: установление правильности машинных программ.

6. Валидация: оценка требуемой точности и адекватности имитационной модели.

7. Планирование: определение условий проведения машинного эксперимента с имитационной моделью.

8. Экспериментирование: многократный прогон имитационной модели на ЭВМ для получения требуемой информации.

9. Анализ результатов: изучение результатов имитационного эксперимента для подготовки выводов и рекомендаций по решению проблемы.

10. Реализация и документирование: реализация рекомендаций, полученных на основе имитации, и составление документации по модели и ее использованию.

Имитационные модели часто используются для принятия решений в условиях риска (напомним, что в классе моделей принятия решений в условиях риска мы можем сделать предположения об исходах случайных параметров и вероятностях наступления каждого возможного состояния).

В этом случае в основе имитационного моделирования лежит метод статистического моделирования (метод Монте-Карло). Этот метод позволяет воспроизводить на компьютере случайные величины (с.в.) с заданными законами распределения. Так как отдельные реализации этих с.в. получены искусственно, то их реализации называют псевдослучайными числами. Процедуры получения псевдослучайных чисел называют *генераторами (датчиками) псевдослучайных чисел*.

Например, при проведении имитационных экспериментов в среде MS Excel (при использовании стандартных офисных средств) могут быть использованы встроенные датчики, генерирующие семь типов распределений: равномерное, Пуассона, нормальное, Бернулли, биномиальное, модельное и дискретное (опции **Инструменты анализа/Генерация случайных чисел** надстройки **Пакет анализа**).

Примеры имитационного моделирования систем управления запасами и массового обслуживания средствами MS Excel приведены, например, в литературе[1].

---

[1] *Гармаш А. Н., Орлова И. В.* Математические методы в управлении: учеб. пособие. М. : Вузовский учебник, 2012.

Обязательным элементом «моделирующего алгоритма» в реальных имитационных экспериментах является оценка точности результатов, полученных методом статистических испытаний.

Для любых законов распределения случайной величины с помощью неравенства Чебышева можно получить верхнюю оценку ошибки, т.е. ошибка не может быть больше результата, полученного методом статистических испытаний. При нормальном законе распределения результата, полученного методом статистических испытаний, ошибка будет определяться неравенством

$$\left|\bar{x} - m\right| \leq \frac{3\bar{\sigma}}{\sqrt{N}}.$$

Это неравенство дает точную оценку возможной ошибки. Таким образом, для определения ошибки результата моделирования случайной величины необходимо вначале по данным N испытаний вычислить среднее статистическое $\bar{x}$, статистическое среднеквадратичное отклонение $\bar{\sigma}$ и лишь затем оценить ошибку результата на основе приведенного неравенства. Если ошибка окажется больше приемлемой, то потребуется увеличение числа испытаний N.

Так, имитируя работу одноканальной СМО с отказами, надо получить N раз реализации x_i соответственно с.в. длительности интервалов между отдельными поступлениями требований и с.в. длительности интервалов по обслуживанию требований и с помощью «моделирующего алгоритма» зафиксировать число отказов в такой системе. При большом числе испытаний N получим, например, *достаточно точную* статистическую оценку вероятности отказа в обслуживании (конечно, предварительно проводится статистическая оценка гипотез о характере законов распределения входного потока требований и длительности интервалов обслуживания).

3.6. Модели сетевого планирования и управления

Во многих областях экономики, технологии, проектирования, строительства, научных исследований важное значение имеют задачи оптимизации распределения ресурсов (трудовых, финансовых и др.). Особую значимость приобретают эти задачи в условиях реализации новых проектов, когда выполняется огромное количество взаимозаменяемых операций, в работу вовлекается множество работников, предприятий, организаций, так как в этих случаях управление работами усложняется новизной разработки, трудностью точного определения сроков и затрат ресурсов на том или ином этапе. Высокоэффективными инструментами для решения таких задач являются *сетевые методы и модели*.

Основные понятия сетевого моделирования

Сетевой моделью (другие названия: *сетевой график, сеть*) называется экономико-математическая модель, отражающая комплекс работ (операций) и событий, связанных с реализацией некоторого проекта (научно-исследовательского, производственного и др.), в их логической и технологической последовательности и связи. Анализ сетевой модели, представленной в графической или табличной (матричной) форме, позволяет, во-первых, более четко выявить взаимосвязи этапов реализации проекта и, во-вторых, определить наиболее оптимальный порядок выполнения этих этапов в целях, например, сокращения сроков выполнения всего комплекса работ. Таким образом, методы сетевого моделирования можно отнести к методам принятия оптимальных решений.

Математический аппарат сетевых моделей базируется на теории графов. *Графом* называется совокупность двух конечных множеств: множества точек, которые называются *вершинами*, и множества пар вершин, которые называются *ребрами*. Представление о графе можно получить, если рассмотреть некоторый геометрический многогранник, например куб; в кубе можно выделить два конечных множества, состоящих соответственно из восьми вершин и двенадцати ребер.

Если рассматриваемые пары вершин являются упорядоченными, т.е. на каждом ребре задается направление, то граф называется *ориентированным*; в противном случае — *неориентированным*. Последовательность неповторяющихся ребер, ведущая от некоторой вершины к другой, образует *путь*. Граф называется *связным*, если для любых двух его вершин существует путь, их соединяющий; в противном случае граф называется *несвязным*. В экономике чаще всего используется два вида графов: дерево и сеть. *Дерево* представляет собой связный граф без циклов, имеющий исходную вершину (*корень*) и крайние вершины; пути от исходной вершины к крайним вершинам называются *ветвями*. *Сеть* — это ориентированный конечный связный граф, имеющий начальную вершину (*источник*) и конечную вершину (*сток*). Таким образом, сетевая модель представляет собой граф вида «сеть».

В экономических исследованиях сетевые модели возникают при моделировании экономических систем и процессов методами *сетевого планирования и управления* (*СПУ*).

Объектом управления в системах сетевого планирования и управления являются коллективы исполнителей, располагающие определенными ресурсами и выполняющие заданный комплекс операций, который призван обеспечить достижение намеченной цели, например разработку нового изделия, строительство объекта и т.п.

Основой СПУ служит сетевая модель (СМ), в которой моделируется совокупность взаимосвязанных работ и событий, отображающих процесс достижения определенной цели. Она может быть представлена в виде графика или таблицы.

Основными понятиями СМ являются следующие: работа, событие, путь. На рис. 3.7 графически представлена СМ, состоящая из 5 событий (кружочки) и 6 работ (стрелки); продолжительность выполнения работ в некоторых единицах времени указана над стрелками.

Работа характеризует материальное действие, требующее использования ресурсов, или логическое, требующее лишь взаимосвязи событий. При графическом распределении работа изображается стрелкой, которая соединяет два

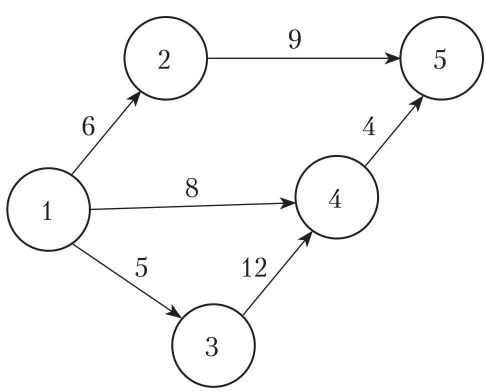


Рис. 3.7. **Сетевая модель**

события. Она обозначается парой заключенных в скобки чисел (i, j), где i — номер события, из которого работа выходит, а j — номер события, в которое она входит. Работа не может начаться раньше, чем свершится событие, из которого она выходит. Каждая работа имеет определенную продолжительность $t(i, j)$. Например, запись $t(2, 5) = 9$ означает, что работа (2, 5) имеет продолжительность 9 единиц времени (см. рис. 3.7). К работам относятся также такие процессы, которые не требуют ни ресурсов, ни времени выполнения. Они заключаются в установлении логической взаимосвязи работ и показывают, что одна из них непосредственно зависит от другой и не может выполняться, прежде чем эта другая будет завершена; такие работы называются *фиктивными* и на графике изображаются пунктирными стрелками.

Событиями называются результаты выполнения одной или нескольких работ. Они не имеют протяженности во времени. Событие свершается в тот момент, когда оканчивается последняя из работ, входящая в него. События обозначаются одним числом и при графическом представлении СМ изображаются кружком (или иной геометрической фигурой), внутри которого проставляется его порядковый номер ($i = 1, 2, \ldots N$). В СМ имеется *начальное событие* (с номером 1), из которого работы только выходят, и *конечное событие* (с номером N), в которое работы только входят.

Путь в СМ — это цепочка следующих друг за другом работ, соединяющих начальную и конечную вершины, например, в приведенной на рис. 3.7 модели путями являются $L_1 = (1, 2, 5)$, $L_2 = (1, 4, 5)$ и др. Продолжительность пути определяется суммой продолжительностей составляющих его работ. Путь, имеющий максимальную длину, называют *критическим* и обозначают $L_{кр}$, а его продолжительность — $t_{кр}$. Работы, принадлежащие критическому пути, называются *критическими*. Их несвоевременное выполнение ведет к срыву сроков всего комплекса работ.

Правила построения сетевых моделей

СМ имеют ряд характеристик, которые позволяют определить степень напряженности выполнения отдельных работ, а также всего их комплекса и принять решение о перераспределении ресурсов в случае необходимости. Однако перед расчетом СМ следует убедиться, что она удовлетворяет следующим основным требованиям.

1. События правильно пронумерованы, т.е. для каждой работы (i, j) $i < j$. При невыполнении этого требования необходимо использовать алгоритм перенумерации событий, который заключается в следующем:

— нумерация событий начинается с исходного события, которому присваивается № 1;

— из исходного события вычеркивают все исходящие из него работы (стрелки) и на оставшейся сети находят событие, в которое не входит ни одна работа, ему и присваивают № 2;

— затем вычеркиваются работы, выходящие из события № 2, и вновь находят событие, в которое не входит ни одна работа, и ему присваивают № 3, и так продолжается до завершающего события, номер которого должен быть равен количеству событий в сетевом графике;

— если при очередном вычеркивании работ одновременно несколько событий не имеют входящих в них работ, то их нумеруют очередными номерами в произвольном порядке.

2. Отсутствуют тупиковые события (кроме завершающего), т.е. такие, за которыми не следует хотя бы одна работа.

3. Отсутствуют события (за исключением исходного), которым не предшествует хотя бы одна работа.

4. Отсутствуют циклы, т.е. замкнутые пути, соединяющие событие с ним же самим.

Более подробно эти требования описаны, например, в [13], [17]. При невыполнении указанных требований не имеет смысла приступать к вычислениям характеристик событий, работ и критического пути.

Расчет характеристик сетевых моделей

Для событий рассчитывают три основные характеристики: ранний и поздний срок совершения события, а также его резерв.

Ранний срок свершения события $t_р$ определяется величиной наиболее длительного отрезка пути от исходного до рассматриваемого события, причем $t_р(1) = 0$, а $t_р(N) = t_{кр}(L_{кр})$:

$$t_р(j) = \max_i\{t_р(i) + t(i, j)\}; \quad j = \overline{2, N}. \tag{3.46}$$

Поздний срок свершения события $t_п$ характеризует самый поздний допустимый срок, к которому должно совершаться событие, не вызывая при этом срыва срока свершения конечного события:

$$t_п(i) = \min_j\{t_п(j) - t(i, j)\}; \quad i = \overline{2, N-1}. \quad (3.47)$$

Этот показатель определяется «обратным ходом», начиная с завершающего события, с учетом соотношения $t_п(N) = t_р(N)$.

Все события, за исключением событий, принадлежащих критическому пути, имеют *резерв* $R(i)$:

$$R(i) = t_п(i) - t_р(i). \quad (3.48)$$

Резерв показывает, на какой предельно допустимый срок можно задержать наступление этого события, не вызывая при этом увеличения срока выполнения всего комплекса работ.

Для всех работ (i, j) на основе ранних и поздних сроков свершения событий можно определить следующие показатели (здесь и в дальнейшем, где это целесообразно, для упрощения записей все подстрочные символы заменяются строчными):

Ранний срок начала — $t_{рн}(i,j) = t_р(i)$, (3.49)

Ранний срок окончания — $t_{ро}(i,j) = t_р(i) + t(i,j)$, (3.50)

Поздний срок окончания — $t_{по}(i,j) = t_п(j)$, (3.51)

Поздний срок начала — $t_{пн}(i,j) = t_п(j) - t(i,j)$, (3.52)

Полный резерв времени — $R_п(i,j) = t_п(j) - t_р(i) - t(i,j)$, (3.53)

Независимый резерв времени —

$$R_н(i,j) = \max\{0; t_р(j) - t_п(i) - t(i,j)\} \quad (3.54)$$

или

$$R_н(i,j) = \max\{0; R_п(i,j) - R(i) - R(j)\}. \quad (3.55)$$

Полный резерв времени показывает, на сколько можно увеличить время выполнения конкретной работы при условии, что срок выполнения всего комплекса работ не изменится.

Независимый резерв времени соответствует случаю, когда все предшествующие работы заканчиваются в поздние сроки, а все последующие — начинаются в ранние сроки. Использование этого резерва не влияет на величину резервов времени других работ.

Путь характеризуется двумя показателями — продолжительностью и резервом. Продолжительность пути определяется суммой продолжительностей составляющих его работ.

Резерв определяется как разность между длинами критического и рассматриваемого путей. Из этого определения следует, что работы, лежащие на критическом пути, и сам критический путь имеют нулевой резерв времени. Резерв времени пути показывает, на сколько может увеличиться продолжительность работ, составляющих данный путь, без изменения продолжительности общего срока выполнения всех работ.

Названные выше характеристики СМ могут быть получены на основе аналитических формул (3.46)—(3.55), а процесс вычислений отображается либо на самом графике СМ, либо в матрице размерности $N \times N$, либо в таблице. Рассмотрим алгоритмы табличного метода расчета сетевого графика на конкретном примере решения задачи организации труда при реализации некоторого проекта.

Пример 3.11. Пусть реализуется проект по строительству АЗС, при этом привлекаемые рабочие могут выполнять любую из выделенных по принятой технологии работ. Менеджер проекта установил, что в данном проекте от начала до завершения работ можно выделить 5 событий, и существуют 6 различных видов работ, связывающих эти события. При первоначальном распределении рабочих по видам работ по имеющимся нормативам трудоемкости были рассчитаны длительности выполнения работ (в днях), таким образом, сетевой график реализации данного проекта имеет вид модели, представленной ранее на рис. 3.7.

Требуется рассчитать основные характеристики событий, работ и всей сетевой модели в целом, а также определить наличие резерва для некоторых работ в целях оптимизации модели и сокращения сроков выполнения проекта за счет перераспределения рабочих по видам работ.

Рассмотрим этапы табличного метода расчета данной сетевой модели; результаты этого расчета приведены в табл. 3.17 в графах 1—9.

Этап 1. Перечень работ и их продолжительность запишем во вторую и третью графы табл. 3.17, при этом работы записываются последовательно в гр. 2: сперва начинающиеся с номера 1, затем с номера 2 и т.д.

Этап 2. В первой графе поставим число $K_{пр}$, показывающее количество работ, непосредственно предшествующих событию i, с которого начинается рассматриваемая работа. Для работ, начинающихся с номера 1, предшествующих работ нет ($K_{пр} = 0$). Для работы, начинающейся на номер k, просматриваются все верхние строчки второй графы и отыскиваются работы, заканчивающиеся на этот номер k. Количество найденных работ записывается

Таблица 3.17

$K_{пр}$	(i, j)	$t(i, j)$	$t_{рн}(i, j) = t_р(i)$	$t_{ро}(i, j)$	$t_{пн}(i, j)$	$t_{по}(i, j) = t_п(j)$	$R_п$	$R_н$	$K_н$
1	2	3	4	5 = 4 + + 3	6 = 7 – – 3	7	8	9	10
0	(1, 2)	6	0	6	0	6	0	0	0,71
0	(1, 3)	5	0	5	0	5	0	0	1
0	(1, 4)	8	0	8	9	17	9	9	0,47
1	(2, 5)	9	6	15	12	21	6	0	0,71
1	(3, 4)	12	5	17	5	17	0	0	1
2	(4, 5)	4	17	21	17	21	0	0	1

в первую графу во все строчки, соответствующие работам, начинающимся с номера k. Например, для работы (4, 5) в графу 1 поставим цифру 2, так как на номер 4 оканчиваются две работы: (1, 4) и (3, 4).

Этап 3. Заполнение таблицы начинается с расчета раннего срока работ $t_р(i)$. Для работ, имеющих цифру «ноль» в первой графе, в графе 4 также заносятся нули и рассчитываются соответствующие значения для графы 5 (ранний срок окончания $t_{ро}(i, j)$) как суммы соответствующих чисел в графе 4 и графе 3 (см. формулу (3.51)). В нашей модели таких работ три; в первой строчке графы 5 ставим 0 + 6 = 6, аналогично по второй и третьей строке.

Для заполнения следующих строк графы 4 для работ (i, j) просматриваются заполненные строки графы 5, содержащие работы, оканчивающиеся на номер i, и максимальное из найденных значений (если их несколько) переносится в графу 4 для обрабатываемых строк. Так, в нашем примере в четвертой строке в графе 4 ставим 6, а в графе 5 — 15 (6 + 9 = 15). Аналогично в пятой строке в графе 4 и графе 5 ставим соответственно 5 и 17 (5 + 12 = 17). В последней шестой строке в графе 4 ставим 17 (наибольшее из чисел 8 и 17 в графе 5) и соответственно в графу 5 ставим 21 (17 + 4 = 21).

Этап 4. Графы 7 и 6 заполняются «обратным ходом», т.е. снизу вверх. Для этого просматриваются строки (работы), оканчивающиеся на номер N последнего события, и из графы 5 выбирается максимальная величина; эта величина записывается в графу 7 по всем строчкам, оканчивающимся на N (см. формулу (3.51)), с учетом равенства $t_п(N) = t_р(N)$. Затем заполняется графа 6 по этим строкам как разность между графой 7 и графой 3 (см. формулу (3.52)). В нашем примере таких строк две (четвертая и шестая), для которых в графе 5 стоят числа 15 и 21; выбираем наибольшее из них (21) и записываем его в графу 7 по этим строкам, после чего заносим соответствующие числа в графу 6.

Далее просматриваются строки, оканчивающиеся на номер события, предшествующего завершающему, т.е. на $(N-1)$. Для этих строк просматриваются все строчки графы 6, лежащие ниже и начинающиеся с номера $(N-1)$; среди них в графе 6 выбирается минимальная величина, которая переносится в графу 7 по обрабатываемым строкам, после чего заполняется графа 6. В нашем примере таких строк две (третья и пятая); ниже их с номера 4 начинается одна (последняя) работа, и в графе 6 стоит 17, следовательно, в графе 7 по этим строчкам ставим число 17, после чего заполняется графа 6.

Затем аналогичный процесс повторяется для строк, оканчивающихся на $(N-2)$, $(N-3)$ и т.д. до тех пор, пока не будут заполнены все строки по графы 7 и 6. В нашем примере результаты приведены в соответствующих графах табл. 3.17.

Этап 5. Показатели графы 8 рассчитываются как разности соответствующих показателей граф 6 и 4 или граф 7 и 5 (см. формулы (3.48) или (3.53)). Чтобы заполнить графу 9, можно предварительно рассчитать резервы времени каждого события по формуле (3.48), а затем воспользоваться формулой (3.55). В нашем примере резервы времени для каждого из пяти событий равны соответственно: $R(1) = 0$; $R(2) = 12 - 6 = 6$; $R(3) = 5 - 5 = 0$; $R(4) = 17 - 17 = 0$; $R(5) = 0$. Последующие результаты по формуле (3.55) приведены в графе 9 табл. 3.17.

Этап 6. На этом этапе подводятся основные итоги расчета. Учитывая, что нулевой резерв времени имеют только работы ($R_п = 0$) и события ($R(i) = 0$), которые принадлежат критическому пути, получаем, что критическим является путь $L_{кр} = (1, 3, 4, 5)$, продолжительность которого ($t_{кр}$) равна 21 дню. Так как работы (1, 2), (1, 4) и (2, 5) имеют ненулевые резервы $R_п$, то очевидно, что путем перевода некоторого числа рабочих с этих работ на работы, принадлежащие критическому пути, можно сократить продолжительность этого пути и тем самым сократить сроки выполнения проекта в целом, т.е. осуществить оптимизацию сетевого графика.

Оптимизация сетевых моделей

Для оптимизации сетевой модели, выражающейся в перераспределении ресурсов с ненапряженных работ на критические для ускорения их выполнения, необходимо как можно более точно оценить степень трудности своевременного выполнения всех работ, а также «цепочек» пути. Более точным инструментом решения этой задачи по сравнению с полным резервом является *коэффициент напряженности*, который может быть вычислен одним их двух способов в соответствии с приведенной ниже формулой:

$$K_н(i,j)=\frac{t(L_{\max})-t_{кр}}{t_{кр}-t'_{кр}}=1-\frac{R_п(i,j)}{t_{кр}-t'_{кр}}, \qquad (3.56)$$

где $t(L_{max})$ — продолжительность максимального пути, проходящего через работу (i, j);
$t'_{кр}$ — общая продолжительность отрезков пути, проходящего через работу (i, j), совпадающих с критическим путем.

Коэффициент напряженности изменяется от нуля до единицы, причем чем он ближе к единице, тем сложнее выполнить данную работу в установленный срок. Самыми напряженными являются работы критического пути, для которых коэффициент напряженности равен 1. На основе значений этого коэффициента все работы СМ могут быть разделены на три группы:

— напряженные ($K_н(i, j) > 0{,}8$);
— подкритические ($0{,}6 < K_н(i, j) < 0{,}8$);
— резервные ($K_н(i, j) < 0{,}6$).

В результате перераспределения ресурсов стараются максимально уменьшить общую продолжительность работ, что возможно при переводе всех работ в первую группу.

Пример 3.12. Рассчитаем коэффициенты напряженности всех работ сетевой модели, рассматриваемой выше в примере 3.11. При расчете этих показателей по формуле (3.56) целесообразно пользоваться графиком данной СМ, представленным на рис. 3.7. Для работ критического пути (1, 3), (3, 4) и (4, 5) коэффициенты напряженности $K_н = 1$. Для других работ:

$K_н(1, 2) = 1 - (6 : (21 - 0)) = 0{,}71$;
$K_н(1, 4) = 1 - (9 : (21 - 4)) = 0{,}47$;
$K_н(2, 5) = 1 - (6 : (21 - 0)) = 0{,}71$.

Результаты расчетов приведены в графе 10 ($K_н$) табл. 3.17. Они показывают, что напряженными являются работы критического пути (1, 3), (3, 4) и (4, 5); работы (1, 2) и (2, 5) являются подкритическими, а работа (1, 4) — резервная. Следовательно, оптимизация рассматриваемой СМ возможна в основном за счет резервной работы (1, 4) и частично за счет подкритических работ (1, 2) и (2, 5).

Если данный расчет СМ проведен до начала работ, менеджер проекта может скорректировать первоначальное распределение рабочих по видам работ, сняв некоторое количество рабочих с работы (1, 4) и, возможно, с работ (1, 2) и (2, 5) и распределив их по работам критического пути (1, 3), (3, 4) и (4, 5).

В результате работы, не лежащие на критическом пути, несколько увеличат свою продолжительность, а продолжительность работ критического пути сократится; тем самым при той же численности рабочих сократится срок реализации проекта в целом. Получив новый вариант сетевого графика, менеджер может повторить аналогичные расчеты, добиваясь путем перераспределения рабочих наиболее оптимального варианта СМ и, следовательно, наилучшего распределения рабочих по видам работ.

Сетевое планирование в условиях неопределенности

В рассмотренных выше примерах сроки выполнения отдельных видов работ определялись при условии взаимозаменяемости рабочих и при наличии нормативов трудоемкости для данных работ. На практике во многих случаях трудно точно определить продолжительность работ, поэтому задаются две оценки этой продолжительности — минимальная и максимальная. Минимальная оценка $t_{\min}(i, j)$ дает продолжительность работ при наиболее благоприятных обстоятельствах, а максимальная $t_{\max}(i, j)$ — при наиболее неблагоприятных. Продолжительность работы в этом случае рассматривается как случайная величина, которая при реализации может принять любое значение в заданном интервале. Такие оценки являются вероятностными, и их ожидаемые значения $t_{\text{ож}}(i, j)$ оцениваются по-разному в зависимости от принятого закона распределения. Так, при бета-распределении плотности вероятности ожидаемое значение продолжительности работ (математическое ожидание) задается формулой

$$t_{\text{ож}}(i,j) = (3\,t_{\min}(i,j) + 2\,t_{\max}(i,j)) : 5 \qquad (3.57)$$

Для характеристики степени разброса возможных значений относительно ожидаемого уровня используется показатель дисперсии:

$$S^2(i,j) = (t_{\max}(i,j) - t_{\min}(i,j))^2 : 5^2 = 0{,}04(t_{\max}(i,j) - t_{\min}(i,j))^2 \qquad (3.58)$$

На основе этих оценок можно рассчитать все характеристики СМ, однако они будут выступать как средние характеристики. При достаточно большом количестве работ можно утверждать, что общая продолжительность любого пути, включая критический, имеет нормальный закон распределения со средним значением, равным сумме средних значений продолжительности составляющих его работ, и дисперсией, равной сумме дисперсий этих же работ.

Кроме основных характеристик СМ, при вероятностном задании продолжительности работ можно решить две важные задачи:

1. определить вероятность того, что продолжительность критического пути $t_{кр}$ не превысит заданного (директивного) уровня T;

2. определить максимальный срок выполнения всего комплекса работ T при заданном уровне вероятности (надежности) p.

Более подробно вопросы сетевого планирования в условиях неопределенности с решением конкретных примеров рассмотрены в ряде учебных и научных изданий (см., например, параграф 3.6 в [15]).

Вопросы и задания

1. Что такое двойственная задача в линейном программировании? Сформулируйте основные теоремы теории двойственности.
2. Поясните экономический смысл теорем двойственности, дайте экономическую интерпретацию свойств двойственных оценок.
3. Опишите экономико-математическую модель транспортной задачи. Какие методы решения транспортных задач вы знаете?
4. Дайте экономическую интерпретацию метода потенциалов решения транспортной задачи.
5. Поясните суть задач целочисленного программирования. Приведите конкретные примеры таких задач и назовите известные вам методы их решения.
6. В чем сущность задач многокритериальной оптимизации? Дайте характеристику метода последовательных уступок.
7. Опишите общую постановку задачи нелинейного программирования. В чем суть метода Лагранжа решения классической оптимизационной задачи?
8. Дайте краткую характеристику задач динамического программирования и методов их решения.
9. Раскройте основные понятия имитационного моделирования и перечислите этапы машинной имитации как экспериментального метода изучения социально-экономических процессов.
10. В чем суть методов сетевого планирования и управления? Дайте содержательную характеристику элементов сетевого графика.
11. Какие задачи решаются на основе сетевых моделей? Раскройте сущность сетевого планирования в условиях неопределенности.

Упражнения

1. Сформулируйте двойственные задачи для задач линейного программирования, приведенных в примерах 3.6, 3.7, 3.8 и 3.9 и в упражнениях 2, 3 и 4 главы 2.

2. При решении задачи оптимального использования ресурсов на максимум прибыли, исходные данные которой приведены в таблице,

Тип ресурса	Нормы затрат ресурсов на единицу продукции			Наличие ресурсов
	1	2	3	
1. Труд	1	4	3	200
2. Сырье	1	1	2	80
3. Оборудование	1	2	2	130
Цена единицы продукции	40	60	80	

был получен оптимальный план: $x_1 = 40$; $x_2 = 40$; $x_3 = 0$. Требуется:

а) сформулировать прямую оптимизационную задачу, указать оптимальную производственную программу;

б) сформулировать двойственную задачу и найти ее оптимальный план на основе теорем двойственности;

в) проанализировать использование ресурсов в оптимальном плане, найти норму относительной заменяемости дефицитных ресурсов;

г) определить, как изменится максимум общей стоимости продукции и план выпуска при увеличении запасов сырья на 9 ед. и одновременном уменьшении трудовых ресурсов на 3 ед.;

д) оценить целесообразность включения в план выпуска продукции четвертого вида, если цена единицы этой продукции составляет 70 ден. ед., а на ее производство расходуется по 2 ед. ресурсов каждого типа.

3. Решить следующие транспортные задачи (здесь A — вектор мощностей поставщиков; B — вектор мощностей потребителей; C — матрица транспортных издержек на единицу груза):

а) $A = (100; 150; 50)$, $B = (75; 80; 60; 85)$,

$$C = \begin{pmatrix} 6 & 7 & 3 & 5 \\ 1 & 2 & 5 & 6 \\ 8 & 10 & 20 & 1 \end{pmatrix};$$

б) $A = (300; 350; 150; 200)$, $B = (400; 400; 200)$,

$$C = \begin{pmatrix} 4 & 1 & 2 \\ 3 & 4 & 2 \\ 1 & 3 & 1 \\ 1 & 4 & 3 \end{pmatrix};$$

в) $A = (20; 30; 40; 20)$, $B = (40; 40; 20)$,

$$C = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ 4 & 1 & 2 \\ 3 & 1 & 1 \end{pmatrix}.$$

4. Найти целочисленные решения следующих задач линейного программирования методом Гомори:

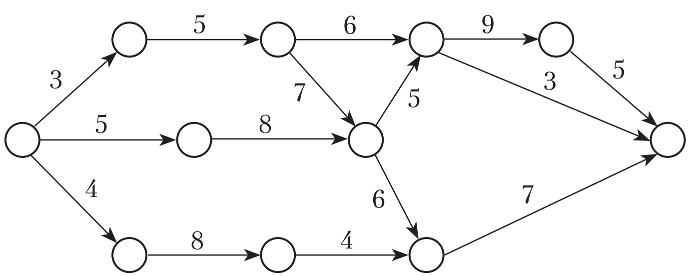
а) $\max f(X) = 3x_1 + 3x_2$,
$x_1 + 3x_2 \geq 6$,
$3x_1 + 2x_2 \leq 36$,
$x_2 \leq 13$,
$x_1, x_2 \geq 0$;

б) $\max f(X) = 3x_1 + 4x_2$,
$3x_1 + 2x_2 \leq 8$,
$x_1 + 4x_2 \leq 10$,
$x_1, x_2 \geq 0$.

5. С помощью метода Лагранжа найти условный экстремум функционала Z;

а) $Z = x_1 \cdot x_2$ при $x_1^2 + x_2^2 = 2$;
б) $Z = x_1^2 + x_2^2$ при $x_1 + x_2 = 2$,
$x_1, x_2 \geq 0$;
в) $Z = x_1 + x_2$ при $1/x_1 + 1/x_2 = 1$.

6. Сетевой график с указанием продолжительности работ в днях приведен на рисунке:



Требуется:

а) пронумеровать события;
б) выделить критический путь и найти его длину;
в) определить резервы времени каждого события;
г) определить полные резервы времени некритических работ.

Глава 4
МЕТОДЫ И МОДЕЛИ АНАЛИЗА ДИНАМИКИ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

— Понятия экономических рядов динамики
— Предварительный анализ и сглаживание временных рядов экономических показателей
— Расчет показателей динамики развития экономических процессов
— Анализ сезонных колебаний в экономике

Изучение данной главы должно позволить студентам:

***знать*:**
- основные понятия временных рядов экономической динамики;
- методы предварительного анализа и сглаживания временных экономических рядов;
- методы расчета основных показателей динамики экономических процессов;
- особенности анализа сезонности в экономике;

***уметь*:**
- дать общую характеристику временного экономического ряда и выделить его структурно образующие элементы;
- осуществить предварительный анализ и сглаживание экономического ряда динамики на основе статистических методов;
- рассчитать основные показатели экономической динамики, а также провести анализ сезонных колебаний в экономических процессах;

***владеть*:**
- понятийным аппаратом математико-статистического анализа временных экономических рядов;
- основными методами предварительной обработки временных рядов и выявления сезонной волны в этих рядах.

4.1. Понятия экономических рядов динамики

Динамические процессы, происходящие в экономических системах, чаще всего проявляются в виде ряда последовательно расположенных в хронологическом порядке значений того или иного показателя, который в своих изменениях отражает ход развития изучаемого явления в экономике. Эти значения, в частности, могут служить для обоснования (или отрицания) различных моделей социально-экономических систем, в том числе изученных в предыдущих главах. Они служат также основой для разработки прикладных моделей особого вида, называемых трендовыми моделями, которые будут рассмотрены в главе 5.

Прежде всего, дадим ряд определений. Последовательность наблюдений одного показателя (признака), упорядоченных в зависимости от последовательно возрастающих или убывающих значений другого показателя (признака), называют *динамическим рядом,* или *рядом динамики*. Если в качестве признака, в зависимости от которого происходит упорядочение, берется время, то такой динамический ряд называется *временным рядом*. Так как в экономических процессах, как правило, упорядочение происходит в соответствии со временем, то при изучении последовательных наблюдений экономических показателей все три приведенных выше термина используются как равнозначные. Составными элементами рядов динамики являются, таким образом, цифровые значения показателя, называемые *уровнями* этих рядов, и моменты или интервалы времени, к которым относятся уровни.

Временные ряды, образованные показателями, характеризующими экономическое явление на определенные моменты времени, называются *моментными;* пример такого ряда представлен в табл. 4.1.

Таблица 4.1

Списочная численность предприятия

Дата	1/I	1/II	1/III	1/IV	30/IV
Списочная численность рабочих	4100	4400	4200	4600	4800

Если уровни временного ряда образуются путем агрегирования за определенный промежуток (интервал) времени, то такие ряды называются *интервальными* временными радами; пример приведен в табл. 4.2.

Таблица 4.2

Фонд заработной платы рабочих предприятия

Месяц	Январь	Февраль	Март	Апрель
Фонд заработной платы рабочих, тыс. руб.	37 187,5	38 270,0	39 380,0	42 535,0

Временные ряды могут быть образованы как из абсолютных значений экономических показателей, так и из средних или относительных величин — это производные ряды; пример такого ряда дан в табл. 4.3.

Таблица 4.3

Среднемесячная заработная плата рабочих предприятия

Месяц	Январь	Февраль	Март	Апрель
Средняя заработная плата рабочих, тыс. руб.	8750	8900	8950	9050

Под *длиной* временного ряда понимают время, прошедшее от начального момента наблюдения до конечного. Таким образом, длина всех приведенных выше временных рядов равна четырем месяцам. Часто длиной ряда называют количество уровней, входящих во временной ряд; длина ряда из табл. 4.1 равна пяти, а из табл. 4.2 и 4.3 — четырем.

Если во временном ряду проявляется длительная («вековая») тенденция изменения экономического показателя, то говорят, что имеет место тренд. Таким образом, под *трендом* понимается изменение, определяющее общее направление развития, основную тенденцию временных рядов. В связи с этим экономико-математическая динамическая модель, в которой развитие моделируемой экономической системы отражается через тренд ее основных показателей, называется *трендовой моделью*. Для выявления тренда во временных рядах, а также для построения и анализа трендовых моделей используется аппарат, разработанный для простых статистических совокупностей. Отличие временных экономических рядов от простых статистических совокупностей заключается, прежде всего, в том, что последовательные значения уровней временного ряда зависят друг от друга. Поэтому применение выводов и формул теории вероятностей и математической статистики требует известной осторожности при анализе временных рядов, особенно при экономической интерпретации результатов анализа.

Предположим, имеется временной ряд, состоящий из n уровней: $y_1, y_2, y_3, ..., y_n$.

В самом общем случае временной ряд экономических показателей можно разложить на четыре структурно образующих элемента:

— тренд, составляющие которого будем обозначать U_t, $t = 1, 2, ..., n$;

— сезонная компонента, обозначаемая через V_t, $t = 1, 2, ..., n$;

— циклическая компонента, обозначаемая через C_t, $t = 1, 2, ..., n$;

— случайная компонента, которую будем обозначать ε_t, $t = 1, 2,..., n$.

Под трендом, как уже отмечалось выше, понимается устойчивое систематическое изменение процесса в течение продолжительного времени.

Во временных рядах экономических процессов могут иметь место более или менее регулярные колебания. Если они носят строго периодический или близкий к нему характер и завершаются в течение одного года, то их называют *сезонными колебаниями*. В тех случаях, когда период колебаний составляет несколько лет, то говорят, что во временном ряде присутствует *циклическая компонента*.

Тренд, сезонная и циклическая компоненты называются *регулярными*, или систематическими, компонентами временного ряда. Составная часть временного ряда, остающаяся после выделения из него регулярных компонент, представляет собой *случайную*, нерегулярную компоненту. Она является обязательной составной частью любого временного ряда в экономике, так как случайные отклонения неизбежно сопутствуют любому экономическому явлению. Если систематические компоненты временного ряда определены правильно, что как раз и составляет одну из главных целей при разработке трендовых моделей, то остающаяся после выделения из временного ряда этих компонент так называемая *остаточная последовательность* (ряд остатков) будет случайной компонентой ряда, т.е. будет обладать следующими свойствами:

— случайностью колебаний уровней остаточной последовательности;

— соответствием распределения случайной компоненты нормальному закону распределения;

— равенством математического ожидания случайной компоненты нулю;

— независимостью значений уровней случайной последовательности, т.е. отсутствием существенной автокорреляции.

Проверка адекватности трендовых моделей основана на проверке выполняемости у остаточной последовательности указанных четырех свойств. Если не выполняется хотя бы одно из них, модель признается неадекватной; при выполнении всех четырех свойств модель адекватна. Данная проверка осуществляется с использованием ряда статистических критериев и рассмотрена более подробно ниже. Отметим, что в дальнейшем мы не будем рассматривать циклическую компоненту временных рядов; укажем только, что для моделирования и прогнозирования сезонных и циклических экономических процессов используются специальные методы (индексный и спектральный анализы, выравнивание по ряду Фурье и др.). Методы анализа сезонности и тренд-сезонных экономических процессов рассмотрим в параграфе 4.4.

4.2. Предварительный анализ и сглаживание временных рядов экономических показателей

Предварительный анализ временных рядов экономических показателей заключается в основном в выявлении и устранении аномальных значений уровней ряда, а также в определении наличия тренда в исходном временном ряде. Рассмотрим эти операции более подробно.

Под *аномальным уровнем* понимается отдельное значение уровня временного ряда, которое не отвечает потенциальным возможностям исследуемой экономической системы и которое, оставаясь в качестве уровня ряда, оказывает существенное влияние на значения основных характеристик временного ряда, в том числе на соответствующую трендовую модель. Причинами аномальных наблюдений могут быть ошибки технического порядка, или *ошибки первого рода*: ошибки при агрегировании и дезагрегировании показателей, при передаче информации и другие технические причины. Ошибки первого рода подлежат выявлению и устранению. Кроме того, аномальные уровни во временных рядах могут возникать из-за воздействия факторов, имеющих объективный характер, но проявляющихся эпизодически, очень редко — *ошибки второго рода*; они устранению не подлежат.

Для выявления аномальных уровней временных рядов используются методы, рассчитанные для статистических совокупностей.

Метод Ирвина, например, предполагает использование следующей формулы:

$$\lambda_t = \frac{\left| y_t - y_{t-1} \right|}{\sigma_y}, \quad t = 2, 3, \ldots, n, \tag{4.1}$$

где среднеквадратическое отклонение σ_y рассчитывается в свою очередь с использованием формул:

$$\sigma_y = \sqrt{\frac{\sum_{t=1}^{n} \left(y_t - \overline{y} \right)^2}{n-1}}; \quad \overline{y} = \frac{\sum_{t=1}^{n} y_t}{n}.$$

Расчетные значения λ_2, λ_3 и т.д. сравниваются с табличными значениями критерия Ирвина λ_α, и если оказываются больше табличных, то соответствующее значение y_t уровня ряда считается аномальным. Значения критерия Ирвина для уровня значимости $\alpha = 0{,}05$, т.е. с 5%-ной ошибкой, приведены в табл. 4.4.

Таблица 4.4

n	2	3	10	20	30	60	100
λ_α	2,8	2,3	1,5	1,3	1,2	1,1	1,0

После выявления аномальных уровней ряда обязательно определение причин их возникновения. Если точно установлено, что они вызваны ошибками первого рода, то они устраняются либо заменой аномальных уровней простой средней арифметической двух соседних уровней ряда, либо заменой аномальных уровней соответствующими значениями по кривой, аппроксимирующей данный временной ряд. Порядок нахождения такой кривой, т.е. трендовой модели, рассматривается в гл. 5.

Для определения *наличия тренда* в исходном временном ряду применяется несколько методов; рассмотрим два из них.

Метод проверки разности средних уровней. Реализация этого метода состоит из четырех этапов.

На *первом этапе* исходный временной ряд y_1, y_2, y_3, ... y_n разбивается на две примерно равные по числу уровней части: в первой части n_1 первых уровней исходного ряда, во второй — n_2 остальных уровней ($n_1 + n_2 = n$).

На *втором этапе* для каждой из этих частей вычисляются средние значения и дисперсии:

$$\bar{y}_1 = \frac{\sum_{t=1}^{n_1} y_t}{n_1}; \quad \sigma_1^2 = \frac{\sum_{t=1}^{n_1}\left(y_t - \bar{y}_1\right)^2}{n_1 - 1};$$

$$\bar{y}_2 = \frac{\sum_{t=n_1+1}^{n} y_t}{n_2}; \quad \sigma_2^2 = \frac{\sum_{t=n_1+1}^{n}\left(y_t - \bar{y}_2\right)^2}{n_2 - 1}.$$

Третий этап заключается в проверке равенства (однородности) дисперсий обеих частей ряда с помощью ***F*-критерия Фишера**, которая основана на сравнении расчетного значения этого критерия:

$$F = \begin{cases} \sigma_1^2/\sigma_2^2, & \text{если} \quad \sigma_1^2 > \sigma_2^2 \\ \sigma_2^2/\sigma_1^2, & \text{если} \quad \sigma_1^2 < \sigma_2^2 \end{cases}$$

с табличным (критическим) значением критерия Фишера F_α с заданным *уровнем значимости* (уровнем ошибки) α. В качестве α чаще всего берут значения 0,1 (10%-ная ошибка), 0,05 (5%-ная ошибка), 0,01 (1%-ная ошибка). Величина $1 - \alpha$ называется *доверительной вероятностью*.

Если расчетное значение F меньше табличного F_α, то гипотеза о равенстве дисперсий принимается и переходят к четвертому этапу. Если F больше или равно F_α, гипотеза о равенстве дисперсий отклоняется и делается вывод, что данный метод для определения наличия тренда ответа не дает.

На *четвертом этапе* проверяется гипотеза об отсутствии тренда с использованием ***t*-критерия Стьюдента**. Для этого определяется расчетное значение критерия Стьюдента по формуле

$$t = \frac{\left|\bar{y}_1 - \bar{y}_2\right|}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_1}}}, \tag{4.2}$$

где σ — среднеквадратическое отклонение разности средних:

$$\sigma = \sqrt{\frac{\left(n_1 - 1\right)\sigma_1^2 + \left(n_2 - 1\right)\sigma_2^2}{n_1 + n_2 - 2}}.$$

Если расчетное значение t меньше табличного значения статистики Стьюдента t_α с заданным уровнем значимости α, гипотеза принимается, т.е. тренда нет, в противном случае тренд есть. Заметим, что в данном случае табличное значение t_α берется для числа степеней свободы, равного $n_1 + n_2 - 2$, при этом данный метод применим только для рядов с монотонной тенденцией.

Метод Фостера — Стьюарта. Этот метод обладает большими возможностями и дает более надежные результаты по сравнению с предыдущим. Кроме тренда самого ряда (как говорят, тренда в среднем), он позволяет установить наличие тренда дисперсии временного ряда: если тренда дисперсии нет, то разброс уровней ряда постоянен; если дисперсия увеличивается, то ряд «раскачивается» и т.д.

Реализация метода также содержит четыре этапа.

На *первом этапе* производится сравнение каждого уровня исходного временного ряда, начиная со второго уровня, со всеми предыдущими, при этом определяются две числовые последовательности:

$$k_t = \begin{cases} 1, \text{ если } y_t \text{ больше всех предыдущих уровней} \\ 0 - \text{в противном случае,} \end{cases}$$

$$l_t = \begin{cases} 1, \text{ если } y_t \text{ меньше всех предыдущих уровней} \\ 0 - \text{в противном случае,} \end{cases}$$

$$t = 2, 3, \ldots, n.$$

На *втором этапе* вычисляются величины s и d:

$$s = \sum_{t=2}^{n} \left(k_t - l_t\right);$$

$$d = \sum_{t=2}^{n} \left(k_t + l_t\right).$$

Нетрудно заметить, что величина s, характеризующая изменение временного ряда, принимает значения от $-(n-1)$ (ряд монотонно убывает) до $(n-1)$ (ряд монотонно возрастает). Величина d характеризует изменение дисперсии уровней временного ряда и изменяется от 0 (все уровни ряда равны) до $(n-1)$ (ряд монотонный или с чередованием подъемов и падений уровней).

Третий этап заключается в проверке гипотез: можно ли считать случайными:

1) отклонение величины d от величины μ — математического ожидания величины d для ряда, в котором уровни расположены случайным образом;

2) отклонение величины s от нуля.

Эта проверка проводится с использованием расчетных значений t-критерия Стьюдента для дисперсии и для средней:

$$t_d = \frac{|d-\mu|}{\sigma_1}; \quad \sigma_1 = \sqrt{2\sum_{t=2}^{n} 1/t - 4\sum_{t=2}^{n} 1/t^2};$$

$$t_s = \frac{|s-0|}{\sigma_2}; \quad \sigma_2 = \sqrt{2\sum_{t=2}^{n} 1/t},$$

где μ — математическое ожидание величины d, определенной для ряда, в котором уровни расположены случайным образом;

σ_1 — среднеквадратическое отклонение для величины d;

σ_2 — среднеквадратическое отклонение для величины s.

Для удобства имеются табулированные значения величин μ, σ_1 и σ_2; фрагмент этих значений представлен в табл. 4.5.

Таблица 4.5

n	10	20	30	40
μ	3,858	5,195	5,990	6,557
σ_1	1,288	1,677	1,882	2,019
σ_2	1,964	2,279	2,447	2,561

На *четвертом этапе* расчетные значения t_s и t_d сравниваются с табличным значением t-критерия Стьюдента с заданным уровнем значимости t_α. Если расчетное значение меньше табличного, то гипотеза об отсутствии соответствующего тренда принимается; в противном случае тренд есть. Например, если t_s больше табличного значения t_α, а t_d меньше t_α, то для данного временного ряда имеется тренд в среднем, а тренда дисперсии уровней ряда нет. Пример определения наличия тренда методом Фостера — Стьюарта будет приведен в параграфе 4.4.

Перейдем к вопросу сглаживания временных рядов экономических показателей. Очень часто уровни экономических рядов динамики колеблются, при этом тенденция развития экономического явления во времени скрыта случайными отклонениями уровней в ту или иную сторону.

С целью более четко выявить тенденцию развития исследуемого процесса, в том числе для дальнейшего применения методов прогнозирования на основе трендовых моделей, производят *сглаживание* (*выравнивание*) временных рядов.

Методы сглаживания временных рядов делятся на две основные группы:

1) аналитическое выравнивание с использованием кривой, проведенной между конкретными уровнями ряда так, чтобы она отображала тенденцию, присущую ряду, и одновременно освобождала его от незначительных колебаний;

2) механическое выравнивание отдельных уровней временного ряда с использованием фактических значений соседних уровней.

Методы аналитического выравнивания на основе кривых роста рассматриваются в главе 5. Суть методов механического сглаживания заключается в следующем. Берется несколько первых уровней временного ряда, образующих *интервал сглаживания*. Для них подбирается полином, степень которого должна быть меньше числа уровней, входящих в интервал сглаживания; с помощью полинома определяются новые, выравненные значения уровней в середине интервала сглаживания. Далее интервал сглаживания сдвигается на один уровень ряда вправо, вычисляется следующее сглаженное значение и т.д. Самым простым методом механического сглаживания является **метод простой скользящей средней**.

Сначала для временного ряда $y_1, y_2, y_3, ..., y_n$ определяется интервал сглаживания m ($m < n$). Если необходимо сгладить мелкие беспорядочные колебания, то интервал сглаживания берут по возможности большим; интервал сглаживания уменьшают, если нужно сохранить более мелкие колебания. При прочих равных условиях интервал сглаживания рекомендуется брать нечетным. Для первых m уровней временного ряда вычисляется их средняя арифметическая; это будет сглаженное значение уровня ряда, находящегося в середине интервала сглаживания. Затем интервал сглаживания сдвигается на один уровень вправо, повторяется вычисление средней арифметической и т.д. Для вычисления сглаженных уровней ряда y применяется формула

$$\overline{y}_t = \frac{t = \sum\limits_{t=t-p}^{t+p} y_t}{m}, \quad t > p,$$

где $p = \frac{m-1}{2}$ (при нечетном m); для четных m формула усложняется.

В результате такой процедуры получаются $n - m + 1$ сглаженных значений уровней ряда; при этом первые p и последние p уровней ряда теряются (не сглаживаются).

Другой недостаток метода в том, что он применим лишь для рядов, имеющих линейную тенденцию.

Метод взвешенной скользящей средней отличается от предыдущего метода сглаживания тем, что уровни, входящие в интервал сглаживания, суммируются с разными весами. Это связано с тем, что аппроксимация ряда в пределах интервала сглаживания осуществляется с использованием полинома не первой степени, как в предыдущем случае, а степени, начиная со второй и выше. Используется формула средней арифметической взвешенной:

$$\overline{y}_t = \frac{\sum\limits_{t=t-p}^{t+p} \rho_t y_t}{\sum\limits_{t=t-p}^{t+p} \rho_t},$$

причем веса ρ_t определяются с помощью метода наименьших квадратов. Эти веса рассчитаны для различных степеней аппроксимирующего полинома и различных интервалов сглаживания. Так, для полиномов второго и третьего порядков числовая последовательность весов при интервале сглаживания $m = 5$ имеет вид: $\{-3; 12; 17; 12; -3\}$, а при $m = 7$ имеет вид: $\{-2; 3; 6; 7; 6; 3; -2\}$. Для полиномов четвертой и пятой степеней и при интервале сглаживания $m = 7$ последовательность весов выглядит следующим образом: $\{5; -30; 75; 131; 75; -30; 5\}$.

К этой же группе методов выравнивания временных рядов примыкает **метод экспоненциального сглаживания**. Его особенность заключается в том, что в процедуре нахождения сглаженного уровня используются значения только предшествующих уровней ряда, взятые с определенным весом, причем вес наблюдения уменьшается по мере удаления его от момента времени, для которого определяется сглаженное значение уровня ряда. Если для исходного временного ряда $y_1, y_2, y_3, ..., y_n$ соответствующие сглаженные значения уровней обозначить через S_t,

$t = 1, 2, ..., n$, то экспоненциальное сглаживание осуществляется по формуле

$$S_t = \alpha y_t + (1-\alpha)S_{t-1}, \tag{4.3}$$

где α — параметр сглаживания ($0 < \alpha < 1$); величина $1 - \alpha$ называется коэффициентом дисконтирования.

Используя приведенное выше рекуррентное соотношение для всех уровней ряда, начиная с первого и кончая моментом времени t, можно получить, что экспоненциальная средняя, т.е. сглаженное данным методом значение уровня ряда, является взвешенной средней всех предшествующих уровней:

$$S_t = \alpha \sum_{i=0}^{t-1} (1-\alpha)^i y_{t-i} + (1-\alpha)^t S_0,$$

здесь S_0 — величина, характеризующая начальные условия.

В практических задачах обработки экономических временных рядов рекомендуется (необоснованно) выбирать величину параметра сглаживания в интервале от 0,1 до 0,3. Других точных рекомендаций для выбора оптимальной величины параметра α пока нет. В отдельных случаях Р. Браун предлагает определять величину α исходя из длины сглаживаемого ряда:

$$\alpha = \frac{2}{n+1}.$$

Что касается начального параметра S_0, то в конкретных задачах его берут или равным значению первого уровня ряда y_1, или равным средней арифметической нескольких первых членов ряда, например членов y_1, y_2, y_3:

$$S_0 = \frac{y_1 + y_2 + y_3}{3}.$$

Указанный выше порядок выбора величины S_0 обеспечивает хорошее согласование сглаженного и исходного рядов для первых уровней. Если при подходе к правому концу временного ряда сглаженные этим методом значения при выбранном параметре α начинают значительно отличаться от соответствующих значений исходного ряда, необходимо перейти на другой параметр сглаживания. Заметим, что при этом методе сглаживания не теряются ни начальные, ни конечные уровни сглаживаемого временного ряда.

4.3. Расчет показателей динамики развития экономических процессов

Этот расчет проводится на основе статистического анализа одномерных временных рядов экономической динамики. Для статистического анализа одномерных временных рядов экономических показателей вида $y_1, y_2, y_3, ..., y_n$ абсолютные уровни моментных и интервальных рядов (см. для примера табл. 4.1 и 4.2), а также уровни из средних величин (см. табл. 4.3) должны быть преобразованы в относительные величины. Их можно получить соотнесением уровней ряда с одним и тем же уровнем, взятым за базу (за базу сравнения чаще всего принимают начальный уровень временного ряда y_1), либо последовательными сопоставлениями с предыдущим уровнем. В первом случае получают *базисные* показатели, во втором — *цепные*.

Временной ряд тогда правильно отражает объективный процесс развития экономического явления, когда уровни этого рада состоят из однородных, *сопоставимых* величин. Для несопоставимых величин вести расчет рассматриваемых ниже статистических показателей динамики неправомерно. Причины несопоставимости уровней временного ряда могут быть различными. В экономике чаще всего такими причинами является несопоставимость:

— по территории ввиду изменения границ региона, по которому собираются статистические данные;

— по кругу охватываемых объектов по подчинению или форме собственности ввиду перехода, например, части предприятий данного объединения в другое объединение;

— по временным периодам, когда, например, данные за различные годы приведены по состоянию на разные даты;

— уровней, вычисленных в различном масштабе измерения;

— уровней ряда из-за различий в структуре совокупности, для которой они вычислены. Например, данные о рождаемости населения зависят не только от изменений числа родившихся и численности населения, но и от изменения возрастного состава населения в течение периода наблюдения.

Возможны и другие причины несопоставимости.

При анализе временных рядов для определения изменений, происходящих в данном явлении, прежде всего вычисляют скорость развития этого явления во времени. Показателем скорости служит *абсолютный прирост,* вычисляемый по формуле

$$\Delta y_i = y_i - y_{i-k}, \tag{4.4}$$

где y_i — i-й уровень временного ряда (i = 2, 3, ..., n); индекс k = 1, 2, ..., n – 1 определяет начальный уровень и может быть выбран любым в зависимости от целей исследования: при k = 1 получаются цепные показатели, при $k = i - 1$ получаются базисные показатели с начальным уровнем ряда в качестве базисного и т.д.

Абсолютный прирост выражает величину изменения показателя за интервал времени между сравниваемыми периодами. Если подходить более строго, то скоростью называют прирост в единицу времени; эта величина носит название *среднего абсолютного прироста*:

$$\overline{\Delta y_i} = \frac{y_i - y_{i-k}}{k}. \tag{4.5}$$

В частности, средний абсолютный прирост за весь период наблюдения для данного временного ряда равен

$$\overline{\Delta y} = \frac{y_n - y_1}{n-1} \tag{4.6}$$

и характеризует среднюю скорость изменения временного ряда.

Для определения относительной скорости изменения изучаемого явления в единицу времени используют относительные показатели: коэффициенты роста и прироста (если эти показатели выражены в процентах, то их называют соответственно темпами роста и прироста). Заметим, что во всех последующих формулах индекс начального уровня, по отношению к которому осуществляется сопоставление, определяется точно так же с помощью индекса k, как и ранее для показателя абсолютного прироста.

Коэффициент роста для i-го периода вычисляется по формуле

$$K_{i(\rho)} = \frac{y_i}{y_{i-k}}. \tag{4.7}$$

$K_{i(\rho)} > 1$, если уровень повышается; $K_{i(\rho)} < 1$, если уровень понижается; при $K_{i(\rho)} = 1$ уровень не меняется.

Коэффициент прироста равен

$$K_{i(\text{пр})} = K_{i(\rho)} - 1,$$

или

$$K_{i(\text{пр})} = \frac{y_i - y_{i-k}}{y_{i-k}}. \tag{4.8}$$

На практике чаще применяют показатели *темпа роста* и *темпа прироста*:

$$T_{i(\rho)} = \frac{y_i}{y_{i-k}} \cdot 100\%, \tag{4.9}$$

где $T_{i(\rho)}$ — темп прироста для i-го периода;

$$T_{i(\text{пр})} = T_{i(\rho)} - 100\%,$$

или

$$T_{i(\text{пр})} = \frac{y_i - y_{i-k}}{y_{i-k}} \cdot 100\%, \tag{4.10}$$

где $T_{i(\text{пр})}$ — темп прироста для i-го периода.

Темп прироста показывает, на сколько процентов уровень одного периода увеличился (уменьшился) по сравнению с уровнем другого периода, т.е. этот показатель выражает относительную величину прироста в процентах. Сравнение абсолютного прироста и темпа прироста за одни и те же промежутки времени показывает, что в реальных экономических процессах замедление темпа прироста часто не сопровождается уменьшением абсолютных приростов.

Абсолютное значение одного процента прироста определяется как отношение абсолютного прироста Δy_t к темпу прироста в процентах $T_{i(\text{пр})}$.

Среднюю скорость изменения изучаемого явления за рассматриваемый период характеризует также *средний темп роста*. Обычно он рассчитывается по формуле средней геометрической:

$$\overline{T}_{(\text{р})} = \sqrt[n-1]{T_{1(\text{р})} \cdot T_{2(\text{р})} \cdot \ldots \cdot T_{n(\text{р})}} = \sqrt[n-1]{\frac{y_n}{y_1}} \cdot 100\%, \tag{4.11}$$

где $T_{1(\text{р})}$, ..., $T_{n(\text{р})}$ — средние темпы роста за отдельные интервалы времени.

Соответственно *средний темп прироста* определяется как

$$T_{(\text{пр})} = \overline{T}_{(\text{пр})} - 100\%. \tag{4.12}$$

Показатель среднего темпа роста, рассчитываемый по приведенной выше формуле средней геометрической, имеет существенные недостатки, так как основан на сопоставлении конечного и начального уровней временного ряда, промежуточные уровни во внимание не принимаются. В случае сильной колеблемости уровней использование для статистического анализа среднего геометрического темпа роста может привести к серьезным просчетам в результате искажения реальной тенденции временного ряда.

В настоящее время предложены другие способы расчета среднего темпа роста, в той или иной мере лишенные недостатков средней геометрической. Например, предлагается использовать для расчета среднего темпа роста формулу

$$\overline{T}_{(\text{р})} = \sqrt[n-1]{\frac{\hat{y}_n}{\hat{y}_1}}, \qquad (4.13)$$

где $\hat{y}_1$ и $\hat{y}_n$ — сглаженные по уравнению тренда (уравнению кривой роста) начальный и конечный уровни временного ряда. Порядок получения уравнения тренда, т.е. порядок построения трендовой модели, рассмотрен в главе 5. Трендовая модель учитывает колеблемость промежуточных уровней временного ряда, поэтому вычисленные по ней значения $\hat{y}_1$ и $\hat{y}_n$, а следовательно, и средний темп роста, вычисляемый по последней формуле, будут более точно характеризовать изменения изучаемого экономического явления за рассматриваемый интервал времени.

Важной характеристикой временного ряда является также *средний уровень ряда*. В интервальном ряду динамики с равноотстоящими во времени уровнями расчет среднего уровня ряда производится по формуле простой средней арифметической (здесь и далее суммирование ведется по всем периодам наблюдения):

$$\overline{y} = \frac{\sum y_t}{n}. \qquad (4.14)$$

Если интервальный ряд имеет неравноотстоящие во времени уровни, то средний уровень ряда (так называемая средняя хронологическая) вычисляется по формуле взвешенной арифметической средней, где роль весов играет продолжительность времени (например, количество лет), в течение которого уровень постоянен:

$$\overline{y} = \frac{\sum y_t t}{\sum t}, \qquad (4.15)$$

где t — число периодов времени, при которых значение уровня y не изменяется.

Для моментного ряда с равноотстоящими уровнями средняя хронологическая рассчитывается по формуле

$$\overline{y} = \frac{1/2y_1 + y_2 + y_3 + \ldots + y_{n-1} + 1/2y_n}{n-1}, \tag{4.16}$$

где n — число уровней ряда.

Средняя хронологическая для моментного временного ряда с равноотстоящими во времени уровнями вычисляется по формуле

$$\overline{y} = \frac{(y_1 + y_2)t_1 + (y_2 + y_3)t_2 + \ldots + (y_{n-1} + y_n)t_{n-1}}{2\sum t}. \tag{4.17}$$

Здесь n — число уровней ряда, а t_i — период времени, отделяющий i-й уровень ряда от $(i + 1)$-го уровня.

При статистическом анализе временных рядов часто возникает необходимость, кроме определения основных характеристик ряда, оценить зависимость изучаемого показателя y_t, от его значений, рассматриваемых с некоторым запаздыванием во времени. Зависимость значений уровней временного ряда от предыдущих (сдвиг на 1), предпредыдущих (сдвиг на 2) и так далее уровней того же временного ряда называется *автокорреляцией* во временном ряду. Для получения числовой характеристики такой внутренней зависимости вычисляют взаимную корреляционную функцию между исходным рядом y_t и этим же рядом, сдвинутым во времени на величину τ. Такая функция называется *автокорреляционной*, она характеризует внутреннюю структуру временного ряда и состоит из множества коэффициентов автокорреляции (нециклических), рассчитываемых по формуле

$$r_\tau = \frac{(n-\tau)\sum\limits_{t=1}^{n-\tau} y_t y_{t+\tau} - \sum\limits_{t=1}^{n-\tau} y_t \sum\limits_{t=1}^{n-\tau} y_{t+\tau}}{\sqrt{\left[(n-\tau)\sum\limits_{t=1}^{n-\tau} y_t^2 - \left(\sum\limits_{t=1}^{n-\tau} y_t\right)^2\right]\left[(n-\tau)\sum\limits_{t=1}^{n-\tau} y_{t+\tau}^2 - \left(\sum\limits_{t=1}^{n-\tau} y_{t+\tau}\right)^2\right]}}. \tag{4.18}$$

Задавая различные значения $\tau = 1, 2, 3, \ldots$ получаем последовательность значений $r_1, r_2, r_3, \ldots$ На практике рекомендуется вычислять такие коэффициенты в количестве от $n/4$ до $n/3$.

График автокорреляционной функции называется *коррелограммом* и показывает величину запаздывания, с которым изменение показателя y_t сказывается на его последующих значениях. Величина сдвига τ, которому соответствует наибольший коэффициент автокорреляции, называется *временным лагом*. В ряде случаев используется упрощенная формула для вычисления коэффициента автокорреляции:

$$r_\tau = \frac{\frac{1}{n-\tau}\sum_{t=1}^{n-\tau}(y_t - \overline{y})(y_{t+\tau} - \overline{y})}{\frac{1}{n}\sum_{t=1}^{n}(y_t - \overline{y})^2}, \qquad (4.18')$$

где $\overline{y}$ — средний уровень ряда (см. формулу (4.14)).

4.4. Методы анализа сезонных колебаний в экономике

Для многих явлений общественной жизни характерны внутригодичные повторяющиеся колебания, которые называются *сезонными*. Сезонные колебания наблюдаются и в различных отраслях народного хозяйства России: при производстве большинства сельскохозяйственных продуктов, их переработке, в строительстве, транспорте, торговле и т.д. Так, например, ежегодно суточные удои молока в специализированных агрофирмах с началом весны начинают возрастать, достигают своего максимума в мае — июле, а затем монотонно снижаются.

Сезонные колебания обычно отрицательно сказываются на работе многих отраслей народного хозяйства, так как обычно приводят к неравномерному использованию трудовых ресурсов, к простоям оборудования, что в свою очередь влечет повышение издержек производства. Поэтому изучение закономерностей, складывающихся в области сезонных колебаний, имеет большое практическое и познавательное значение для преодоления или уменьшения их негативных последствий. Кроме того, изучение сезонности необходимо для внутригодичного планирования, в частности для разбивки годового плана по месяцам с учетом сезонных колебаний. Для анализа сезонности необходимы данные по изучаемому явлению в виде временного ряда за несколько лет, при этом эти данные должны быть как минимум квартальными; для более детального исследования необходимы данные за каждый месяц, каждую декаду года.

Существуют две модели сезонности: *аддитивная* и *мультипликативная*. Аддитивная модель предполагает агрегирование (сложение) отдельных компонент уровней временного ряда, и в зависимости от того, существует или нет тренд (тенденция) в этом ряду, модель может иметь следующий вид:

$$y_t = y_{\text{ср}} + S + \varepsilon \text{ — при отсутствии тренда;}$$
$$y_t = \hat{y}_t + S + \varepsilon \text{ — при наличии тренда,}$$

где y_t — уровень временного ряда в момент времени t; $y_{\text{ср}}$ — средний уровень временного ряда; $\hat{y}_t$ — теоретический уровень ряда по модели тренда; S — сезонная составляющая, измеренная в тех же единицах, что и уровни ряда; ε — случайная компонента, представленная в тех же самых единицах.

В мультипликативной модели уровень временного ряда рассматривается как произведение составляющих его компонент:

$$y_t = \hat{y}_t \cdot K_s \cdot E_t,$$

где y_t — фактический уровень ряда; $\hat{y}_t$ — теоретический уровень ряда в соответствии с трендом; K_s — коэффициент сезонности; E_t — коэффициент случайной компоненты. В отличие от аддитивной модели в мультипликативной модели амплитуда сезонных колебаний при наличии тренда во временном ряду меняется; так, при возрастающем тренде в мультипликативной модели амплитуда сезонных колебаний с течением времени увеличивается. В дальнейшем будем рассматривать аддитивные модели сезонности.

Среди многочисленных методов анализа сезонности в экономических процессах можно выделить три основные группы: итерационные методы фильтрации компонент временного ряда, статистические методы расчета сезонной волны (индексов сезонности) и методы гармонического анализа на основе рядов Фурье.

При рассмотрении методов первой из названных групп следует иметь в виду, что под *фильтрацией* (выделением) компонент временного ряда понимается процесс получения оценок всех трех составляющих ряда $\hat{y}_t$, S, и ε; основная идея итерационных процедур указанной фильтрации заключается в многократном применении расчета скользящей средней и оценке сезонной компоненты S в каждом цикле. Детально эти методы анализа сезонности на примере итерационных методов Четверикова и Шискина — Эйзенпресса описаны в параграфе 4.4 учебного пособия [17]. Рассмотрим далее две другие из указанных групп методов анализа сезонных колебаний.

Статистические методы расчета сезонной волны

Для выявления и измерения интенсивности сезонных колебаний в этих методах используются *индексы сезонности*, совокупность которых принято называть *сезонной волной*. В зависимости от характера динамики применяются различные способы расчета индексов сезонности. В тех случаях, когда средний годовой уровень сезонного явления остается от года к году относительно неизменным, другими словами, во временном ряду отсутствует тренд, применяется *метод простых средних*. Суть метода состоит в определении простой средней за одни и те же месяцы (кварталы, декады) всего изучаемого периода и в сопоставлении этих средних со средней за весь изучаемый период. Следует отметить, что при использовании данных только одного года расчеты могут быть слишком ненадежными в силу элемента случайности. Поэтому на практике для выявления закономерности в сезонных колебаниях используются данные за ряд лет (например, месячные данные за три года). Тогда для каждого месяца рассчитывается средняя величина уровня ряда за три года, после чего на основе полученных данных рассчитывается среднемесячный уровень за весь период наблюдения. Отношение средних для каждого месяца к общему среднемесячному уровню (часто выражаемое в процентах) и образует совокупность индексов сезонности. Покажем расчет индексов сезонности методом простых средних на примере.

Пример 4.1. Есть данные о производстве молока в специализированной агрофирме за каждый месяц в течение трех лет. Требуется рассчитать месячные индексы сезонности (сезонную волну) и отобразить полученные результаты на графике. Исходные данные и результаты расчета представлены в табл. 4.6.

Решение. Для получения индексов сезонности, прежде всего, для каждого месяца рассчитываются среднемесячные уровни за три года. Например, для января этот среднемесячный уровень равен $(602 + 618 + 634)/3 = 618$ и т.д. Расчет этих значений дает возможность избавиться от элементов случайности, имевших место в том или ином году. Затем на основе полученных среднемесячных уровней рассчитывается как их простая средняя среднемесячное производство молока за все три года наблюдения; в нашем примере эта величина равна 1054 ц. Искомые месячные индексы

Таблица 4.6

Производство молока в агрофирме за 2005—2007 гг.

Месяц	Производство молока в агрофирме, ц				Индексы сезонности, %
	2005 г.	2006 г.	2007 г.	В среднем за три года	
I	602	618	634	618	59
II	675	682	709	687	65
III	986	998	1017	1000	95
IV	1192	1214	1231	1212	115
V	1458	1472	1486	1472	140
VI	1507	1529	1557	1531	145
VII	1621	1643	1668	1644	156
VIII	1370	1382	1398	1383	131
IX	943	951	975	956	91
X	739	746	760	748	71
XI	687	692	701	693	66
XII	664	674	692	677	64
Среднемесячный уровень				1054	100

сезонности (в процентах) находятся как отношения средних для каждого месяца к среднемесячному уровню за весь период наблюдения, принимаемому за 100%. В качестве примера рассчитаем индекс сезонности для января. Значение этого индекса равно $618/1054 \cdot 100 = 59\%$. Аналогично рассчитываются индексы сезонности для других месяцев; исчисленные индексы сезонности представлены в последней графе табл. 4.6. Эти индексы характеризуют сезонную волну производства молока в рассматриваемой агрофирме и размах ее колеблемости во внутригодовой динамике. Для наглядного представления полученной сезонной волны производства молока результаты расчета представлены в виде графика на рис. 4.1, который свидетельствует о явно выраженной сезонности в рассматриваемом экономическом процессе.

Описанный выше метод простых средних применим для анализа сезонности во временных рядах без явно выраженного тренда. Если уровень экономического явления проявляет тенденцию к росту или снижению, т.е. имеет место тренд, то отклонения от постоянного среднего уровня могут исказить сезонные колебания. В таких случаях применяется *метод помесячных отношений*. Он заключается в том, что вначале вычисляются по каждому году процентные отношения между показателями за каждый данный и пред-

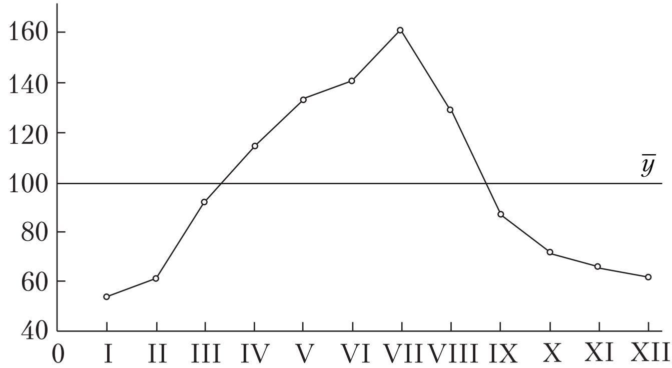


Рис. 4.1

шествующий месяцы, а затем из полученных отношений определяются их средние арифметические. Эти средние показывают динамику изменения показателя в данном месяце по сравнению с предыдущим. Рассмотрим использование метода помесячных отношений на примере анализа сезонности в экономическом процессе, изучаемом в примере 4.1.

Пример 4.2. На основе данных примера 4.1 о производстве молока в специализированной агрофирме за три года требуется проанализировать сезонные колебания производства молока на фирме с использованием метода помесячных отношений. Результаты расчетов представлены в табл. 4.7.

Поясним эти расчеты (как и в предыдущем примере, вычисления делаются с точностью до целых). Все отношения за январь принимаем за 100%. По данным рассматриваемого примера (см. табл. 4.6) для февраля каждого года получим следующие три отношения: в 2005 г. — 675/602·100 = 112%; в 2006 г. — 682/618·100 = 110%; в 2007 г. — 709/634·100 = 112%. Средняя арифметическая этих величин дает среднее помесячное отношение в феврале: (112 +110 + 112)/3 = 111%. Это означает, что в феврале производство молока по сравнению с январем возросло в среднем на 11%. Аналогично рассчитывается помесячное отношение марта к февралю и так по всем месяцам года.

Метод помесячных отношений дает более точное представление о сезонных колебаниях по сравнению с методом простых средних. Следует отметить, что в тех случаях, когда динамический ряд имеет тенденцию к росту или снижению, для исследования сезонности в экономических процессах используются также другие статистические методы,

Таблица 4.7

Определение индексов сезонности методом помесячных отношений

Месяц	Помесячные отношения, %			Средние помесячные отношения, %
	2005 г.	2006 г.	2007 г.	
I	100	100	100	100
II	112	110	112	111
III	146	146	143	145
IV	121	122	121	121
V	122	121	121	121
VI	103	104	105	104
VII	108	107	107	107
VIII	85	84	84	84
IX	69	69	70	69
X	78	78	78	78
XI	93	93	92	92
XII	97	97	99	98

например *метод подвижных (скользящих) средних*, описанные в ряде учебных пособиях по общей теории статистики.

Методы гармонического анализа сезонности

Рассмотрим в заключение параграфа некоторые вопросы применения гармонического анализа при исследовании и моделировании сезонных колебаний в экономике. Если изменение какого-либо показателя носит периодический характер, то такому изменению соответствует периодическая *функция Фурье*. Сезонная волна представляет собой синусоидальную функцию с периодом один год; разложение таких функций в тригонометрический ряд Фурье носит название *гармонического анализа,* и аналитической формой сезонной волны служит тригонометрический многочлен вида

$$\hat{y}_t = a_0 + \sum_{k=1}^{m} (a_k \cdot \cos kt + b_k \cdot \sin kt). \qquad (4.19)$$

В этом многочлене k — порядковый номер гармоники ряда Фурье; m — число гармоник; t — время, принимающее значения 0, $2\pi/n$, $2\cdot 2\pi/n$, ..., $(n-1)\cdot 2\pi/n$ (для месячных данных $n = 12$); параметры a_0, a_k, b_k находятся в соответствии с методом наименьших квадратов и задаются следующими соотношениями:

$$a_0 = \frac{1}{n}\sum y_t; \quad a_k = \frac{2}{n}\sum y_t \cdot \cos kt; \quad b_k = \frac{2}{n}\sum y_t \cdot \sin kt.$$

На практике при выравнивании данных сезонных процессов по ряду Фурье рассчитывают не более четырех гармоник, а затем определяют, при каком числе гармоник наилучшим образом отражается периодичность изменения уровней ряда. Следует иметь в виду, что увеличение числа гармоник, с одной стороны, увеличивает точность аппроксимации, а с другой — может уменьшить значимость модели в результате увеличения дисперсии

$$\sigma^2_{\hat{y}_t} = \frac{\sum(y_t - \hat{y}_t)^2}{n-p},$$

где p — число определяемых параметров аппроксимирующего уравнения (4.19).

Пример 4.3. Покажем процесс выравнивания сезонных колебаний по ряду Фурье на условных месячных данных о численности персонала фирмы, связанной с переработкой сельскохозяйственной продукции. Исходные данные, а также произведения $y_t \cdot \cos t$, $y_t \cdot \sin t$ и $y_t \cdot \sin 2t$, необходимые для определения параметров сглаживающих уравнений по первой и второй гармоникам, приведены в табл. 4.8.

Таблица 4.8

Месяц	t	Численность персонала y_t, чел.	$y_t \cdot \cos t$	$y_t \cdot \sin t$	$y_t \cdot \cos 2t$	$y_t \cdot \sin 2t$
1	0	750	750,00	0	750,00	0
2	π/6	740	640,84	370,00	370,00	640,84
3	π/3	810	405,00	701,46	−405,00	701,46
4	π/2	840	0	840,00	−840,00	0
5	2π/3	990	−495,00	857,34	−495,00	−857,34
6	5π/6	1200	−1039,20	600,00	600,00	−1039,20
7	π	1280	−1280,00	0	1280,00	0
8	7π/6	1240	−1073,84	−620,00	620,00	1073,84
9	4π/3	1150	−575,00	−995,90	−575,00	995,90
10	3π/2	990	0	−990,00	−990,00	0
11	5π/2	880	440,00	−762,08	−440,00	−762,08
12	11π/6	840	727,44	−420,00	420,00	−727,44
		11 710	−1499,76	−419,18	295,00	25,98

Решение. На основе данных этой расчетной таблицы находим:

$$a_0 = \sum y_t / n = 11710 / 12 = 975{,}83 = \overline{y}_t;$$
$$a_1 = 2 \cdot \sum y_t \cdot \cos t / n = -1499{,}76 / 6 = -249{,}96;$$
$$b_1 = 2 \cdot \sum y_t \cdot \sin t / n = -419{,}18 / 6 = -69{,}86;$$
$$a_2 = 2 \cdot \sum y_t \cdot \cos 2t / n = 295{,}00 / 6 = 49{,}17;$$
$$b_2 = 2 \cdot \sum y_t \cdot \sin 2t / n = 25{,}98 / 6 = 4{,}33.$$

Таким образом, аппроксимирующий многочлен Фурье с учетом только первой гармоники имеет вид

$$\hat{y}_{t_1} = 975{,}83 - 249{,}96 \cos t - 69{,}86 \sin t,$$

а с учетом и второй гармоники

$$\hat{y}_{t_2} = 975{,}83 - 249{,}96 \cos t - 69{,}86 \sin t + 49{,}17 \cos 2t + 4{,}33 \sin 2t.$$

Выровненные значения численности персонала ширмы по обоим сглаживающим уравнениям, а также данные для расчета дисперсий аппроксимации приведены в табл. 4.9.

Таблица 4.9

Месяц	y_t	$\hat{y}_{t_1}$	$\hat{y}_{t_2}$	$(y_t - \hat{y}_{t_1})^2$	$(y_t - \hat{y}_{t_2})^2$
1	750	725,83	775,00	584,19	625,00
2	740	724,44	752,77	242,11	163,07
3	810	790,35	769,52	386,12	1638,63
4	840	905,97	856,80	4352,04	282,24
5	990	1040,31	1011,98	2531,10	483,12
6	1200	1157,36	1178,19	1818,17	475,68
7	1280	1225,79	1274,96	2938,72	25,40
8	1240	1227,23	1255,56	163,07	242,11
9	1150	1161,31	1140,48	127,92	90,63
10	990	1045,70	996,53	3102,49	42,64
11	880	911,36	883,03	983,45	9,18
12	840	794,30	815,13	2088,49	618,52
	11 710	11 709,95	11 709,95	19 317,87	4696,22

Из этих расчетов следует, что дисперсии аппроксимации для двух приведенных выше сглаживающих уравнений равны соответственно

$$\sigma^2_{\hat{y}_{t_1}} = \frac{19\,317{,}87}{12-3} = 2146{,}43 \quad \text{и} \quad \sigma^2_{\hat{y}_{t2}} = \frac{4696{,}22}{12-5} = 670{,}89.$$

Соответствующие значения средних квадратических отклонений (стандартных ошибок) аппроксимации равны $\sigma_{\hat{y}_{t_1}} = 46{,}33$ и $\sigma_{\hat{y}_{t2}} = 25{,}90$. Следовательно, лучшим аппроксимирующим тригонометрическим многочленом из двух рассмотренных является многочлен с первой и второй гармониками.

Зная эмпирические и теоретические значения численности персонала, можно определить индексы сезонности для изучаемого показателя, т.е. сезонную волну:

$$J_j = \frac{\hat{y}_{tzj}}{\bar{y}_t} \cdot 100\%.$$

Таким образом, для первого месяца $J_1 = (775{,}00/975{,}83) \cdot 100\% = 79{,}42\%$, для второго $J_2 = (752{,}77/975{,}83) \cdot 100\% = 77{,}14\%$ и т.д. по месяцам года. График сезонной волны исследуемого показателя представлен на рис. 4.2.

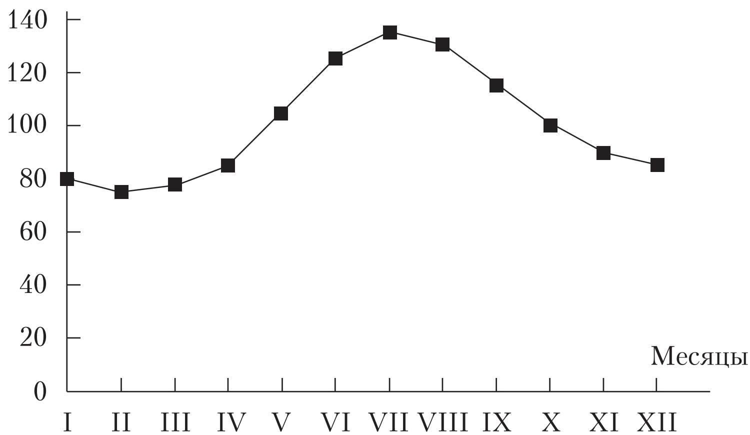


Рис. 4.2. **Сезонная волна численности персонала фирмы**

Вопросы и задания

1. Дайте определение временного экономического ряда и характеристику его структурно образующих элементов.
2. Что такое аномальный уровень временного ряда? Какие методы обнаружения и устранения аномальных уровней вы знаете?
3. Перечислите основные этапы изученных методов определения наличия тренда.
4. Поясните суть методов механического сглаживания временных рядов. Дайте сравнительную характеристику этих методов.
5. Назовите основные показатели экономической динамики, рассчитываемые на основе временных рядов.
6. В чем сущность явления автокорреляции во временных рядах? Что такое временной лаг?
7. Дайте характеристику явления сезонности в экономических процессах. Какие методы выявления и фильтрации сезонной компоненты временного ряда вы знаете?
8. Поясните суть статистических методов анализа сезонности. Что такое сезонная волна?
9. Дайте характеристику методов гармонического анализа сезонности на основе выравнивания по ряду Фурье.

Упражнения

1. В таблице приведены годовые данные о трудоемкости производства 1 т цемента (нормо-смен).

Текущий номер года (t)	1	2	3	4	5	6	7	8	9	10
Трудоемкость 1 т цемента (y_t)	7,9	8,3	7,5	6,9	7,2	6,5	5,8	4,9	5,1	4,4

Определить наличие тренда во временном ряду:

а) методом проверки разности средних уровней;

б) методом Фостера — Стьюарта (табличные значения статистик Стьюдента и Фишера принять равными $t_\alpha = 2{,}23$; $F_\alpha = 3{,}07$; другие необходимые табличные данные приведены в табл. 4.5).

2. Сгладить временной ряд, приведенный в упр. 1, методом простой скользящей средней. Результаты показать на графике.

3. Сгладить временной ряд, приведенный в упр. 1, методом экспоненциального сглаживания, приняв параметр α экспоненциального сглаживания равным: а) 0,1; б) 0,3. Отразить результаты сглаживания на графике. Визуально определить, какое из этих

значений α наиболее соответствует экономическому процессу, заданному временным рядом.

4. Для показателя, заданного временным рядом в упр. 1, определить:

а) средний абсолютный прирост;

б) средние темпы роста и прироста;

в) средний уровень ряда;

г) среднеквадратическое отклонение и дисперсию.

5. Для временного ряда в упражнении 1 рассчитать несколько первых коэффициентов автокорреляции по формуле (4.18') и построить график автокорреляционной функции. Найти величину временного лага.

6. На основе данных о производстве молока в агрофирме в среднем за 3 года (графа 5 табл. 4.6) построить аппроксимирующий многочлен Фурье с первой гармоникой, рассчитать месячные индексы сезонности, построить график сезонной волны и сравнить его с графиком на рис. 4.1.

Глава 5
МОДЕЛИ ПРОГНОЗИРОВАНИЯ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

— Трендовые модели на основе кривых роста
— Оценка адекватности и точности трендовых моделей
— Прогнозирование экономической динамики на основе трендовых моделей
— Адаптивные модели прогнозирования

В процессе усвоения учебного материала главы 5 студенты должны:

***знать*:**

- основные классы моделей прогнозирования на основе временных экономических рядов;
- наиболее распространенные виды трендовых моделей прогнозирования;
- методы оценки качества моделей и прогнозирования экономической динамики на их основе;
- суть адаптивных методов прогнозирования;

***уметь*:**

- выбирать модель, наиболее соответствующую данному временному ряду;
- строить модели временных экономических рядов, оценивать их адекватность и точность;
- находить точечный и интервальный прогнозы рассматриваемого экономического процесса;

***владеть*:**

- методами выбора, построения и оценки качества моделей прогнозирования экономической динамики;
- математико-статистическим инструментарием прогнозирования экономических процессов с использованием трендовых и адаптивных моделей.

5.1. Трендовые модели на основе кривых роста

Основная цель создания трендовых моделей экономической динамики — на их основе сделать прогноз о развитии изучаемого процесса на предстоящий промежуток времени. Прогнозирование на основе временного ряда экономических показателей относится к одномерным методам прогнозирования, базирующимся на экстраполяции, т.е. на продлении на будущее тенденции, наблюдавшейся в прошлом. При таком подходе предполагается, что прогнозируемый показатель формируется под воздействием большого количества факторов, выделить которые либо невозможно, либо по которым отсутствует информация. В этом случае ход изменения данного показателя связывают не с факторами, а с течением времени, что проявляется в образовании одномерных временных рядов.

Рассмотрим метод экстраполяции на основе так называемых кривых роста экономической динамики.

Использование метода экстраполяции на основе кривых роста для прогнозирования базируется на двух предположениях:

— временной ряд экономического показателя действительно имеет тренд, т.е. преобладающую тенденцию;

— общие условия, определявшие развитие показателя в прошлом, останутся без существенных изменений в течение периода упреждения.

В настоящее время насчитывается большое количество типов кривых роста для экономических процессов. Чтобы правильно подобрать наилучшую кривую роста для моделирования и прогнозирования экономического явления, необходимо знать особенности каждого вида кривых. Наиболее часто в экономике используются полиномиальные, экспоненциальные и S-образные кривые роста. Простейшие **полиномиальные** кривые роста имеют вид:

$$\hat{y}_t = a_0 + a_1 t \text{ (полином первой степени)}$$
$$\hat{y}_t = a_0 + a_1 t + a_2 t^2 \text{ (полином второй степени)}$$
$$\hat{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \text{ (полином третьей степени)}$$

и т.д.

Параметр a_1 называют *линейным приростом*, параметр a_2 — *ускорением роста*, параметр a_3 — *изменением ускорения роста*.

Для полинома первой степени характерен постоянный закон роста. Если рассчитать первые приросты по формуле $u_t = y_t - y_{t-1}$, $t = 2, 3, ..., n$, то они будут постоянной величиной и равны a_1.

Если первые приросты рассчитать для полинома второй степени, то они будут иметь линейную зависимость от времени и ряд из первых приростов $u_2, u_3, ...$ на графике будет представлен прямой линией. Вторые приросты $u_t^{(2)} = u_t - u_{t-1}$ для полинома второй степени будут постоянны.

Для полинома третьей степени первые приросты будут полиномами второй степени, вторые приросты будут линейной функцией времени, а третьи приросты, рассчитываемые по формуле $u_t^{(3)} = u_t^{(2)} - u_{t-1}^{(2)}$, будут постоянной величиной.

На основе сказанного можно отметить следующие свойства полиномиальных кривых роста:

— от полинома высокого порядка можно путем расчета последовательных разностей (приростов) перейти к полиному более низкого порядка;

— значения приростов для полиномов любого порядка не зависят от значений самой функции $\hat{y}_t$.

Таким образом, полиномиальные кривые роста можно использовать для аппроксимации (приближения) и прогнозирования экономических процессов, в которых последующее развитие не зависит от достигнутого уровня.

В отличие от использования полиномиальных кривых использование **экспоненциальных** кривых роста предполагает, что дальнейшее развитие зависит от достигнутого уровня, например, прирост зависит от значения функции. В экономике чаще всего применяются две разновидности экспоненциальных (показательных) кривых: простая экспонента и модифицированная экспонента.

Простая экспонента представляется в виде функции

$$\hat{y}_t = ab^t, \tag{5.1}$$

где a и b — положительные числа, при этом если b больше единицы, то функция возрастает с ростом времени t, если b меньше единицы — функция убывает.

Можно заметить, что ордината данной функции изменяется с постоянным темпом прироста. Если взять отношение прироста к самой ординате, оно будет постоянной величиной:

$$\frac{u_t}{y_t} = \frac{y_t - y_{t\text{-}1}}{y_t} = 1 - \frac{1}{b}.$$

Прологарифмируем выражение для данной функции по любому основанию:

$$\log \hat{y}_t = \log a + t \log b.$$

Отсюда можно заметить, что логарифмы ординат простой экспоненты линейно зависят от времени.

Модифицированная экспонента имеет вид

$$\hat{y}_t = k + ab^t, \tag{5.2}$$

где постоянные величины: a меньше нуля, b положительна и меньше единицы, а константа k носит название асимптоты этой функции, т.е. значения функции неограниченно приближаются (снизу) к величине k. Могут быть другие варианты модифицированной экспоненты, но на практике наиболее часто встречается указанная выше функция.

Если прологарифмировать первые приросты данной функции, то получится функция, линейно зависящая от времени, а если взять отношение двух последовательных приростов, то оно будет постоянной величиной:

$$\frac{u_t}{u_{t-1}} = \frac{y_t - y_{t-1}}{y_{t-1} - y_{t-2}} = b.$$

В экономике достаточно распространены процессы, которые сначала растут медленно, затем ускоряются, а затем снова замедляют свой рост, стремясь к какому-либо пределу. В качестве примера можно привести процесс ввода некоторого объекта в промышленную эксплуатацию, процесс изменения спроса на товары, обладающие способностью достигать некоторого уровня насыщения, и др. Для моделирования таких процессов используются так называемые **S-образные** кривые роста, среди которых выделяют кривую Гомперца и логистическую кривую.

Кривая Гомперца имеет аналитическое выражение

$$\hat{y}_t = ka^{b^t}, \tag{5.3}$$

где a, b — положительные параметры, причем b меньше единицы; параметр k — асимптота функции.

В кривой Гомперца выделяются четыре участка: на первом — прирост функции незначителен, на втором — прирост увеличивается, на третьем участке прирост примерно постоянен, на четвертом — происходит замедление темпов прироста, и функция неограниченно приближается к значению k. В результате конфигурация кривой напоминает латинскую букву S.

Логарифм данной функции является экспоненциальной кривой; логарифм отношения первого прироста к самой ординате функции — линейная функция времени.

На основании кривой Гомперца описывается, например, динамика показателей уровня жизни; модификации этой кривой используются в демографии для моделирования показателей смертности и т.д.

Логистическая кривая, или кривая Перла — Рида, — возрастающая функция, наиболее часто выражаемая в виде

$$\hat{y}_t = \frac{k}{1 + ae^{-bt}} \qquad (5.4)$$

другие виды этой кривой:

$$\hat{y}_t = \frac{k}{1 + ab^{-t}}; \; \hat{y}_t = \frac{k}{1 + 10^{a-bt}}.$$

В этих выражениях a и b — положительные параметры; k — предельное значение функции при бесконечном возрастании времени.

Если взять производную данной функции, то можно увидеть, что скорость возрастания логистической кривой в каждый момент времени пропорциональна достигнутому уровню функции и разности между предельным значением k и достигнутым уровнем. Логарифм отношения первого прироста функции к квадрату ее значения (ординаты) есть линейная функция от времени.

Конфигурация графика логистической кривой близка графику кривой Гомперца, но в отличие от последней логистическая кривая имеет точку симметрии, совпадающую с точкой перегиба.

Рассмотрим проблему предварительного выбора вида кривой роста для конкретного временного ряда. Допустим, имеется временной ряд $y_1, y_2, y_3, ..., y_n$. Для выбора вида полиномиальной кривой роста наиболее распространенным методом является **метод конечных разностей (метод Тинтнера)**. Этот метод может быть использован для предварительного выбора полиномиальной кривой, если, во-первых, уровни временного ряда состоят только из двух компонент: тренд и случайная компонента, и во-вторых, тренд является достаточно гладким, чтобы его можно было аппроксимировать полиномом некоторой степени.

На первом этапе этого метода вычисляются разности (приросты) до k-го порядка включительно:

$$u_t = y_t - y_{t-1};$$
$$u_t^{(2)} = u_t - u_{t-1};$$
$$\ldots$$
$$u_t^{(k)} = u_t^{(k-1)} - u_{t-1}^{(k-1)}.$$

Для аппроксимации экономических процессов обычно вычисляют конечные разности до четвертого порядка.

Затем для исходного ряда и для каждого разностного ряда вычисляются дисперсии по следующим формулам:
для исходного ряда

$$\sigma_0^2 = \frac{\sum_{t=1}^{n} y_t^2 - \frac{1}{n}\left(\sum_{t=1}^{n} y_t\right)^2}{n-1};$$

для разностного ряда k-го порядка (k = 1,2, ...)

$$\sigma_k^2 = \frac{\sum_{t=k+1}^{n} \left(u_t^k\right)^2}{(n-k)C_{2k}^k}; \qquad C_{2k}^k \text{ — биноминальный коэффициент.}$$

Производится сравнение отклонений каждой последующей дисперсии от предыдущей, т.е. вычисляются величины

$$\left|\sigma_k^2 - \sigma_{k-1}^2\right|,$$

и если для какого-либо k эта величина не превосходит некоторой наперед заданной положительной величины, т.е. дисперсии одного порядка, то степень аппроксимирующего полинома должна быть равна $k - 1$.

Более универсальным методом предварительного выбора кривых роста, позволяющим выбрать кривую из широкого класса кривых роста, является **метод характеристик приростов**. Он основан на использовании отдельных характерных свойств кривых, рассмотренных выше. При этом методе исходный временной ряд предварительно сглаживается методом простой скользящей средней. Например, для интервала сглаживания m = 3 сглаженные уровни рассчитываются по формуле

$$\overline{y}_t = \frac{y_{t-1} + y_t + y_{t+1}}{3},$$

причем, чтобы не потерять первый и последний уровни, их сглаживают по формулам

$$\overline{y}_1 = \frac{5y_1 + 2y_2 - y_3}{6},$$

$$\overline{y}_n = \frac{-y_{n-2} + 2y_{n-1} + 5y_n}{6}.$$

Затем вычисляются первые средние приросты

$$\overline{u}_t = \frac{\overline{y}_{t+1} - \overline{y}_{t-1}}{2}, \; t = 2, 3, ..., \lambda, n-1;$$

вторые средние приросты

$$\overline{u}_t^{(2)} = \frac{\overline{u}_{t+1} - \overline{u}_{t-1}}{2},$$

а также ряд производных величин, связанных с вычисленными средними приростами и сглаженными уровнями ряда:

$$\frac{\overline{u}_t}{\overline{y}_t}; \quad \log \overline{u}_t; \quad \log \frac{\overline{u}_t}{\overline{y}_t}; \quad \log \frac{\overline{u}_t}{\overline{y}_t^2}.$$

В соответствии с характером изменения средних приростов и производных показателей выбирается вид кривой роста для исходного временного ряда, при этом используется табл. 5.1.

Таблица 5.1

Показатель	Характер изменения показателя во времени	Вид кривой роста
Первый средний прирост $\overline{u}_t$	Примерно одинаковы	Полином первого порядка (прямая)
То же	Изменяются линейно	Полином второго порядка (парабола)
Второй средний прирост $\overline{u}_t^{(2)}$	Изменяются линейно	Полином третьего порядка (кубическая парабола)
$\frac{\overline{u}_t}{\overline{y}_t}$	Примерно одинаковы	Простая экспонента
$\log \overline{u}_t$	Изменяются линейно	Модифицированная экспонента
$\log \frac{\overline{u}_t}{\overline{y}_t}$	Изменяются линейно	Кривая Гомперца
$\log \frac{\overline{u}_t}{\overline{y}_t^2}$	Изменяются линейно	Логистическая кривая

На практике при предварительном выборе отбирают обычно две-три кривые роста для дальнейшего исследования и построения трендовой модели данного временного ряда.

Рассмотрим методы определения параметров отобранных кривых роста. Параметры полиномиальных кривых оцениваются, как правило, **методом наименьших квадратов**, суть которого заключается в том, чтобы сумма квадратов отклонений фактических уровней ряда от соответствующих выравненных по кривой роста значений была наименьшей. Этот метод приводит к системе так называемых *нормальных уравнений* для определения неизвестных параметров отобранных кривых.

Для полинома первой степени

$$\hat{y}_t = a_0 + a_1 t$$

система нормальных уравнений имеет вид

$$\begin{cases} a_0 n + a_1 \sum t = \sum y_t \\ a_0 \sum t + a_1 \sum t^2 = \sum y_t t \end{cases}, \tag{5.5}$$

где знак суммирования распространяется на все моменты наблюдения (все уровни) исходного временного ряда.

Аналогичная система для полинома второй степени

$$\hat{y}_t = a_0 + a_1 t + a_2 t^2$$

имеет вид

$$\begin{cases} a_0 n + a_1 \sum t + a_2 \sum t^2 = \sum y_t, \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 = \sum y_t t, \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 = \sum y_t t^2. \end{cases} \tag{5.6}$$

Для полинома третьей степени

$$\hat{y}_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3$$

система нормальных уравнений записывается следующим образом:

$$\begin{cases} a_0 n + a_1 \sum t + a_2 \sum t^2 + a_3 \sum t^3 = \sum y_t, \\ a_0 \sum t + a_1 \sum t^2 + a_2 \sum t^3 + a_3 \sum t^4 = \sum y_t t, \\ a_0 \sum t^2 + a_1 \sum t^3 + a_2 \sum t^4 + a_3 \sum t^5 = \sum y_t t^2. \\ a_0 \sum t^3 + a_1 \sum t^4 + a_2 \sum t^5 + a_3 \sum t^6 = \sum y_t t^3. \end{cases} \tag{5.7}$$

Параметры экспоненциальных и S-образных кривых находятся более сложными методами. Для простой экспоненты

$$\hat{y}_t = ab^t$$

предварительно логарифмируют выражение по некоторому основанию (например, десятичному или натуральному):

$$\log \hat{y}_t = \log a + t \log b,$$

т.е. для логарифма функции получают линейное выражение, а затем для неизвестных параметров $\log a$ и $\log b$ составляют на основе метода наименьших квадратов систему нормальных уравнений, аналогичную системе для полинома первой степени. Решая эту систему, находят логарифмы параметров, а затем и сами параметры модели.

При определении параметров кривых роста, имеющих асимптоты (модифицированная экспонента, кривая Гомперца, логистическая кривая), различают два случая. Если значение асимптоты k известно заранее, то путем несложной модификации формулы и последующего логарифмирования определение параметров сводят к решению системы нормальных уравнений, неизвестными которой являются логарифмы параметров кривой.

Если значение асимптоты заранее неизвестно, то для нахождения параметров указанных выше кривых роста используются приближенные методы: метод трех точек, метод трех сумм и др.

Таким образом, при моделировании экономической динамики, заданной временным рядом, путем сглаживания исходного ряда, определения наличия тренда, отбора одной или нескольких кривых роста и определения их параметров в случае наличия тренда получают одну или несколько трендовых моделей для исходного временного ряда. Встает вопрос, насколько эти модели близки к экономической реальности, отраженной во временном ряду, насколько обосновано применение этих моделей для анализа и прогнозирования изучаемого экономического явления. Этот вопрос рассматривается в следующем параграфе.

5.2. Оценка адекватности и точности трендовых моделей

Независимо от вида и способа построения экономико-математической модели вопрос о возможности ее применения в целях анализа и прогнозирования экономического явления может быть решен только после установления *адекватности*, т.е. соответствия модели исследуемому процессу или объекту. Так как полного соответствия модели реальному процессу или объекту быть не может, адекватность — в какой-то мере условное понятие. При моделировании имеется в виду адекватность не вообще, а по тем свойствам модели, которые считаются существенными для исследования.

Трендовая модель $\hat{y}_t$, конкретного временного ряда y_t, считается адекватной, если правильно отражает систематические компоненты временного ряда. Это требование эквивалентно требованию, чтобы остаточная компонента $\varepsilon_t = y_t - \hat{y}_t$ $(t = 1, 2, ..., n)$ удовлетворяла свойствам случайной компоненты временного ряда, указанным в параграфе 4.1: случайность колебаний уровней остаточной последовательности, соответствие распределения случайной компоненты нормальному закону распределения, равенство математического ожидания случайной компоненты нулю, независимость значений уровней случайной компоненты. Рассмотрим, каким образом осуществляется проверка этих свойств остаточной последовательности.

Проверка случайности колебаний уровней остаточной последовательности означает проверку гипотезы о правильности выбора вида тренда. Для исследования случайности отклонений от тренда мы располагаем набором разностей

$$\varepsilon_t = y_t - \hat{y}_t \; (t = 1, 2, ..., n).$$

Характер этих отклонений изучается с помощью ряда непараметрических критериев. Одним из таких критериев является *критерий серий*, основанный на медиане выборки. Ряд из величин ε_t располагают в порядке возрастания их значений и находят медиану ε_m полученного вариационного ряда, т.е. срединное значение при нечетном n, или среднюю арифметическую из двух срединных значений, при n четном. Возвращаясь к исходной последовательности ε_t и сравнивая значения этой последовательности с ε_m, будем ставить знак «плюс», если значение ε_t превосходит медиану, и знак «минус», если оно меньше медианы; в случае равенства сравниваемых величин соответствующее значение ε_t опускается. Таким образом, получается последовательность, состоящая из плюсов и минусов, общее число которых не превосходит n. Последовательность подряд идущих плюсов или минусов называется *серией*. Для того чтобы последовательность ε_t была случайной выборкой, протяженность самой длинной серии не должна быть слишком большой, а общее число серий — слишком малым.

Обозначим протяженность самой длинной серии через $K_{\max}$, а общее число серий — через υ. Выборка признается случайной, если выполняются следующие неравенства для 5%-ного уровня значимости:

$$K_{\max} < \left[3{,}3\left(\lg n + 1\right)\right];$$

$$\upsilon > \left[\frac{1}{2}\left(n + 1 - 1{,}96\sqrt{n-1}\right)\right], \qquad (5.8)$$

где квадратные скобки означают целую часть числа.

Если хотя бы одно из этих неравенств нарушается, то гипотеза о случайном характере отклонений уровней временного ряда от тренда отвергается и, следовательно, трендовая модель признается неадекватной.

Другим критерием для данной проверки может служить *критерий пиков (поворотных точек)*. Уровень последовательности ε_t считается максимумом, если он больше двух рядом стоящих уровней, т.е. $\varepsilon_{t-1} < \varepsilon_t > \varepsilon_{t+1}$, и минимумом, если он меньше обоих соседних уровней, т.е. $\varepsilon_{t-1} > \varepsilon_t < \varepsilon_{t+1}$. В обоих случаях ε_t считается поворотной точкой; общее число поворотных точек для остаточной последовательности ε_t обозначим через p. В случайной выборке математическое ожидание числа точек поворота $\overline{p}$ и дисперсия σ_p^2 выражаются формулами:

$$\overline{p} = \frac{2}{3}\left(n - 2\right); \quad \sigma_p^2 = \frac{16n - 29}{90}.$$

Критерием случайности с 5%-ным уровнем значимости, т.е. с доверительной вероятностью 95%, является выполнение неравенства

$$p > \left[\overline{p} - 1{,}96\sqrt{\sigma_p^2}\right], \qquad (5.9)$$

где квадратные скобки означают целую часть числа. Если это неравенство не выполняется, трендовая модель считается неадекватной.

Проверка соответствия распределения остаточной последовательности нормальному закону распределения может быть произведена лишь приближенно с помощью исследования показателей *асимметрии* (γ_1) и эксцесса (γ_2), так как временные ряды, как правило, не очень велики. При нормальном распределении показатели асимметрии и эксцесса некоторой генеральной совокупности равны нулю. Мы предполагаем, что отклонения от тренда представляют собой выборку из генеральной совокупности, поэтому можно определить только выборочные характеристики асимметрии и эксцесса и их ошибки:

$$\hat{\gamma}_1 = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^3}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^3}}; \quad \sigma_{\hat{\gamma}_1} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}};$$
$$\hat{\gamma}_2 = \frac{\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^4}{\left(\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^2\right)^2}; \quad \sigma_{\hat{\gamma}_2} = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}. \tag{5.10}$$

В этих формулах $\hat{\gamma}_1$ — выборочная характеристика асимметрии; $\hat{\gamma}_2$ — выборочная характеристика эксцесса; $\sigma_{\hat{\gamma}_1}$ и $\sigma_{\hat{\gamma}_2}$ — соответствующие среднеквадратические ошибки.

Если одновременно выполняются следующие неравенства:

$$\left|\hat{\gamma}_1\right| < 1{,}5\sigma_{\hat{\gamma}_1}; \quad \left|\hat{\gamma}_2 + \frac{6}{n+1}\right| < 1{,}5\sigma_{\hat{\gamma}_2},$$

то гипотеза о нормальном характере распределения остаточной последовательности принимается.

Если выполняется хотя бы одно из неравенств

$$\left|\hat{\gamma}_1\right| \geq 2\sigma_{\hat{\gamma}_1}; \quad \left|\hat{\gamma}_2 + \frac{6}{n+1}\right| \geq 2\sigma_{\hat{\gamma}_2},$$

то гипотеза о нормальном характере распределения отвергается, трендовая модель признается неадекватной. Другие случаи требуют дополнительной проверки с помощью более сложных критериев.

Кроме рассмотренного метода известен ряд других методов проверки нормальности закона распределения случайной величины: метод Вестергарда, *RS*-критерий и т.д. Рассмотрим наиболее простой из них, основанный на *RS*-критерии. Этот критерий численно равен отношению размаха вариации случайной величины *R* к стандартному отклонению *S*.

В нашем случае $R = \varepsilon_{\max} - \varepsilon_{\min}$, а $S = S_{\hat{y}} = \sqrt{\sum \varepsilon_t^2 / n-1}$. Вычисленное значение *RS*-критерия сравнивается с табличными (критическими) нижней и верхней границами данного отношения, и если это значение не попадает в интервал между критическими границами, то с заданным уровнем значимости гипотеза о нормальности распределения отвергается; в противном случае эта гипотеза принимается. Для иллюстрации приведем несколько пар значений критических границ RS-критерия для уровня значимости $\alpha = 0{,}05$: при $n = 10$ нижняя граница равна 2,67, а верхняя равна 3,685; при $n = 20$ эти числа составляют соответственно 3,18 и 4,49; при $n = 30$ они равны 3,47 и 4,89.

Проверка равенства математического ожидания остаточной последовательности нулю, если она распределена по нормальному закону, осуществляется на основе t-критерия Стьюдента. Расчетное значение этого критерия задается формулой

$$t = \frac{\bar{\varepsilon} - 0}{S_\varepsilon}\sqrt{n}, \qquad (5.11)$$

где $\bar{\varepsilon}$ — среднее арифметическое значение уровней остаточной последовательности ε_t;
S_ε — стандартное (среднеквадратическое) отклонение для этой последовательности.

Если расчетное значение t меньше табличного значения t_α статистики Стьюдента с заданным уровнем значимости α и числом степеней свободы $n - 1$, то гипотеза о равенстве нулю математического ожидания остаточной последовательности принимается; в противном случае эта гипотеза отвергается и модель считается неадекватной.

Следует отметить, что проверку данного свойства на основе t-критерия имеет смысл проводить лишь в том случае, когда это свойство не очевидно из самого значения величины $\bar{\varepsilon}$.

Проверка независимости значений уровней остаточной последовательности, т.е. проверка отсутствия существенной автокорреляции в остаточной последовательности может осуществляться по ряду критериев, наиболее распространенным из которых является d-критерий Дарбина — Уотсона. Расчетное значение этого критерия определяется по формуле

$$d = \frac{\sum_{t=2}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n}\varepsilon_t^2}. \qquad (5.12)$$

Заметим, что расчетное значение критерия Дарбина — Уотсона в интервале от 2 до 4 свидетельствует об отрицательной связи; в этом случае его надо преобразовать по формуле $d' = 4 - d$ и в дальнейшем использовать значение d'.

Расчетное значение критерия d (или d') сравнивается с верхним d_2 и нижним d_1 критическими значениями статистики Дарбина — Уотсона, фрагмент табличных значений которых для различного числа уровней ряда n и числа определяемых параметров модели k представлен для наглядности в табл. 5.2 (уровень значимости 5%).

Таблица 5.2

n	$k = 1$		$k = 2$		$k = 3$	
	d_1	d_2	d_1	d_2	d_1	d_2
15	1,08	1,36	0,95	1,54	0,82	1,75
20	1,20	1,41	1,10	1,54	1,00	1,68
30	1,35	1,49	1,28	1,57	1,21	1,65

Если расчетное значение критерия d больше верхнего табличного значения d_2, то гипотеза о независимости уровней остаточной последовательности, т.е. об отсутствии в ней автокорреляции, принимается. Если значение d меньше нижнего табличного значения d_1, то эта гипотеза отвергается и модель неадекватна. Если значение d находится между значениями d_1 и d_2, включая сами эти значения, то считается, что нет достаточных оснований сделать тот или иной вывод и необходимы дальнейшие исследования, например, по большему числу наблюдений. В этом случае можно использовать также *первый коэффициент автокорреляции:*

$$r_1 = \frac{\sum_{t=2}^{n} \varepsilon_t \cdot \varepsilon_{t-1}}{\sum_{t=1}^{n} \varepsilon_t^2}.$$

Если расчетное значение этого коэффициента по модулю меньше табличного (критического) значения $r_{1\text{кр}}$, то гипотеза об отсутствии автокорреляции принимается; в противном случае эта гипотеза отвергается. В частности, при $n = 10$ можно принять $r_{1\text{кр}} = 0{,}36$.

Вывод об адекватности трендовой модели делается, если *все* указанные выше четыре проверки свойств остаточной

последовательности дают положительный результат. Для адекватных моделей имеет смысл ставить задачу оценки их *точности*. Точность модели характеризуется величиной отклонения выхода модели от реального значения моделируемой переменной (экономического показателя). Для показателя, представленного временным рядом, точность определяется как разность между значением фактического уровня временного ряда и его оценкой, полученной расчетным путем с использованием модели, при этом в качестве статистических показателей точности применяются следующие:
среднее квадратическое отклонение

$$\sigma_\varepsilon = \sqrt{\frac{1}{n-k}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}, \tag{5.13}$$

средняя относительная ошибка аппроксимации

$$\overline{\varepsilon}_{\text{отн}} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \cdot 100\%, \tag{5.14}$$

коэффициент сходимости

$$\varphi^2 = \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \overline{y})^2}, \tag{5.15}$$

коэффициент детерминации

$$R^2 = 1 - \varphi^2 \tag{5.16}$$

и другие показатели; в приведенных формулах n — количество уровней ряда, k — число определяемых параметров модели, $\hat{y}_t$ — оценка уровней ряда по модели, $\overline{y}$ — среднее арифметическое значение уровней ряда.

На основании указанных показателей можно сделать выбор из нескольких адекватных трендовых моделей экономической динамики наиболее точной, хотя может встретиться случай, когда по некоторому показателю более точна одна модель, а по другому — другая.

Данные показатели точности моделей рассчитываются на основе всех уровней временного ряда и поэтому отражают лишь точность аппроксимации. Для оценки прогнозных свойств модели целесообразно использовать так называемый ретроспективный прогноз — подход, основанный на выделении участка из ряда последних уровней исходного временного ряда в количестве, допустим, n_2 уровней в качестве проверочного, а саму трендовую модель в этом случае следует строить по первым точкам, количество которых будет равно $n_1 = n - n_2$. Тогда для расчета показателей точности модели по ретроспективному прогнозу применяются те же формулы, но суммирование в них будет вестись не по всем наблюдениям, а лишь по последним n_2 наблюдениям. Например, формула для среднего квадратического отклонения будет иметь вид

$$\sigma_\varepsilon = \sqrt{\frac{1}{n_2 - k} \sum_{t=n_1+1} (y_t - \hat{y}_t)^2},$$

где $\hat{y}_t$ — значения уровней ряда по модели, построенной для первых n_1 уровней.

Оценивание прогнозных свойств модели на ретроспективном участке весьма полезно, особенно при сопоставлении различных моделей прогнозирования из числа адекватных. Однако надо помнить, что оценки ретропрогноза — лишь приближенная мера точности прогноза и модели в целом, так как прогноз на период упреждения делается по модели, построенной по всем уровням ряда.

Пример 5.1. Для временного ряда, представленного в первых двух графах табл. 5.3, построена трендовая модель в виде полинома первой степени (линейная модель):

$$\hat{y}_t = 87{,}8 - 3{,}4t.$$

Требуется оценить адекватность и точность построенной модели.

Решение. Прежде всего, сформируем остаточную последовательность (ряд остатков), для чего из фактических значений уровней ряда вычтем соответствующие расчетные значения по модели: остаточная последовательность приведена в графе 4 табл. 5.3.

Проверку случайности уровней ряда остатков проведем на основе критерия пиков (поворотных точек). Точки пиков отмечены в графе 5 табл. 5.3; их количество равно шести ($p = 6$). Правая часть неравенства (5.9) равняется в данном случае двум, т.е. это неравенство выполняется. Следовательно, можно сделать вывод, что свойство случайности ряда остатков подтверждается.

Результаты предыдущей проверки дают возможность провести проверку соответствия остаточной последовательности нормальному закону распределения. Воспользуемся RS-критерием. В нашем случае размах вариации $R = \varepsilon_{max} - \varepsilon_{min} = 2{,}7 - (-2{,}1) = 4{,}8$, а среднее квадратическое отклонение

$$S_{\hat{y}} = \sqrt{\sum_{t=1}^{n} \varepsilon_t^2 / (n-1)} = \sqrt{16{,}51 : 8} = 1{,}44.$$

Таблица 5.3

t	Фактическое y_t	Расчетное $\hat{y}_t$	Отклонение ε_t	Точки пиков	ε_t^2	$\varepsilon^t - \varepsilon_{t-1}$	$\left(\varepsilon^t - \varepsilon_{t-1}\right)^2$	$\left\|\varepsilon_t\right\| : y_t \times$ $\times 100$
1	2	3	4	5	6	7	8	9
1	85	84,4	0,6	–	0,36	–	–	0,71
2	81	81,0	0,0	1	0,00	–0,6	0,36	0,00
3	78	77,6	0,4	1	0,16	0,4	0,16	0,49
4	72	74,1	–2,1	1	4,41	–2,5	6,25	2,69
5	69	70,7	–1,7	0	2,89	0,4	0,16	2,46
6	70	67,3	2,7	1	7,29	4,4	19,36	3,86
7	64	63,8	0,2	1	0,04	–2,5	6,25	0,31
8	61	60,4	0,6	1	0,36	0,4	0,16	0,98
9	56	57,0	–1,0	–	1,00	–1,6	2,56	1,79
45	636	636,3	–0,3	6	16,51		35,26	13,29

Следовательно, критерий RS = 4,8 : 1,44 = 3,33, и это значение попадает в интервал между нижней и верхней границами табличных значений данного критерия (эти границы для n = 10 и уровня значимости α = 0,05 составляют соответственно 2,7 и 3,7). Это позволяет сделать вывод, что свойство нормальности распределения выполняется.

Переходя к проверке равенства (близости) нулю математического ожидания ряда остатков, заметим, что по результатам вычислений в табл. 5.3 это математическое ожидание равно (–0,3) : 9 ≈ – 0,03 и, следовательно, можно подтвердить выполнение данного свойства, не прибегая к статистике Стьюдента.

Для проверки независимости уровней ряда остатков (отсутствия автокорреляции) вычислим значение критерия Дарбина — Уотсона.

Расчеты по формуле (5.12), представленные в графах 6, 7, 8 табл. 5.3, дают следующее значение этого критерия: d = 35,26 : : 15,51 = 2,27. Эта величина превышает 2, что свидетельствует об отрицательной автокорреляции (при наличии последней), поэтому критерий Дарбина — Уотсона необходимо преобразовать: d' = 4 – d = 4 – 2,27 = 1,73. Данное значение сравниваем с двумя критическими табличными значениями критерия, которые для линейной модели в нашем случае можно принять равными d_1 = = 1,08 и d_2 = 1,36. Так как расчетное значение попадает в интервал от d_2 до 2, то делается вывод о независимости уровней остаточной последовательности.

Из сказанного выше следует, что остаточная последовательность удовлетворяет всем свойствам случайной компоненты временного ряда, следовательно, построенная линейная модель является адекватной.

Для характеристики точности модели воспользуемся показателем средней относительной ошибки аппроксимации, который рассчитывается по формуле (5.14): $\bar{\varepsilon}_{\text{отн}} = 13{,}29 : 9 = 1{,}48$ (%). Полученное значение средней относительной ошибки говорит о достаточно высоком уровне точности построенной модели (ошибка менее 5% свидетельствует об удовлетворительном уровне точности; ошибка в 10 и более процентов считается очень большой).

5.3. Прогнозирование экономической динамики на основе трендовых моделей

Прогнозирование экономических показателей на основе трендовых моделей, как и большинство других методов экономического прогнозирования, основано на идее экстраполяции. Как уже сказано выше, под экстраполяцией обычно понимают распространение закономерностей, связей и соотношений, действующих в изучаемом периоде, за его пределы. В более широком смысле слова ее рассматривают как получение представлений о будущем на основе информации, относящейся к прошлому и настоящему. В процессе построения прогнозных моделей в их структуру иногда закладываются элементы будущего предполагаемого состояния объекта или явления, но в целом эти модели отражают закономерности, наблюдаемые в прошлом и настоящем, поэтому достоверный прогноз возможен лишь относительно таких объектов и явлений, которые в значительной степени детерминируются прошлым и настоящим.

Существуют две основные формы детерминации: внутренняя и внешняя. Внутренняя детерминация, или самодетерминация, более устойчива, ее проще идентифицировать с использованием экономико-математических моделей. Внешняя детерминация определяется большим числом факторов, поэтому учесть их все практически невозможно. Если некоторые методы моделирования, например адаптивные, отражают общее совокупное влияние на экономическую систему внешних факторов, т.е. отражают внешнюю детерминацию, то методы, базирующиеся на использовании трендовых моделей экономических процессов, представленных одномерными временными рядами, отражают внутреннюю детерминацию объектов и явлений.

При экстраполяционном прогнозировании экономической динамики на основе временных рядов с использованием трендовых моделей выполняются следующие основные этапы:

1) предварительный анализ данных;
2) формирование набора моделей (например, набора кривых роста), называемых функциями-кандидатами;
3) численное оценивание параметров моделей;
4) определение адекватности моделей;
5) оценка точности адекватных моделей;
6) выбор лучшей модели;
7) получение точечного и интервального прогнозов;
8) верификация прогноза.

Порядок реализации первых шести этапов из перечисленных описан в предыдущих параграфах данной главы. Рассмотрим более подробно два заключительных этапа.

Прогноз на основании трендовых моделей (кривых роста) содержит два элемента: точечный и интервальный прогнозы. *Точечный прогноз* — это прогноз, которым называется единственное значение прогнозируемого показателя. Это значение определяется подстановкой в уравнение выбранной кривой роста величины времени t, соответствующей периоду упреждения: $t = n + 1$; $t = n + 2$ и т.д. Такой прогноз называется точечным, так как на графике его можно изобразить в виде точки.

Очевидно, что точное совпадение фактических данных в будущем и прогностических точечных оценок маловероятно. Поэтому точечный прогноз должен сопровождаться двусторонними границами, т.е. указанием интервала значений, в котором с достаточной долей уверенности можно ожидать появления прогнозируемой величины. Установление такого интервала называется *интервальным прогнозом*.

Интервальный прогноз на базе трендовых моделей осуществляется путем расчета *доверительного интервала* — такого интервала, в котором с определенной вероятностью можно ожидать появления фактического значения прогнозируемого экономического показателя. Расчет доверительных интервалов при прогнозировании с использованием кривых роста опирается на выводы и формулы теории регрессий. Перенесение выводов теории регрессий на временные экономические ряды не совсем правомерно, так как динамические ряды, как выше уже отмечали, отличаются от статистических совокупностей. Поэтому к оцениванию доверительных интервалов для кривых роста следует подходить с известной долей осторожности.

Методы, разработанные для статистических совокупностей, позволяют определить доверительный интервал, зависящий от стандартной ошибки оценки прогнозируемого показателя, от времени упреждения прогноза, от количества уровней во временном ряду и от уровня значимости (ошибки) прогноза.

Стандартная (средняя квадратическая) ошибка оценки прогнозируемого показателя $S_{\hat{y}}$ определяется по формуле

$$S_{\hat{y}} = \sqrt{\frac{\sum\left(y_t - \hat{y}_t\right)^2}{n-k}}, \tag{5.17}$$

где:

y_t — фактическое значение уровня временного ряда для времени t;

$\hat{y}_t$ — расчетная оценка соответствующего показателя по модели (например, по уравнению кривой роста);

n — количество уровней в исходном ряду;

k — число параметров модели.

В случае прямолинейного тренда для расчета доверительного интервала можно использовать аналогичную формулу для парной регрессии, таким образом, доверительный интервал прогноза U_y в этом случае будет иметь вид

$$U_y = \hat{y}_{n+L} \pm t_\alpha S_{\hat{y}} \sqrt{1 + \frac{1}{n} + \frac{3(n+2L-1)^2}{n(n^2-1)}}, \tag{5.18}$$

где:

L — период упреждения;

$\hat{y}_{n+L}$ — точечный прогноз по модели на $(n + L)$-й момент времени;

n — количество наблюдений во временном ряду;

$S_{\hat{y}}$ — стандартная ошибка оценки прогнозируемого показателя, рассчитанная по ранее приведенной формуле для числа параметров модели, равного двум;

t_α — табличное значение критерия Стьюдента для уровня значимости α и для числа степеней свободы, равного $n - 2$.

Если выражение

$$t_\alpha \cdot \sqrt{1+\frac{1}{n}+\frac{3(n+2L-1)^2}{n(n^2-1)}}$$

обозначить через K, то формула для доверительного интервала примет вид

$$U_y = \hat{y}_{n+L} \pm S_{\hat{y}} K\,. \qquad (5.19)$$

Значения величины K для оценки доверительных интервалов прогноза относительно линейного тренда табулированы. Фрагмент такой таблицы для уровня значимости $\alpha = 0{,}20$ представлен для иллюстрации в табл. 5.4.

Таблица 5.4

Число уровней в ряду (n)	Период упреждения L					
	1	2	3	4	5	6
7	1,932	2,106	2,300	2,510	2,733	2,965
10	1,692	1,774	1,865	1,964	2,069	2,180
13	1,581	1,629	1,682	1,738	1,799	1,863
15	1,536	1,572	1,611	1,653	1,697	1,745

Иногда для расчета доверительных интервалов прогноза относительно линейного тренда применяют приведенную выше формулу в несколько преобразованном виде:

$$U_y = \hat{y}_{n+L} \pm t_\alpha S_{\hat{y}} \sqrt{1+\frac{1}{n}+\frac{(t_L-\bar{t})^2}{\sum(t-\bar{t})^2}}. \qquad (5.20)$$

Здесь t — порядковый номер уровня ряда ($t = 1, 2, ..., n$); $t_L = n + L$ — время, для которого делается прогноз; $\bar{t}$ — время, соответствующее середине периода наблюдений для исходного ряда, например $\bar{t} = (n + 1) : 2$; суммирование ведется по всем наблюдениям.

Эту формулу можно упростить, если, как часто делается на практике, перенести начало отсчета времени на середину периода наблюдений ($\bar{t} = 0$):

$$U_y = \hat{y}_{n+L} \pm t_\alpha S_{\hat{y}} \sqrt{1+\frac{1}{n}+\frac{t_L^2}{\sum t^2}}\,. \qquad (5.21)$$

Формула для расчета доверительных интервалов прогноза относительно тренда, имеющего вид полинома второго или третьего порядка, выглядит следующим образом:

$$U_y = \hat{y}_{n+L} \pm t_\alpha S_{\hat{y}} \sqrt{1+\frac{1}{n}+\frac{t_L^2}{\sum t^2}+\frac{\sum t^4 - 2t_L^2\sum t^2 + nt_L^4}{n\sum t^4 - \left(\sum t^2\right)^2}}. \qquad (5.22)$$

Аналогично вычисляются доверительные интервалы для экспоненциальной кривой роста, а также для кривых роста, имеющих асимптоту (модифицированная экспонента, кривая Гомперца, логистическая кривая), если значение асимптоты известно.

Таким образом, формулы расчета доверительного интервала для трендовых моделей разного класса различны, но каждая из них отражает динамический аспект прогнозирования, т.е. увеличение неопределенности прогнозируемого процесса с ростом периода упреждения проявляется в постоянном расширении доверительного интервала.

Несмотря на громоздкость некоторых формул, расчет точечных и интервальных прогнозов на основе трендовых моделей в форме кривых роста технически является достаточно простой процедурой. Однако не следует обольщаться технической простотой процедуры экстраполяции и пытаться заглянуть слишком далеко, это неизбежно приведет к грубым ошибкам. Оптимальная длина периода упреждения определяется отдельно для каждого экономического явления с учетом статистической колеблемости изучаемых данных на основе содержательного суждения о стабильности явления. Эта длина, как правило, не превышает для рядов годовых наблюдений одной трети объема данных, а для квартальных и месячных рядов — двух лет.

При выравнивании временных рядов с использованием кривых роста приходится решать вопрос о том, какой длины должен быть ряд, выбираемый для прогнозирования. Очевидно, что если период ряда экономической динамики слишком короткий, можно не обнаружить тенденцию его развития. С другой стороны, очень длительный временной ряд может охватывать периоды с различными трендами и его описание с помощью одной кривой роста не даст положительных результатов. Поэтому рекомендуется поступать следующим образом. Если нет никаких соображений качественного порядка, следует выбирать возможно больший промежуток времени.

Если развитие обнаруживает циклический характер, следует брать период от середины первого до середины последнего периода цикла. Если ряд охватывает периоды с разными трендами, лучше сократить ряд, отбросив наиболее ранние уровни, которые относятся к периоду с иной тенденцией развития.

При экстраполяционном прогнозировании экономической динамики с использованием трендовых моделей весьма важным является заключительный этап — **верификация прогноза**. Верификация любых дескриптивных моделей, к которым относятся трендовые модели, сводится к сопоставлению расчетных результатов по модели с соответствующими данными действительности — массовыми фактами и закономерностями экономического развития. Верификация прогнозной модели представляет собой совокупность критериев, способов и процедур, позволяющих на основе многостороннего анализа оценивать качество получаемого прогноза. Однако чаще всего на этапе верификации в большей степени осуществляется оценка метода прогнозирования, с помощью которого был получен результат, чем оценка качества самого результата. Это связано с тем, что до сих пор не найдено эффективного подхода к оценке качества прогноза до его реализации.

Даже в тех случаях, когда прогноз не оправдался, нельзя категорически утверждать, что он был бесполезен, поскольку пользователь, если он хотя бы частично контролирует ход событий и может воздействовать на экономический процесс, может использовать прогнозную информацию желаемым для себя образом. Так, получив прогноз событий, определяющих нежелательное направление перспективного развития, пользователь может принять меры, чтобы прогноз не оправдался; такой прогноз называется *самодеструктивным*. Если прогноз предсказал ход событий, устраивающий пользователя, то он может использовать свои возможности для увеличения вероятности правильного прогноза; подобный прогноз называется *саморегулирующим*. Таким образом, показателем ценности прогноза является не только его достоверность, но и полезность для пользователей.

О точности прогноза принято судить по величине ошибки прогноза — разности между фактическим значением исследуемого показателя и его прогнозным значением. Очевидно, что определить указанную разность можно лишь в двух случаях: либо если период упреждения уже окончился и известно фактическое значение прогнозируемого показателя (известна его реализация), либо если прогнозирование осуществлялось для некоторого момента времени в прошлом, для которого известны фактические данные. Во втором из названных случаев информация делится на две части. Часть, охватывающая более ранние данные, служит для оценивания параметров прогностической кривой роста, другая, более поздняя, рассматривается как реализация прогноза. Полученные таким образом ошибки прогноза в какой-то мере характеризуют точность применяемой методики прогнозирования.

Проверка точности одного прогноза недостаточна для оценки качества прогнозирования, так как она может быть результатом случайного совпадения. Наиболее простой мерой качества прогнозов при условии, что имеются данные об их реализации, является отношение числа случаев, когда фактическая реализация охватывалась интервальным прогнозом, к общему числу прогнозов. Данную меру качества прогнозов k можно вычислить по формуле

$$k=\frac{p}{p+q},$$

где p — число прогнозов, подтвержденных фактическими данными;

q — число прогнозов, не подтвержденных фактическими данными.

Однако в практической работе проблему качества прогнозов чаще приходится решать, когда период упреждения еще не закончился и фактическое значение прогнозируемого показателя неизвестно. В этом случае более точной считается модель, дающая более узкие доверительные интервалы прогноза. На практике не всегда удается сразу построить достаточно хорошую модель прогнозирования, поэтому описанные в данной главе этапы построения трендовых моделей экономической динамики выполняются неоднократно.

Рассмотрим пример расчета точечного и интервального прогноза на основе трендовых моделей, используя данные задачи, решаемой в предыдущем параграфе данной главы.

Пример 5.2. Пусть для временного ряда, представленного в табл. 5.3, требуется дать прогноз на два шага вперед (t = 10 и t = 11) на основе адекватной линейной модели

$$\hat{y}_t = 87{,}8 - 3{,}4t.$$

Решение. Точечные прогнозы получим, подставляя в уравнение модели значения $t = 10$ и $t = 11$:

$$\hat{y}_{10} = 87{,}8 - 3{,}4 \cdot 10 = 53{,}8;$$

$$\hat{y}_{11} = 87{,}8 - 3{,}4 \cdot 11 = 50{,}4.$$

При расчете доверительных интервалов прогноза учтем, что в процессе решения упомянутой задачи предыдущего параграфа было найдено значение средней квадратической ошибки оценки прогнозируемого показателя $S_{\hat{y}} = 1{,}39$, а значения величины K в формуле (5.19) для ряда из девяти уровней можно получить при уровне значимости $\alpha = 0{,}20$ из табл. 5.4 путем линейной интерполяции приведенных значений для $n = 7$ и $n = 10$: для $t = 10$ ($L = 1$) $K = 1{,}77$; для $t = 11$ ($L = 2$) $K = 1{,}88$. Результаты расчета по формуле (5.19) представлены в табл. 5.5.

Таблица 5.5

Время (t)	Шаг (L)	Точечный прогноз ($\hat{y}_{n+L}$)	Доверительный интервал прогноза	
			Нижняя граница	Верхняя граница
10	1	53,8	51,3	56,3
11	2	50,4	47,8	53,0

Так как модель, на основе которой осуществлялся прогноз, признана адекватной, то с принятым уровнем значимости 0,20, другими словами, с доверительной вероятностью 0,80 (или 80%) можно утверждать, что при сохранении сложившихся закономерностей развития прогнозируемая величина попадет в интервал, образованный нижней и верхней границами.

5.4. Адаптивные модели прогнозирования

Как уже выше отмечено, в основе экстраполяционных методов прогнозирования лежит предположение о том, что основные факторы и тенденции, имевшие место в прошлом, сохраняются в будущем. Сохранение этих тенденций — непременное условие успешного прогнозирования. При этом необходимо, чтобы учитывались лишь те тенденции, которые еще не устарели и до сих пор оказывают влияние на изучаемый процесс.

При краткосрочном прогнозировании, а также при прогнозировании в ситуации изменения внешних условий, когда наиболее важными являются последние реализации исследуемого процесса, наиболее эффективными оказываются адаптивные методы, учитывающие неравноценность уровней временного ряда.

Адаптивные модели прогнозирования — это модели дисконтирования данных, способные быстро приспосабливать свою структуру и параметры к изменению условий. Инструментом прогноза в адаптивных моделях, как и в кривых роста, является математическая модель с единственным фактором «время».

При оценке параметров адаптивных моделей в отличие от рассматриваемых ранее моделей «кривых роста» наблюдениям (уровням ряда) присваиваются различные веса в зависимости от того, насколько сильным признается их влияние на текущий уровень. Это позволяет учитывать изменения в тенденции, а также любые колебания, в которых прослеживается закономерность. Все адаптивные модели базируются на двух схемах: скользящего среднего (СС-модели) и авторегрессии (АР-модели).

Согласно схеме скользящего среднего оценкой текущего уровня является взвешенное среднее всех предшествующих уровней, причем веса при наблюдениях убывают по мере удаления от последнего уровня, т.е. информационная ценность наблюдений признается тем большей, чем ближе они к концу интервала наблюдений. Такие модели хорошо отражают изменения, происходящие в тенденции, но в чистом виде не позволяют отражать колебания.

Реакция на ошибку прогноза и дисконтирование уровней временного ряда в моделях, базирующихся на схеме СС, определяется с помощью параметров сглаживания (адаптации), значения которых могут изменяться от нуля до единицы. Высокое значение этих параметров (свыше 0,5) означает придание большего веса последним уровням ряда, а низкое (менее 0,5) — предшествующим наблюдениям. Первый случай соответствует быстроизменяющимся динамичным процессам, второй — более стабильным.

В авторегрессионной схеме оценкой текущего уровня служит взвешенная сумма не всех, а нескольких предшествующих уровней, при этом весовые коэффициенты при наблюдениях не ранжированы. Информационная ценность наблюдений определяется не их близостью к моделируемому уровню, а теснотой связи между ними. Общая схема построения адаптивных моделей может быть представлена следующим образом. По нескольким первым уровням ряда оцениваются значения параметров модели. По имеющейся модели строится прогноз на один шаг вперед, причем его отклонение от фактических уровней ряда расценивается как ошибка прогнозирования, которая учитывается в соответствии с принятой схемой корректировки модели. Далее по модели со скорректированными параметрами рассчитывается прогнозная оценка на следующий момент времени и т.д. Таким образом, модель постоянно «впитывает» новую информацию и к концу периода обучения отражает тенденцию развития процесса, существующую в данный момент.

В практике статистического прогнозирования наиболее часто используются две базовые СС-модели — Брауна и Хольта, первая из них является частным случаем второй. Эти модели представляют процесс развития как линейную тенденцию с постоянно изменяющимися параметрами.

Модель Брауна (модель экспоненциального сглаживания). Модель Брауна может отображать развитие не только в виде линейной тенденции, но также в виде случайного процесса, не имеющего тенденции, а также в виде изменяющейся параболической тенденции. Соответственно различают модели Брауна:

- нулевого порядка, которая описывает процессы, не имеющие тенденции развития. Она имеет лишь один параметр a_0 (оценка текущего уровня). Прогноз развития на τ шагов вперед осуществляется согласно формуле $y_{t+1} = a_0$. Такая модель еще называется «наивной» («будет, как было»). Доверительный интервал прогноза получают по формуле[1] $y_{t+\tau} \pm t_\alpha \hat{\sigma}\sqrt{\alpha / (2-\alpha)}$;
- первого порядка $y_{t+\tau} = (a_0 + a_1 \cdot \tau)$. Коэффициент a_0 — значение, близкое к последнему уровню, и представляет как бы закономерную составляющую этого уровня. Коэффициент a_1 определяет прирост, сформировавшийся в основном к концу периода наблюдений, но отражающий также (правда, в меньшей степени) скорость роста на более ранних этапах. Прогноз осуществляется по формуле $y_{t+k} = a_0 + a_1 \cdot \tau$. Доверительный интервал прогноза получают по формуле

$$y_{t+\tau} \pm t_\alpha \hat{\sigma}\sqrt{\frac{\alpha\left[1+4(1-\alpha)+5(1-\alpha)+2\alpha(4-3\alpha)\tau+2\alpha^2\tau^2\right]}{(2-\alpha)^3}};$$

[1] *Колемаев В. А.* Эконометрика : учеб. / В. А. Колемаев. М. : ИНФРА-М, 2004. 160 с.

- второго порядка, отражающую развитие в виде параболической тенденции с изменяющимися «скоростью» и «ускорением». Она имеет три параметра (a_2 — оценка текущего прироста или «ускорение»). Прогноз осуществляется по формуле $y_{t+k} = a_0 + a_1 \cdot \tau + a_2 \cdot \tau^2$.

Рассмотрим этапы построения линейной адаптивной модели Брауна.

Этап 1. По первым пяти точкам временного ряда оцениваются значения a_0 и a_1 параметров модели с помощью метода наименьших квадратов для линейной аппроксимации по формуле

$$\hat{y}_t = \hat{a}_0 + \hat{a}_1 t.$$

Этап 2. С использованием параметров $\hat{a}_0$ и $\hat{a}_1$, которые соответствуют нулевому моменту времени, по модели Брауна находим прогноз на первый шаг $(\tau = 1)$:

$$\hat{y}_1 = \hat{a}_{0(0)} + \hat{a}_{1(0)} \cdot \tau = \hat{a}_{0(0)} + \hat{a}_{1(0)}.$$

Этап 3. Расчетное значение $\hat{y}$ экономического показателя сравнивают с фактическим значением y_i и находят величину отклонения e_i:

$$e_1 = y_1 - \hat{y}_1.$$

Для всех остальных членов ряда отклонение (остаточная компонента) находится по формуле $e_{(t)} = y_{(t)} - \hat{y}_{(t)}$, которое используют для корректировки параметров модели в соответствии с принятой схемой.

Этап 4. Корректируют параметры модели $a_{0(t)}$ и по следующим формулам:

$$\begin{aligned} a_{0(t)} &= a_{0(t-1)} + a_{1(t-1)} + \left(1 - \beta^2\right) \cdot e_{(t)}; \\ a_{1(t)} &= a_{1(t-1)} + \left(1 - \beta\right)^2 \cdot e_{(t)}, \end{aligned} \qquad (5.23)$$

где β — коэффициент дисконтирования данных, отражающий большую степень доверия более поздним наблюдениям; α — параметр сглаживания $(1 - \beta = \alpha)$. Оптимальное значение β находится итеративным путем, т.е. многократным построением модели при разных значениях β и выбором наилучшей. Параметры вычисляются последовательно, от уровня к уровню, и их значения для последнего уровня определяют окончательный вид модели.

Этап 5. По модели со скорректированными параметрами $a_{0(t)}$ находят прогноз на следующий момент времени $(\tau = 1)$:

$$\hat{y}_{(t)}(\tau) = a_{0(t)} + \tau \cdot a_{1(t)} = \hat{y}_{(t+1)} = a_{0(t)} + a_{1(t)}.$$

Этап 6. Возврат на пункт 3, если $t = n$.

Если $t = n$, то построенную модель можно использовать для прогнозирования на будущее. Точечный прогноз рассчитывается по формуле

$$\hat{y}_{(t+\tau)} = a_{0(t)} + \tau \cdot a_{1(n)},$$
$$\tau = 1, 2, \ldots$$

Пример 5.3. Построим прогноз по линейной модели Брауна.

По имеющейся информации (y в табл. 5.6) об объемах продаж нового товара (тыс. руб.) в течение 35 недель построить адаптивную модель Брауна с линейной тенденцией. Построить прогноз на три шага вперед, используя значение параметра сглаживания 0,3. Результаты моделирования и прогнозирования привести на графике.

Таблица 5.6

t	1	2	3	4	5	6	7
y	27,30	41,80	42,80	56,20	72,50	56,00	79,00

t	8	9	10	11	12	13	14
y	74,90	103,30	111,30	125,20	189,30	169,10	193,50

t	15	16	17	18	19	20	21
t	207,40	221,20	267,20	264,00	273,80	321,00	317,40

t	22	23	24	25	26	27	28
t	342,00	350,60	368,50	397,00	382,90	400,60	409,40

t	29	30	31	32	33	34	35
y	426,00	402,00	398,70	418,10	424,60	435,10	439,8

1) По первым пяти точкам временного ряда оцениваются значения $a_{0(t)}$ и параметров модели с помощью метода наименьших квадратов для линейной аппроксимации:

$$\hat{y}_t = \hat{a}_0 + \hat{a}_1 t.$$

Получаем начальные значения параметров модели $\hat{a}_0$ = 16,68 и $\hat{a}_1$ = 10,48, которые соответствуют моменту времени t = 0.

2) С использованием параметров $\hat{a}_{0(0)}$ и $\hat{a}_{1(0)}$ по модели Брауна находим прогноз на первый шаг $(\tau = 1)$: $\hat{y}$ = $\hat{a}_{0(0)}$ + $\hat{a}_{1(0)}$ = = 16,68 + 10,48 = 27,16.

3) Находим величину отклонения:

$$e_t = y_t - \hat{y}_t = e_1 = y_1 - \hat{y}_1 = 27{,}3 - 27.16 = 0{,}14\ .$$

Все расчеты в табл. 5.7 показаны для $\alpha = 0{,}3,\ \beta = 0{,}7.$

4) Корректируем параметры модели $\hat{a}_{0(t)}$ и $\hat{a}_{1(t)}$ по следующим формулам:

$t = 1,$

$$a_{0(t)} = a_{0(t-1)} + a_{1(t-1)} + \left(1 - \beta^2\right) \cdot e_{(t)} = 16{,}68 + 10{,}48 + \left(1 - 0{,}7^2\right) \cdot 0{,}14 =$$
$$= 27{,}231;$$

$$a_{1(t)} = a_{1(t-1)} + \left(1 - \beta\right)^2 \cdot e_{(t)} = 10{,}48 + \left(1 - 0{,}7\right)^2 \cdot 0{,}14 = 10{,}493.$$

Таблица 5.7

Оценка параметров модели Брауна

t	y_t	a_0	a_1	$\hat{y}_t$	$e_{(t)} = y_{(t)} - \hat{y}_{(t)}$	$\frac{\lvert e_t \rvert}{y_t} \cdot 100\%$
0		16,68	10,48			
1	27,3	27,231	10,493	27,160	0,140	0,513
2	41,8	39,803	10,859	37,724	4,076	9,751
3	42,8	46,652	10,152	50,662	–7,862	18,370
4	56,2	56,496	10,097	56,804	–0,604	1,075
5	72,5	69,606	10,629	66,594	5,906	8,147
...	...	...	...	...	...	...
32	418,1	419,625	6,628	421,213	–3,113	0,744
33	424,6	425,410	6,479	426,254	–1,654	0,389
34	435,1	433,527	6,768	431,890	3,210	0,738
35	439,8	440,043	6,724	440,295	–0,495	0,113

5) По модели со скорректированными параметрами $a_{0(t)}$ находим прогноз на следующий момент времени $\tau = 1$, $t = 2$:

$$\hat{y}_{(t)}(\tau) = a_{0(t)} + \tau \cdot a_{1(t)} = \hat{y}_{(t+1)} = \hat{y}_{(2)}(1) = a_{0(1)} + a_{1(1)} = 27{,}231 + 10{,}493 = 37{,}724.$$

Возврат к пункту 3.

Вычисления повторяем до конца наблюдений.

6) Параметры модели, полученные в последний момент времени ($t = 35$), используем для построения прогноза на три недели вперед $(\tau = 1,\ 2,\ 3)$;

$$\hat{y}_{35}(1) = \hat{y}_{36} = a_{0(35)} + a_{1(35)} \cdot 1 = 440{,}043 + 6{,}724 \cdot 1 = 446{,}767;$$

$$\hat{y}_{35}(2) = \hat{y}_{37} = a_{0(35)} + a_{1(35)} \cdot 2 = 440{,}043 + 6{,}724 \cdot 2 = 453{,}490;$$

$$\hat{y}_{35}(3) = \hat{y}_{38} = a_{0(35)} + a_{1(35)} \cdot 3 = 440{,}043 + 6{,}724 \cdot 3 = 460{,}214.$$

Результаты прогнозирования по модели Брауна представлены графически на рис. 5.1.

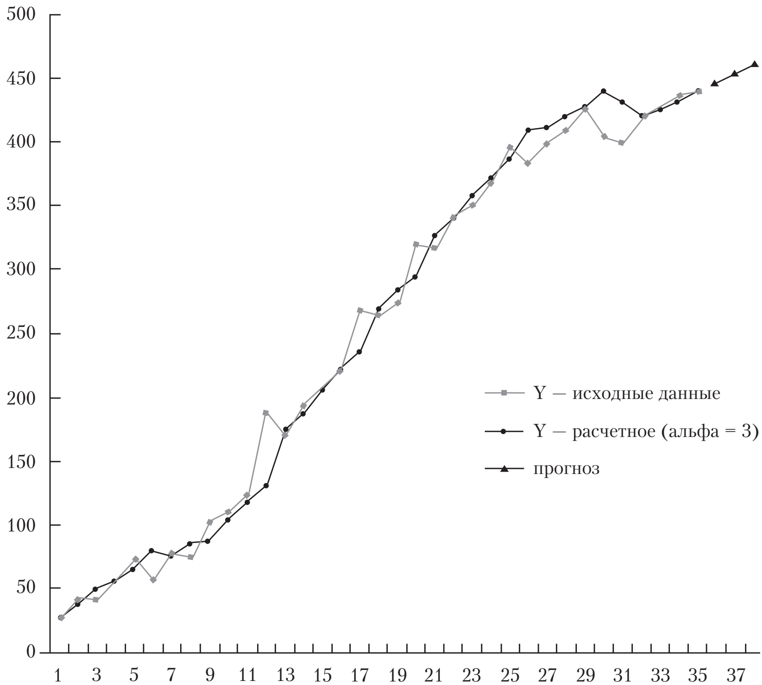


Рис. 5.1. **Исходные данные, результаты моделирования (альфа = 0,3) и точечный прогноз**

В моделях Брауна и Хольта параметры сглаживания характеризуют степень адаптации модели к изменению ряда наблюдений. Они определяют скорость реакции модели на изменения, происходящие в развитии. Чем они больше, тем быстрее реагирует модель на изменения. Обычно для устойчивых рядов их величина большая, а для неустойчивых — маленькая. В различных методах прогнозирования используется различный подход к их определению. Их можно взять фиксированными, а наилучшее значение определить методом подбора, чтобы ошибка прогноза на один шаг вперед была наименьшей. При использовании компьютера это не представляет труда. Альтернативу этому подходу составляет динамическое изменение параметров сглаживания. В методах эволюции и симплекс-планирования параметры адаптации постоянно меняются на каждом шаге. Для каждого параметра сглаживания формируется несколько значений.

Модели и методы авторегрессии

Авторегрессионные модели широко используются для описания стационарных случайных процессов. Характерной особенностью стационарных временных рядов является то, что их вероятностные свойства рядов не изменяются во времени. Иначе говоря, функции распределения стационарных динамических рядов не меняются при сдвиге времени.

Модель, в которой расчетные значения уровней ряда определяются как линейная функция от предыдущих наблюдений, называют авторегрессионной. Если текущая величина уровня ряда y_t зависит только от одного предшествующего значения y_{t-1}, то такая модель является авторегрессионной моделью первого порядка АР (1), если y_t зависит от двух предшествующих уровней y_{t-1} и y_{t-2}, моделью второго порядка АР (2) и т.д. до порядка p, т.е. АР (p).

Если АР-модель имеет первый ($p = 1$) или второй порядок ($p = 2$), то она приобретает соответственно следующий вид:

$$y_t = a_1 y_{t-1} + \varepsilon_t;$$

$$y_t = a_1 \cdot y_{t-1} + a_2 \cdot y_{t-2} + \varepsilon_t.$$

Идентификация АР(p) модели состоит в определении ее порядка p. Одной из предпосылок построения модели этого типа является применение их к стационарному процессу. Поэтому в более широком смысле идентификация модели включает также выбор способа трансформации исходного ряда наблюдений, как правило, имеющего некоторую тенденцию, в стационарный (или близкий к нему) ряд. Один из наиболее распространенных способов решения этой проблемы — последовательное взятие разностей, т.е. переход от исходного ряда к ряду первых, а затем и вторых разностей.

«Чистые» авторегрессионные процессы имеют плавно затухающую автокорреляционную функцию (АКФ). В этом случае в качестве порядка модели выбирается лаг, после которого все частные автокорреляционные функции (ЧАКФ) имеют незначительную величину. Однако на практике редко встречаются процессы, которые легко было бы идентифицировать. Поэтому порядок модели обычно определяется методом проб из нескольких альтернатив. В число кандидатов включаются модели, у которых порядок соответствует ЧАКФ, превышающей стандартное отклонение $1/N$. При обработке разностных рядов иногда ориентируются на АКФ, выбирая модели, у которых порядок соответствует максимальному ее значению, при условии, что оно превышает стандартное отклонение.

Ряды без тенденции, как правило, не представляют интереса для экономистов. АР-модели вообще не предназначены для описания процессов с тенденцией, однако они хорошо описывают колебания, что весьма важно для отображения развития неустойчивых показателей.

Чтобы применить АР-модели к экономическим процессам с тенденцией, на первом этапе формируют стационарный ряд, исключая тенденцию, путем перехода от исходного ряда к ряду разностей соседних значений членов ряда. Например, переход от исходного ряда $y_t (t = 1, 2, ..., n)$ к ряду $\Delta y_t (t = 1, 2, ..., n - d)$ первых ($d = 1$) или вторых ($d = 2$) разностей осуществляется следующим образом:

$$\Delta^0 y_t = y_t, \qquad t = 1, 2, ..., n \qquad \text{при } d = 0;$$

$$\Delta^1 y_t = y_{t+1} - y_t, \qquad t = 1, 2, ..., n - 1 \qquad \text{при } d = 1;$$

$$\Delta^2 y_t = y_{t+1} - \Delta y_t, \qquad t = 1, 2, ..., n - 2 \qquad \text{при } d = 2.$$

Первоначальный (исходный) ряд является интегрированным рядом первого порядка, когда его первые разности образуют стационарный ряд динамики. Если для формирования стационарного временного ряда требуется получить ряд вторых разностей, то исходный ряд называется интегрированным рядом второго порядка и т.д.

После перехода к разностным рядам авторегрессионная модель порядка m первых разностей (приростов) и вторых разностей приобретает соответственно такой вид:

$$\Delta\hat{y}_t = a_0 + a_1\Delta y_{t-1} + a_2\Delta y_{t-2} + \ldots + a_m\Delta y_{t-m};$$
$$\Delta^2\hat{y}_t = a_0 + a_1\Delta^2 y_{t-1} + a_2\Delta^2 y_{t-2} + \ldots + a_m\Delta^2 y_{t-m}.$$

Таким образом, АР-модели разностных временных рядов характеризуются двумя параметрами: p (порядок авторегрессии) и d (порядок конечных разностей), поэтому записываются как АР (p, d).

Простейшим способом определения наиболее подходящего разностного ряда является вычисление для каждого ряда ($d = 0, 1, 2$) его дисперсии. Для дальнейшей обработки выбирается ряд, у которого величина этого показателя минимальна.

Для идентификации порядка модели может быть использована частная автокорреляционная функция. Если для m лагов частные коэффициенты автокорреляции статистически значимы, а затем для лагов $m + 1$ и далее резко падают до нулевого значения, то это указывает на авторегрессионный процесс порядка m.

При моделировании нестационарных по своей природе экономических процессов авторегрессионная функция объединяется с другими методами анализа динамики: скользящей средней, трендом, сезонной волной. Объединение разных моделей в единое целое существенно расширяет сферу их использования.

Построение моделей авторегрессии и получение прогнозных оценок исходного ряда возможно только с использованием программных средств. При грамотном использовании хорошие результаты дает модель Бокса — Дженкинса.

Данную модель также называют авторегрессионной интегрированной моделью скользящего среднего или сокращенно АРИСС (p, d, q).

При построении АРИСС-моделей с использованием программных средств осуществляются:

а) идентификация динамического ряда (определение размерности операторов конечной разности d, авторегрессии p и скользящего среднего q);

б) оценивание параметров модели;

в) проверка адекватности модели.

Пример 5.4. Рассмотрим построение прогноза объема продаж с использованием программы SPSS.

В результате расчетов в качестве лучшей выбрана модель ARIMA (1, 1, 0). Ниже в табл. 5.8 и 5.9 приведены расчеты построения прогноза объема продаж, выполненные с использованием программы SPSS. На рис. 5.2 представлены результаты аппроксимации и прогнозирования по этой модели. Построенная модель характеризуется высоким коэффициентом детерминации 0,986 и низким значением средней относительной ошибки аппроксимации 6,48%.

Таблица 5.8

Параметры модели АРПСС

					Оценка	Стандартная ошибка	t	Знач.
Y-Модель_1	Y	Нет преобразования	Константа		12,175	2,144	5,679	,000
			АР	Лаг 1	-,390	,162	-2,400	,022
			Дифференцирование		1			

Таблица 5.9

Прогноз объема продаж, границы доверительного интервала

Неделя	Прогноз	Нижняя граница $p = 95\%$	Верхняя граница $p = 95\%$
36	454,89	419,80	489,98
37	465,93	424,82	507,04
38	478,55	429,50	527,59
39	490,55	435,65	545,45

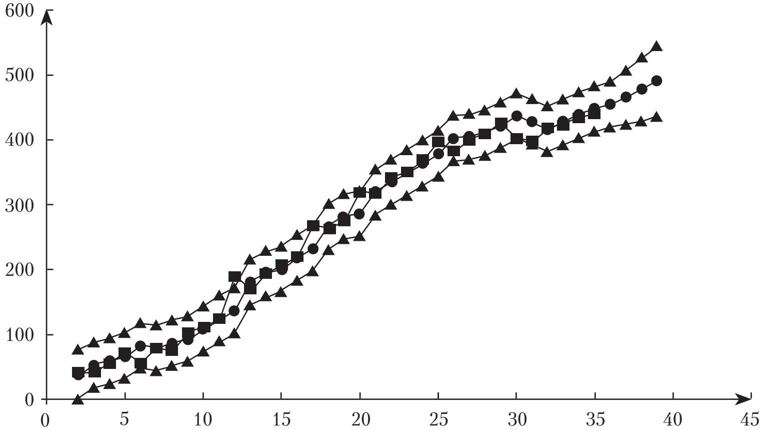


Рис. 5.2. **Прогнозирование объема продаж по модели ARIMA (1, 1, 0)**

Вопросы и задания

1. В чем суть прогнозирования экономических процессов на основе метода экстраполяции?
2. Дайте характеристику основных типов кривых роста, наиболее часто используемых при построении трендовых моделей прогнозирования.
3. Укажите методы предварительного выбора кривой роста. Как находятся параметры этих кривых?
4. Каким образом проводится оценка адекватности трендовых моделей? Какие статистические критерии при этом используются?
5. Назовите статистические критерии оценки точности моделей прогнозирования в экономике.
6. Перечислите основные этапы прогнозирования экономической динамики на основе одномерных временных рядов с использованием трендовых моделей.
7. Опишите порядок получения точечного и интервального прогноза экономического показателя на основе трендовых моделей. От каких факторов зависит ширина доверительного интервала прогноза?
8. Поясните суть адаптивных методов прогнозирования. Какие типы адаптивных моделей вы знаете?
9. Укажите этапы построения и использования адаптивной модели Брауна. Как влияет параметр сглаживания на скорость адаптации моделей этого типа к изменениям в прогнозируемом процессе?
10. Дайте краткую характеристику авторегрессионных моделей прогнозирования. Для каких экономических процессов применимы методы авторегрессии?

Упражнения

1. Временной ряд задан в таблице:

t	1	2	3	4	5	6	7	8	9	10
y_t	43	47	50	48	54	57	61	59	65	62

Сделайте предварительный выбор наилучшей кривой роста:

а) методом конечных разностей (Тинтнера);

б) методом характеристик прироста.

2. Для ряда, приведенного в упр. 1, построить линейную модель $\hat{y}_t = a_0 + a_1 t,$ определив ее параметры методом наименьших квадратов.

3. Для временного ряда из упр. 1 построить адаптивную модель Брауна с параметром сглаживания $\alpha = 0{,}4$ и $\alpha = 0{,}7$; выбрать наилучшую модель Брауна $\hat{y}(k) = a_0 + a_1 k,$ где k — период упреждения (количество шагов вперед).

4. Оценить адекватность моделей, построенных в упр. 2 и 3, на основе исследования:

а) близости математического ожидания остаточной компоненты нулю; критическое значение статистики Стьюдента принять $t_\alpha = 1{,}09$ (для доверительной вероятности 0,70);

б) случайности отклонений остаточной компоненты по критерию пиков (поворотных точек); расчеты выполнить на основе соотношения (5.9);

в) независимости (отсутствия автокорреляции) уровней ряда остатков либо по критерию Дарбина — Уотсона (в качестве критических используйте уровни $d_1 = 1{,}08$ и $d_2 = 1{,}36$), либо по первому коэффициенту автокорреляции (критический уровень принять равным $r_1 = 0{,}36$);

г) нормальности закона распределения остаточной компоненты на основе *RS*-критерия (в качестве критических уровней принять интервал 2,7 — 3,7).

5. Оценить точность моделей, построенных в упр. 2 и 3, используя показатели среднего квадратического отклонения и средней относительной ошибки аппроксимации.

6. На основе сравнительного анализа адекватности и точности моделей по результатам упр. 4 и 5 выбрать лучшую модель, по которой построить точечный и интервальный прогнозы на два шага вперед ($t_\alpha = 1{,}09$). Результату прогнозирования отразить графически. Построить и проиллюстрировать на графике точечный и интервальный прогнозы на два шага вперед по линейной трендовой модели, полученной в упр. 2.

Глава 6
БАЛАНСОВЫЕ МОДЕЛИ

— Балансовый метод. Принципиальная схема межпродуктового баланса

— Экономико-математическая модель межотраслевого баланса

— Коэффициенты прямых и полных материальных затрат

— Межотраслевые балансовые модели в анализе экономических показателей

— Динамическая межотраслевая балансовая модель

В результате усвоения материалов данной главы студенты должны:

***знать*:**

- сущность балансовых методов в экономике;
- структуру общей схемы межотраслевого баланса и экономический смысл параметров модели В. Леонтьева;
- возможности использования балансовых моделей для анализа различных экономических показателей;
- особенности построения и использования динамических межотраслевых балансовых моделей;

***уметь*:**

- строить экономико-математическую модель межотраслевого баланса (ЭММ МОБ);
- проводить на основе ЭММ МОБ оценку продуктивности экономики и расчеты объемов конечной и валовой продукции отраслей;
- использовать балансовые модели для построения межотраслевых балансов затрат труда и затрат фондов;

***владеть*:**

- понятийным аппаратом балансового метода в экономике, включая аппарат ЭММ МОБ в статической и динамической постановке;
- методами расчета коэффициентов межотраслевых прямых, косвенных и полных материальных затрат, а также затрат труда и фондов.

6.1. Балансовый метод. Принципиальная схема межпродуктового баланса

Балансовые модели, как статические, так и динамические, широко применяются при экономико-математическом моделировании экономических систем и процессов. В основе создания этих моделей лежит балансовый метод, т.е. метод взаимного сопоставления имеющихся материальных, трудовых и финансовых ресурсов и потребностей в них. Если описывать экономическую систему в целом, то под балансовой моделью понимается система уравнений, каждое из которых выражает требование баланса между производимым отдельными экономическими объектами количеством продукции и совокупной потребностью в этой продукции. При таком подходе рассматриваемая система состоит из экономических объектов, каждый из которых выпускает некоторый продукт, часть его потребляется другими объектами системы, а другая часть выводится за пределы системы в качестве ее конечного продукта.

Если вместо понятия **продукт** ввести более общее понятие **ресурс**, то под *балансовой моделью* следует понимать систему уравнений, которые удовлетворяют требованиям соответствия наличия ресурса и его использования. Кроме приведенного выше требования соответствия производства каждого продукта и потребности в нем, можно указать такие примеры балансового соответствия, как соответствие наличия рабочей силы и количества рабочих мест, платежеспособного спроса населения и предложения товаров и услуг и т.д. При этом соответствие понимается либо как равенство, либо менее жестко — как достаточность ресурсов для покрытия потребности и, следовательно, наличие некоторого резерва.

Важнейшие виды балансовых моделей:

— частные материальные, трудовые и финансовые балансы для народного хозяйства и отдельных отраслей;

— межотраслевые балансы;

— матричные техпромфинпланы предприятий и фирм.

Балансовый метод и создаваемые на его основе балансовые модели служат основным инструментом поддержания пропорций в народном хозяйстве. Балансовые модели на базе отчетных балансов характеризуют сложившиеся пропорции, в них ресурсная часть всегда равна расходной. Для выявления диспропорций используются балансовые модели, в которых фактические ресурсы сопоставлялись бы не с их фактическим потреблением, а с потребностью в них. В связи с этим необходимо отметить, что балансовые модели не содержат какого-либо механизма сравнения отдельных вариантов экономических решений и не предусматривают взаимозаменяемости разных ресурсов, что не позволяет сделать выбор оптимального вариант развития экономической системы. Этим определяется ограниченность балансовых моделей и балансового метода в целом.

Основу информационного обеспечения балансовых моделей и экономике составляет матрица коэффициентов затрат ресурсов но конкретным направлениям их использования. Например, в модели межотраслевого баланса такую роль играет так называемая технологическая матрица — таблица межотраслевого баланса, составленная из коэффициентов (нормативов) прямых затрат на производство единицы продукции в натуральном выражении. По многим причинам исходные данные реальных хозяйственных объектов не могут быть использованы в балансовых моделях непосредственно, поэтому подготовка информации для ввода в модель является весьма серьезной проблемой. Так, при построении модели межотраслевого баланса используется специфическое понятие чистой (или технологической) отрасли, т.е. условной отрасли, объединяющей все производство данного продукта независимо от ведомственной (административной) подчиненности и форм собственности предприятий и фирм. Переход от хозяйственных отраслей к чистым отраслям требует специального преобразования реальных данных хозяйственных объектов, например агрегирования отраслей, исключения внутриотраслевого оборота и др. В этих условиях понятия «межпродуктовый баланс» и «межотраслевой баланс» практически идентичны, отличие заключается лишь в единицах измерения элементов баланса.

Как отмечено выше, балансовые модели строятся в виде числовых матриц — прямоугольных таблиц чисел. В связи с этим балансовые модели относятся к тому типу экономико-математических моделей, которые называются матричными. В матричных моделях балансовый метод получает строгое математическое выражение. Таким образом, матричную структуру имеют межотраслевой и межрайонный балансы производства и распределения продукции в народном хозяйстве, модели развития отраслей, межотраслевые балансы производства и распределения продукции отдельных регионов, модели промфинпланов предприятий и фирм. Несмотря на специфику этих моделей, их объединяет не только общий формальный (матричный) принцип построения и единство системы расчетов, но и аналогичность ряда экономических характеристик. Это позволяет рассматривать структуру, содержание и основные зависимости матричных моделей на примере одной из них, а именно на примере межотраслевого баланса производства и распределения продукции в народном хозяйстве. Данный баланс отражает производство и распределение общественного продукта в отраслевом разрезе, межотраслевые производственные связи, использование материальных и трудовых ресурсов, создание и распределение национального дохода.

Принципиальная схема межотраслевого баланса производства и распределения совокупного общественного продукта в стоимостном выражении приведена в табл. 6.1. В основу этой схемы положено разделение совокупного продукта на две части: промежуточный и конечный продукты; все народное хозяйство представлено в виде совокупности n отраслей (имеются в виду чистые отрасли), при этом каждая отрасль фигурирует в балансе как производящая и как потребляющая.

Таблица 6.1

Принципиальная схема межотраслевого баланса (МОБ)

Производящие отрасли	Потребляющие отрасли					Конечный продукт	Валовой продукт
	1	2	3	...	n		
1	x_{11}	x_{12}	x_{13}	...	x_{1n}	Y_1	X_1
2	x_{21}	x_{22}	x_{23}	...	x_{2n}	Y_2	X_2
3	x_{31}	x_{32}	x_{33}	...	x_{3n}	Y_3	X_3
...	...	...		...	...	...	...
...	...	...		I	...	II	...
...	...	...		...	...	...	...
n	x_{n1}	x_{n2}		...	x_{nn}	Y_n	X_n
Амортизация	c_1	c_2	c_3	...	c_n		
Оплата труда	v_1	v_2	v_3	III	v_n	IV	
Чистый доход	m_1	m_2	m_3	...	m_n		
Валовой продукт	X_1	X_2	X_3	...	X_n		$\sum_{i=1}^{n} X_i = \sum_{j=1}^{n} Xj$

Рассмотрим схему МОБ в разрезе его крупных составных частей. Выделяются четыре части, имеющие различное экономическое содержание, они называются квадрантами баланса и на схеме обозначены римскими цифрами.

Первый квадрант МОБ — это шахматная таблица межотраслевых материальных связей. Показатели, помещенные на пересечениях строк и столбцов, представляют собой величины межотраслевых потоков продукции и в общем виде обозначаются x_{ij}, где i и j — соответственно номера отраслей производящих и потребляющих. Так, величина x_{32} понимается как стоимость средств производства, произведенных в отрасли с номером 3 и потребленных в качестве материальных затрат в отрасли с номером 2. Таким образом, первый квадрант по форме представляет собой квадратную матрицу порядка n, сумма всех элементов которой равняется годовому фонду возмещения затрат средств производства в материальной сфере.

Во *втором квадранте* представлена конечная продукция всех отраслей материального производства, при этом под конечной понимается продукция, выходящая из сферы производства в область конечного использования (на потребление и накопление). В табл. 6.1 этот раздел дан укрупненно в виде одного столбца величин Y_i, в развернутой схеме баланса, конечный продукт каждой отрасли показан дифференцированно по направлениям использования — на личное потребление населения, общественное потребление, на накопление, возмещение потерь, экспорт и др. Второй квадрант характеризует отраслевую материальную структуру национального дохода, а в развернутом виде — также распределение национального дохода на фонд накопления и фонд потребления, структуру потребления и накопление по отраслям производства и потребителям.

Третий квадрант МОБ также характеризует национальный доход, но со стороны его стоимостного состава как сумму чистой продукции и амортизации; чистая продукция понимается при этом как сумма оплаты труда и чистого дохода отраслей. Сумму амортизации (c_j) и чистой продукции ($v_j + m_j$) некоторой j-й отрасли будем называть условно чистой продукцией этой отрасли и обозначать в дальнейшем Z_j. В системе национального счетоводства данные этого квадранта соответствуют валовой добавленной стоимости.

Четвертый квадрант баланса находится на пересечении столбцов второго квадранта (конечной продукции) и строк третьего квадранта (условно чистой продукции). Этим определяется содержание квадранта: он отражает конечное распределение и использование национального дохода, а также содержит амортизационные расходы. В результате перераспределения первоначально созданного национального дохода образуются конечные доходы населения, предприятий, государства. Данные четвертого квадранта важны для отражения в межотраслевой модели баланса доходов и расходов населения, источников финансирования капиталовложений, текущих затрат непроизводственной сферы, для анализа общей структуры конечных доходов по группам потребителей. Более детально составляющие элементы этого квадранта в данном пособии не рассматриваются, однако очень важным является тот факт, что общий итог четвертого квадранта, так же как второго и третьего, должен быть равен созданному за год национальному доходу плюс амортизационные отчисления.

Таким образом, в целом межотраслевой баланс в рамках единой модели объединяет балансы отраслей материального производства, баланс совокупного общественного продукта, балансы национального дохода, финансовый, баланс доходов и расходов населения. Следует особо отметить, что хотя валовая продукция отраслей не входит в рассмотренные выше четыре квадранта, она представлена на принципиальной схеме МОБ в двух местах в виде столбца, расположенного справа от второго квадранта, и в виде строки ниже третьего квадранта. Эти столбец и строка валовой продукции замыкают схему МОБ и играют важную роль как для проверки правильности заполнения квадрантов (т.е. проверки самого баланса), так и для разработки экономико-математической модели межотраслевого баланса. Если, как показано на схеме, обозначить валовой продукт некоторой отрасли буквой X с нижним индексом, равным номеру данной отрасли, то можно записать два важнейших соотношения, отражающих сущность МОБ и являющихся основой его экономико-математической модели.

Во-первых, рассматривая схему баланса по столбцам, можно сделать очевидный вывод, что итог материальных затрат любой потребляющей отрасли и ее условно чистой продукции равен валовой продукции этой отрасли. Данный вывод можно записать в виде соотношения

$$X_j = \sum_{i=1}^{n} x_{ij} + Z_j, \ \ j = \overline{1,n}. \tag{6.1}$$

Напомним, что величина условно чистой продукции Z_j равна сумме амортизации, оплаты труда и чистого дохода j-й отрасли. Соотношение (6.1) охватывает систему из n уравнений, отражающих стоимостный состав продукции всех отраслей материальной сферы.

Во-вторых, рассматривая схему МОБ по строкам для каждой производящей отрасли, можно видеть, что валовая продукция той или иной отрасли равна сумме материальных затрат потребляющих ее продукцию отраслей и конечной продукции данной отрасли:

$$X_i = \sum_{j=1}^{n} x_{ij} + Y_i,\ i = \overline{1,n}. \tag{6.2}$$

Формула (6.2) описывает систему из n уравнений, которые называются уравнениями распределения продукции отраслей материального производства по направлениям использования.

Просуммируем по всем отраслям уравнения (6.1), в результате получим

$$\sum_{j=1}^{n} X_j = \sum_{j=1}^{n}\sum_{i=1}^{n} x_{ij} + \sum_{j=1}^{n} Z_j.$$

Аналогичное суммирование уравнений (6.2) дает:

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} + \sum_{i=1}^{n} Y_i.$$

Левые части обоих равенств равны, так как представляют собой весь валовой общественный продукт. Первые слагаемые правых частей этих равенств также равны, их величина равна итогу первого квадранта. Следовательно, должно соблюдаться соотношение

$$\sum_{j=1}^{n} Z_j = \sum_{i=1}^{n} Y_i. \tag{6.3}$$

Левая часть уравнения (6.3) есть сумма третьего квадранта, а правая часть — итог второго квадранта. В целом же это уравнение показывает, что в межотраслевом балансе соблюдается важнейший принцип единства материального и стоимостного состава национального дохода.

6.2. Экономико-математическая модель межотраслевого баланса

В параграфе 6.1 отмечено, что основу информационного обеспечения модели межотраслевого баланса составляет технологическая матрица, содержащая коэффициенты прямых материальных затрат на производство единицы продукции. Эта матрица является также основой экономико-математической модели межотраслевого баланса. Предполагается, что для производства единицы продукции в j-й отрасли требуется определенное количество затрат промежуточной продукции i-й отрасли, равное a_{ij}. Оно не зависит от объема производства в отрасли и является довольно стабильной величиной во времени. Величины a_{ij} называются *коэффициентами прямых материальных затрат* и рассчитываются следующим образом:

$$a_{ij} = \frac{x_{ij}}{X_j}, \; i, j = \overline{1, n}. \tag{6.4}$$

Определение 6.1. Коэффициент прямых материальных затрат показывает, какое количество продукции i-й отрасли необходимо, если учитывать только прямые затраты, для производства единицы продукции j-й отрасли.

С учетом формулы (6.4) систему уравнений баланса (6.2) можно переписать в виде

$$X_i = \sum_{j=1}^{n} a_{ij} X_j + Y_i, \; i = \overline{1, n}. \tag{6.5}$$

Если ввести в рассмотрение матрицу коэффициентов прямых материальных затрат $A = (a_{ij})$, вектор-столбец валовой продукции X и вектор-столбец конечной продукции Y:

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}; \; Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix},$$

то система уравнений (6.5) в матричной форме примет вид

$$X = AX + Y. \tag{6.6}$$

Система уравнений (6.5), или в матричной форме (6.6), называется *экономико-математической моделью межотраслевого баланса* (*моделью Леонтьева*, *моделью «затраты — выпуск»*). С помощью этой модели можно выполнять три варианта расчетов:

задав в модели величины валовой продукции каждой отрасли (X_i), можно определить объемы конечной продукции каждой отрасли (Y_i):

$$Y = (E - A)X; \tag{6.7}$$

задав величины конечной продукции всех отраслей (Y_i,), можно определить величины валовой продукции каждой отрасли (X_i):

$$X = (E - A)^{-1} Y \,. \qquad (6.8)$$

Для ряда отраслей, задав величины валовой продукции, а для всех остальных отраслей задав объемы конечной продукции, можно найти величины конечной продукции первых отраслей и объемы валовой продукции вторых; в этом варианте расчета удобнее пользоваться не матричной формой модели (6.6), а системой линейных уравнений (6.5). В формулах (6.7) и (6.8) E обозначает единичную матрицу n-го порядка, а $(E - A)^{-1}$ обозначает матрицу, обратную к матрице $(E - A)$. Если определитель матрицы $(E - A)$ не равен нулю, т.е. эта матрица невырожденная, то обратная к ней матрица существует. Обозначим эту обратную матрицу через B, тогда систему уравнений в матричной форме (6.8) можно записать в виде

$$X = BY. \qquad (6.8')$$

Элементы матрицы B будем обозначать через b_{ij}, тогда из матричного уравнения (6.8') для любой i-й отрасли можно получить следующее соотношение:

$$X_i = \sum_{j=1}^{n} b_{ij} Y_j, \quad i = \overline{1, n} \,. \qquad (6.9)$$

Из соотношений (6.9) следует, что валовая продукция выступает как взвешенная сумма величин конечной продукции, причем весами являются коэффициенты b_{ij}, которые показывают, сколько всего нужно произвести продукции i-й отрасли для выпуска в сферу конечного использования единицы продукции j-й отрасли. В отличие от коэффициентов прямых затрат a_{ij} коэффициенты b_{ij} называются *коэффициентами полных материальных затрат* и включают в себя как прямые, так и косвенные затраты всех порядков. Если прямые затраты отражают количество средств производства, израсходованных непосредственно при изготовлении данного продукта, то косвенные относятся к предшествующим стадиям производства и входят в производство продукта не прямо, а через другие (промежуточные) средства производства. Более детально этот вопрос рассматривается в параграфе 6.3.

Определение 6.2. Коэффициент полных материальных затрат b_{ij} показывает, какое количество продукции i-й отрасли нужно произвести, чтобы с учетом прямых и косвенных затрат этой продукции получить единицу конечной продукции j-й отрасли.

Коэффициенты полных материальных затрат можно применять, когда необходимо определить, как скажется на валовом выпуске некоторой отрасли предполагаемое изменение объемов конечной продукции всех отраслей:

$$\Delta X_i = \sum_{j=1}^{n} b_{ij} \Delta Y_j \,, \tag{6.10}$$

где ΔX_i и ΔY_j — изменения (приросты) величин валовой и конечной продукции соответственно.

6.3. Коэффициенты прямых и полных материальных затрат

Переходя к анализу модели межотраслевого баланса, необходимо прежде всего рассмотреть основные свойства матрицы коэффициентов прямых материальных затрат A. Коэффициенты прямых затрат по определению являются неотрицательными, следовательно, матрица A в целом может быть названа неотрицательной: $A \geq 0$. Так как процесс воспроизводства нельзя было бы осуществлять, если бы для собственного воспроизводства в отрасли затрачивалось большее количество продукта, чем создавалось, то очевидно, что диагональные элементы матрицы A меньше единицы: $a_{ii} < 1$.

Система уравнений межотраслевого баланса является отражением реальных экономических процессов, в которых содержательный смысл могут иметь лишь неотрицательные значения валовых выпусков; таким образом, вектор валовой продукции состоит из неотрицательных компонентов и называется неотрицательным: $X \geq 0$. Встает вопрос, при каких условиях экономическая система способна обеспечить положительный конечный выпуск по всем отраслям. Ответ на этот вопрос связан с понятием продуктивности матрицы коэффициентов прямых материальных затрат.

Будем называть неотрицательную матрицу A *продуктивной*, если существует такой неотрицательный вектор $X \geq 0$, что

$$X > AX. \tag{6.11}$$

Очевидно, что условие (6.11) означает существование положительного вектора конечной продукции $Y > 0$ для модели межотраслевого баланса (6.6).

Для того чтобы матрица коэффициентов прямых материальных затрат A была продуктивной, необходимо и достаточно чтобы выполнялось одно из перечисленных ниже условий:

1) матрица $(E - A)$ неотрицательно обратима, т.е. существует обратная матрица $(E - A)^{-1} \geq 0$;

2) матричный ряд $E + A + A^2 + A^3 + \ldots = \sum_{k=0}^{\infty} A^k$ сходится, причем его сумма равна обратной матрице $(E - A)^{-1}$;

3) наибольшее по модулю собственное значение λ матрицы A, т.е. решение характеристического уравнения $\left|\lambda E - A\right| = 0$, строго меньше единицы;

4) все главные миноры матрицы $(E - A)$, т.е. определители матриц, образованные элементами первых строк и первых столбцов этой матрицы, порядка от 1 до n, положительны.

Более простым, но только достаточным признаком продуктивности матрицы A является ограничение на величину ее *нормы*, т.е. на величину наибольшей из сумм элементов матрицы A в каждом столбце. Если норма матрицы A строго меньше единицы, то эта матрица продуктивна; повторим, что данное условие является только достаточным и матрица А может оказаться продуктивной и в случае, когда ее норма больше единицы.

Наибольший по модулю корень характеристического уравнения, приведенного в условии 3) продуктивности матрицы A (обозначим его через λ^*), может служить оценкой общего уровня коэффициентов прямых материальных затрат, а, следовательно, величина $1 - \lambda^*$ характеризует остаток после затрат, т.е. продуктивность. Чем больше $1 - \lambda^*$, тем больше возможности достижения других целей, кроме текущего производственного потребления. Другими словами, чем выше общий уровень коэффициентов матрицы A, тем больше наибольшее по модулю собственное значение λ^* и ниже уровень продуктивности, и наоборот, чем ниже общий уровень коэффициентов матрицы A, тем меньше наибольшее по модулю собственное значение и выше продуктивность.

Перейдем к анализу матрицы коэффициентов полных материальных затрат, т.е. матрицы $B = (E - A)^{-1}$. Согласно определению 6.2 из предыдущего параграфа коэффициент этой матрицы показывает, сколько всего нужно произвести продукции i-й отрасли, чтобы получить единицу конечной продукции j-й отрасли.

Дадим другое определение коэффициента полных материальных затрат исходя из того, что кроме прямых затрат существуют косвенные затраты той или иной продукции при производстве продукции данной отрасли. Рассмотрим в качестве примера формирование затрат электроэнергии на выпуск стального проката, при этом ограничимся технологической цепочкой «руда — чугун — сталь — прокат». Затраты электроэнергии при получении проката из стали будут называться прямыми затратами, те же затраты при получении стали из чугуна будут называться косвенными затратами 1-го порядка, а затраты электроэнергии при получении чугуна из руды будут называться косвенными затратами электроэнергии на выпуск стального проката 2-го порядка и т.д. В связи со сказанным выше имеет место следующее определение.

Определение 6.3. Коэффициентом полных материальных затрат c_{ij} называется сумма прямых затрат и косвенных затрат продукции i-й отрасли для производства единицы продукции j-й отрасли через все промежуточные продукты на всех предшествующих стадиях производства. Если коэффициент косвенных материальных затрат k-го порядка обозначить через $a_{ij}^{(k)}$, то имеет место формула

$$c_{ij} = a_{ij} + a_{ij}^{(1)} + a_{ij}^{(2)} + \ldots + a_{ij}^{(k)} + \ldots, \qquad (6.12)$$

а если ввести в рассмотрение матрицу коэффициентов полных материальных затрат $C = (c_{ij})$ и матрицы коэффициентов косвенных материальных затрат различных порядков $A^{(k)} = \left(a_{ij}^{(k)}\right)$, то поэлементную формулу (6.12) можно записать в более общем матричном виде:

$$C = A + A^{(1)} + A^{(2)} + \ldots + A^{(k)} + \ldots \qquad (6.13)$$

Исходя из содержательного смысла коэффициентов косвенных материальных затрат можно записать ряд матричных соотношений:

$$A^{(1)} = AA = A^2; \quad A^{(2)} = AA^{(1)} = AA^{(2)} = A^{(3)};$$
$$A^{(k)} = AA^{(k-1)} = AA^k = A^{k+1},$$

с использованием которых матричная формула (6.13) может быть переписана в следующем виде:

$$C = A + A^2 + A^3 + \ldots + A^{k+1} + \ldots = \sum_{k=1}^{\infty} A^k. \qquad (6.14)$$

Если матрица коэффициентов прямых материальных затрат A является продуктивной, то из условия 2) продуктивности существует матрица $B = (E - A)^{-1}$, являющаяся суммой сходящегося матричного ряда:

$$B = (E - A)^{-1} = E + A + A^2 + A^3 + \ldots = \sum_{k=1}^{\infty} A^k. \qquad (6.15)$$

Из сопоставления соотношений (6.14) и (6.15) устанавливается следующая связь между двумя матрицами коэффициентов полных материальных затрат:

$$B = E + C,$$

или, в поэлементной записи:

$$b_{ij=}\begin{cases} c_{ij}, & \text{если} \quad i \neq j, \\ 1 + c_{ij}, & \text{если} \quad i = j. \end{cases}$$

Данная связь определяет экономический смысл различия между коэффициентами матриц B и C: в отличие от коэффициентов матрицы C, учитывающих только затраты на производство продукции, коэффициенты матрицы B включают в себя кроме затрат также саму единицу конечной продукции, которая выходит за сферу производства.

Перейдем теперь к вычислительным аспектам решения задач на основе модели межотраслевого баланса. Основной объем расчетов по этой модели связан с вычислением матрицы коэффициентов полных материальных затрат B. Если матрица коэффициентов прямых материальных затрат A задана и является продуктивной, матрицу B можно находить либо по формулам обращения матриц, рассматриваемым в курсе матричной алгебры (некоторые из этих формул рассмотрены в гл. 2), либо приближенным способом, используя разложение в матричный ряд (6.15).

Рассмотрим *первый способ* нахождения матрицы B. Находят матрицу $(E - A)$, а затем, применяя один из прямых методов обращения невырожденных матриц, вычисляют матрицу $(E - A)^{-1}$. Одним из наиболее употребительных методов обращения матриц является метод Жордана. Часто применяется также метод, основанный на применении формулы матричной алгебры

$$B = (E - A)^{-1} = \frac{\overline{(E - A)}}{\left|E - A\right|}, \tag{6.16}$$

где в числителе матрица, присоединенная к матрице $(E - A)$, элементы которой представляют собой алгебраические дополнения для элементов транспонированной матрицы $(E - A)'$, а в знаменателе — определитель матрицы $(E - A)$. Алгебраические дополнения в свою очередь для элемента с индексами i и j получаются умножением множителя $(-1)^{i+j}$ на минор, получаемый после вычеркивания из матрицы i-й строки и j-го столбца.

При *втором способе* вычисления матрицы коэффициентов полных материальных затрат используется формула (6.15). Обязательным условием корректности этих расчетов является условие продуктивности матрицы A, и при расчетах ограничиваются учетом косвенных материальных затрат до некоторого порядка включительно, например до 2-го, 3-го порядков. В этом способе используется процедура умножения квадратных матриц с их последующим сложением, и коэффициенты полных материальных затрат получаются с известным приближением (с недостатком).

Пример 6.1. Для трехотраслевой экономической системы заданы матрица коэффициентов прямых материальных затрат и вектор конечной продукции:

$$A = \begin{pmatrix} 0{,}3 & 0{,}1 & 0{,}4 \\ 0{,}2 & 0{,}5 & 0{,}0 \\ 0{,}3 & 0{,}1 & 0{,}2 \end{pmatrix}; \quad Y = \begin{pmatrix} 200 \\ 100 \\ 300 \end{pmatrix}.$$

Найти коэффициенты полных материальных затрат и вектор валовой продукции, заполнить схему межотраслевого материального баланса.

1. Определим матрицу коэффициентов полных материальных затрат по второму (приближенному) способу, учитывая косвенные материальные затраты до 2-го порядка включительно. Запишем матрицу коэффициентов косвенных затрат 1-го порядка:

$$A^{(1)} = A^2 = \begin{pmatrix} 0{,}3 & 0{,}1 & 0{,}4 \\ 0{,}2 & 0{,}5 & 0{,}0 \\ 0{,}3 & 0{,}1 & 0{,}2 \end{pmatrix} \cdot \begin{pmatrix} 0{,}3 & 0{,}1 & 0{,}4 \\ 0{,}2 & 0{,}5 & 0{,}0 \\ 0{,}3 & 0{,}1 & 0{,}2 \end{pmatrix} = \begin{pmatrix} 0{,}23 & 0{,}12 & 0{,}20 \\ 0{,}16 & 0{,}27 & 0{,}08 \\ 0{,}17 & 0{,}10 & 0{,}16 \end{pmatrix},$$

матрицу коэффициентов косвенных затрат 2-го порядка:

$$A^{(2)} = AA^{(1)} = \begin{pmatrix} 0,3 & 0,1 & 0,4 \\ 0,2 & 0,5 & 0,0 \\ 0,3 & 0,1 & 0,2 \end{pmatrix} \cdot \begin{pmatrix} 0,23 & 0,12 & 0,20 \\ 0,16 & 0,27 & 0,08 \\ 0,17 & 0,10 & 0,16 \end{pmatrix} =$$

$$= \begin{pmatrix} 0,153 & 0,103 & 0,132 \\ 0,126 & 0,159 & 0,080 \\ 0,119 & 0,083 & 0,100 \end{pmatrix}.$$

Таким образом, матрица коэффициентов полных материальных затрат приближенно равна

$$B \approx E + A + A^2 + A^3 = \begin{pmatrix} 1,683 & 0,323 & 0,732 \\ 0,486 & 1,929 & 0,160 \\ 0,589 & 0,283 & 1,460 \end{pmatrix}.$$

2. Определим матрицу коэффициентов полных материальных затрат с помощью формул обращения невырожденных матриц (первый способ):

а) находим матрицу $(E - A)$:

$$(E - A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0,3 & 0,1 & 0,4 \\ 0,2 & 0,5 & 0,0 \\ 0,3 & 0,1 & 0,2 \end{pmatrix} = \begin{pmatrix} 0,7 & -0,1 & -0,4 \\ -0,2 & 0,5 & 0,0 \\ -0,3 & -0,1 & 0,8 \end{pmatrix};$$

б) вычисляем определитель этой матрицы:

$$|E - A| = \begin{vmatrix} 0,7 & -0,1 & -0,4 \\ -0,2 & 0,5 & 0,0 \\ -0,3 & -0,1 & 0,8 \end{vmatrix} = 0,196;$$

в) транспонируем матрицу $(E - A)$:

$$(E - A)' = \begin{pmatrix} 0,7 & -0,2 & -0,3 \\ -0,1 & 0,5 & -0,1 \\ -0,4 & 0,0 & 0,8 \end{pmatrix};$$

г) находим алгебраические дополнения для элементов матрицы

$$(E - A)' : A_{11} = (-1)^2 \begin{vmatrix} 0,5 & -0,1 \\ 0,0 & 0,8 \end{vmatrix} = 0,40; \quad A_{12} = (-1)^3 \begin{vmatrix} -0,1 & -0,1 \\ -0,4 & 0,8 \end{vmatrix} = 0,12;$$

$$A_{13}=(-1)^4\begin{vmatrix}-0,1 & 0,5\\ -0,4 & 0,0\end{vmatrix}=0,20;\quad A_{21}=(-1)^3\begin{vmatrix}-0,2 & -0,3\\ 0,0 & 0,8\end{vmatrix}=0,16;$$

$$A_{22}=(-1)^4\begin{vmatrix}0,7 & -0,3\\ -0,4 & 0,8\end{vmatrix}=0,44;\quad A_{23}=(-1)^5\begin{vmatrix}0,7 & -0,2\\ -0,4 & 0,0\end{vmatrix}=0,08;$$

$$A_{31}=(-1)^4\begin{vmatrix}-0,2 & -0,3\\ 0,5 & -0,1\end{vmatrix}=0,17;\quad A_{32}=(-1)^5\begin{vmatrix}0,7 & -0,3\\ -0,1 & -0,1\end{vmatrix}=0,10;$$

$$A_{33}=(-1)^6\begin{vmatrix}0,7 & -0,2\\ -0,1 & 0,5\end{vmatrix}=0,33,$$

таким образом, присоединенная к матрице $(E-A)$ матрица имеет вид

$$(\widetilde{E-A})=\begin{pmatrix}0,40 & 0,12 & 0,20\\ 0,16 & 0,44 & 0,08\\ 0,17 & 0,10 & 0,33\end{pmatrix};$$

д) используя формулу (6.16), находим матрицу коэффициентов полных материальных затрат:

$$B=(E-A)^{-1}=\begin{pmatrix}2,041 & 0,612 & 0,020\\ 0,816 & 2,245 & 0,408\\ 0,867 & 0,510 & 1,684\end{pmatrix}.$$

Как отмечено выше, элементы матрицы B, рассчитанные по точным формулам обращения матриц, больше соответствующих элементов матрицы, рассчитанной по второму приближенному способу без учета косвенных материальных затрат порядка выше 2-го.

3. Найдем величины валовой продукции трех отраслей (вектор X), используя формулу (6.8'):

$$X=BY=\begin{pmatrix}2,041 & 0,612 & 1,020\\ 0,816 & 2,245 & 0,408\\ 0,867 & 0,510 & 1,684\end{pmatrix}\cdot\begin{pmatrix}200\\ 100\\ 300\end{pmatrix}=\begin{pmatrix}775,3\\ 510,1\\ 729,6\end{pmatrix}.$$

4. Для определения элементов первого квадранта материального межотраслевого баланса воспользуемся формулой, вытекающей из формулы (6.4): $x_{ij}=a_{ij}X_j$. Из этой формулы следует, что для получения первого столбца первого квадранта нужно элементы первого столбца заданной матрицы A умножить на величину X_1 = 775,3; элементы второго столбца матрицы A умножить на X_2 = 510,1; элементы третьего столбца матрицы A умножить на X_3 = 729,6.

Составляющие третьего квадранта (условно чистая продукция) находятся с учетом формулы (6.1) как разность между объемами валовой продукции и суммами элементов соответствующих столбцов найденного первого квадранта.

Четвертый квадрант в нашем примере состоит из одного показателя и служит, в частности, для контроля правильности расчета: сумма элементов второго квадранта должна в стоимостном материальном балансе совпадать с суммой элементов третьего квадранта. Результаты расчета представлены в табл. 6.2; незначительные расхождения по строкам таблицы объясняются погрешностью из-за округления чисел.

Таблица 6.2

Межотраслевой баланс производства и распределения продукции

Производящие отрасли	Потребляющие отрасли			Конечная продукция	Валовая продукция
	1	2	3		
1 2 3	232,6 155,1 232,6	51,0 255,0 51,0	291,8 0,0 145,9	200,0 100,0 300,0	775,3 510,1 729,6
Условно чистая продукция	155,0	153,1	291,9	600,0	
Валовая продукция	775,3	510,1	729,6		2015,0

6.4. Межотраслевые балансовые модели в анализе экономических показателей

Различные модификации рассмотренной выше модели межотраслевого баланса производства и распределения продукции в народном хозяйстве позволяют расширить круг показателей, охватываемых моделью. Рассмотрим применение межотраслевого балансового метода для анализа таких важных экономических показателей, как труд, фонды и цены.

К числу важнейших аналитических возможностей данного метода относится определение прямых и полных затрат труда на единицу продукции и разработка на этой основе балансовых продуктово-трудовых моделей, исходной моделью при этом служит отчетный межпродуктовый баланс в натуральном выражении. В этом балансе по строкам представлено распределение каждого отдельного продукта на производство других продуктов и конечное потребление (первый и второй квадранты схемы межотраслевого баланса). Отдельной строкой дается распределение затрат живого труда в производстве всех видов продукции; предполагается, что трудовые затраты выражены в единицах труда одинаковой степени сложности.

Обозначим затраты живого труда в производстве j-го продукта через L_j а объем производства этого продукта (валовой выпуск), как и раньше, через X_j. Тогда прямые затраты труда на единицу j-го вида продукции (*коэффициент прямой трудоемкости*) можно задать следующей формулой:

$$t_j = \frac{L_j}{X_j}, \quad j = \overline{1, n}. \tag{6.17}$$

Введем понятие полных затрат труда как суммы прямых затрат живого труда и затрат овеществленного труда, перенесенных на продукт через израсходованные средства производства. Если обозначить величину полных затрат труда на единицу продукции j-го вида через T_j, то произведения вида $a_{ij}T_i$ отражают затраты овеществленного труда, перенесенного на единицу j-го продукта через i-е средство производства; при этом предполагается, что коэффициенты прямых материальных затрат a_{ij} выражены в натуральных единицах. Тогда полные трудовые затраты на единицу j-го вида продукции (*коэффициент полной трудоемкости*) будут равны

$$T_j = \sum_{i=1}^{n} a_{ij}T_i + t_j, \quad j = \overline{1, n}. \tag{6.18}$$

Введем в рассмотрение вектор-строку коэффициентов прямой трудоемкости $t = (t_1, t_2, ..., t_n)$ и вектор-строку коэффициентов полной трудоемкости $T = (T_1, T_2, ..., T_n)$.

Тогда с использованием уже рассматриваемой выше матрицы коэффициентов прямых материальных затрат A (в натуральном выражении) систему уравнений (6.18) можно переписать в матричном виде:

$$T = TA + t. \tag{6.19}$$

Произведя очевидные матричные преобразования с использованием единичной матрицы E

$$T - TA = TE - TA = T(E - A) = t,$$

получим следующее соотношение для вектора коэффициентов полной трудоемкости:

$$T = t(E - A)^{-1}. \tag{6.20}$$

Матрица $(E - A)^{-1}$ нам уже знакома, это матрица B коэффициентов полных материальных затрат, так что последнее равенство можно переписать в виде

$$T = tB. \tag{6.20'}$$

Обозначим через L величину совокупных затрат живого труда по всем видам продукции, которая с учетом формулы (6.17) будет равна

$$L = \sum_{j=1}^{n} L_j = \sum_{j=1}^{n} t_j X_j = tX. \tag{6.21}$$

Используя соотношения (6.21), (6.8') и (6.20'), приходим к следующему равенству:

$$tX = TY, \tag{6.22}$$

здесь t и T — вектор-строки коэффициентов прямой и полной трудоемкости, а X и Y — вектор-столбцы валовой и конечной продукции соответственно.

Соотношение (6.22) представляет собой основное балансовое равенство в теории межотраслевого баланса труда. В данном случае его конкретное экономическое содержание заключается в том, что стоимость конечной продукции, оцененной по полным затратам труда, равна совокупным затратам живого труда. Сопоставляя потребительский эффект различных взаимозаменяемых продуктов с полными трудовыми затратами на их выпуск, можно судить о сравнительной эффективности их производства. С помощью показателей полной трудоемкости более полно и точно, чем при использовании существующих стоимостных показателей, выявляется структура затрат на выпуск различных видов продукции и прежде всего соотношение между затратами живого и овеществленного труда.

На основе коэффициентов прямой и полной трудоемкости могут быть разработаны межотраслевые и межпродуктовые балансы затрат труда и использования трудовых ресурсов. Схематически эти балансы строятся по общему типу матричных моделей, однако все показатели в них (межотраслевые связи, конечный продукт, условно чистая продукция и др.) выражены в трудовых измерителях.

Пример 6.2. Пусть в дополнение к исходным данным примера 6.1 из параграфа 6.3 заданы затраты живого труда (трудовые ресурсы) в трех отраслях: $L_1 = 1160$, $L_2 = 460$, $L_3 = 875$ в некоторых единицах измерения трудовых затрат. Требуется определить коэффициенты прямой и полной трудоемкости и составить межотраслевой баланс затрат труда.

1. Воспользовавшись формулой (6.17) и результатами примера 1, находим коэффициенты прямой трудоемкости:

$$t_1 = \frac{1160}{775{,}3} = 1{,}5; \quad t_2 = \frac{460}{510{,}1} = 0{,}9; \quad t_3 = \frac{875}{729{,}6} = 1{,}2.$$

2. По формуле (6.20'), в которой в качестве матрицы B берется матрица коэффициента полных материальных затрат, найденная в примере 6.1, находим коэффициенты полной трудоемкости:

$$T = (1{,}5; 0{,}9; 1{,}2) \cdot \begin{pmatrix} 2{,}041 & 0{,}612 & 1{,}020 \\ 0{,}816 & 2{,}245 & 0{,}408 \\ 0{,}867 & 0{,}510 & 1{,}684 \end{pmatrix} = (4{,}84; 3{,}55; 3{,}92).$$

3. Умножая первую, вторую и третью строки первого и второго квадрантов межотраслевого материального баланса, построенного в примере 1, на соответствующие коэффициенты прямой трудоемкости, получаем схему межотраслевого баланса затрат труда (в трудовых измерителях) (табл. 6.3).

Таблица 6.3

Межотраслевой баланс затрат труда

Производящие отрасли	Потребляющие отрасли. Межотраслевые затраты овеществленного труда			Затраты труда на конечную продукцию	Затраты труда в отраслях (трудовые ресурсы)
	1	2	3		
1	348,9	76,5	437,7	300,0	1163,1
2	139,6	229,5	0,0	90,0	459,1
3	279,1	61,2	175,1	360,0	875,4

Незначительные расхождения между данными таблицы и исходными данными вызваны погрешностями округления при вычислениях.

Развитие основной модели межотраслевого баланса достигается также путем включения в нее показателей фондоемкости продукции. В простейшем случае модель дополняется отдельной строкой, в которой указаны в стоимостном выражении объемы производственных фондов Φ_j, занятые в каждой j-й отрасли. На основании этих данных и объемов валовой продукции всех отраслей определяются *коэффициенты прямой фондоемкости продукции* j-й отрасли:

$$f_j = \frac{\Phi_j}{X_j}, \quad j = \overline{1, n}. \tag{6.23}$$

Коэффициент прямой фондоемкости показывает величину производственных фондов, непосредственно занятых в производстве данной отрасли, в расчете на единицу ее валовой продукции. В отличие от этого показателя *коэффициент полной фондоемкости* F_j отражает объем фондов, необходимых во всех отраслях для выпуска единицы конечной продукции j-й отрасли. Если a_{ij} — коэффициент прямых материальных затрат, то для коэффициента полной фондоемкости справедливо равенство, аналогичное равенству (6.18) для коэффициента полной трудоемкости:

$$F_j = \sum_{i=1}^{n} a_{ij} F_i + f_j, \quad j = \overline{1, n}. \tag{6.24}$$

Если ввести в рассмотрение вектор-строку коэффициентов прямой фондоемкости $f = (f_1, f_2, ..., f_n)$ и вектор-строку коэффициентов полной фондоемкости $F = (F_1, F_2, ..., F_n)$, то систему уравнений (6.24) можно переписать в матричной форме:

$$F = FA + f, \tag{6.25}$$

откуда с помощью преобразований, аналогичных применяемым выше для коэффициентов трудоемкости, можно получить матричное соотношение

$$F = fB, \tag{6.26}$$

где $B = (E - A)^{-1}$ — матрица коэффициентов полных материальных затрат.

Для более глубокого анализа необходимо дифференцировать фонды на основные и оборотные, а в пределах основных — на здания, сооружения, производственное оборудование, транспортные средства и т.д.

Пусть в целом все производственные фонды разделены на m групп. Тогда характеристика занятых в народном хозяйстве фондов задается матрицей показателей Φ_{kj}, отражающих объем фондов k-й группы, занятых в j-й отрасли:

$$(\Phi_{kj}) = \begin{pmatrix} \Phi_{11} & \Phi_{12} & \dots & \Phi_{1n} \\ \Phi_{21} & \Phi_{22} & \dots & \Phi_{2n} \\ \dots & \dots & \dots & \dots \\ \Phi_{m1} & \Phi_{m2} & \dots & \Phi_{mn} \end{pmatrix}.$$

Коэффициенты прямой фондоемкости также образуют матрицу размерности $m \times n$, элементы которой определяют величину производственных фондов k-й группы, непосредственно используемых при производстве единицы продукции j-й отрасли:

$$f_{kj} = \frac{\Phi_{kj}}{X_j}.$$

Для каждой j-й отрасли могут быть вычислены коэффициенты полной фондоемкости F_{kj}, отражающие полную потребность в фондах k-й группы для выпуска единицы конечной продукции этой отрасли:

$$F_{kj} = \sum_{i=1}^{n} a_{ij} F_{ki} + f_{kj}, \quad k = \overline{1, m}; \quad j = \overline{1, n}.$$

Решение системы данных уравнений позволяет представить коэффициенты полной фондоемкости по каждой из m групп фондов как функцию коэффициентов прямой фондоемкости:

$$F_{kj} = \sum_{i=1}^{n} b_{ij} f_{ki}, \quad k = \overline{1, m}; \quad j = \overline{1, n}.$$

В этих формулах величины a_{ij} и b_{ij} — уже известные коэффициенты прямых и полных материальных затрат.

Коэффициенты фондоемкости в межотраслевом балансе позволяют увязать планируемый выпуск продукции с имеющимися производственными мощностями. Так, потребность в функционирующих фондах k-й группы для достижения заданного объема материального производства X_j по всем отраслям задается формулой

$$\Phi_k = \sum_{j=1}^{n} f_{kj} X_j, \quad k = \overline{1, m}.$$

6.5. Динамическая межотраслевая балансовая модель

Рассмотренные выше межотраслевые балансовые модели являются *статическими*, т.е. такими, в которых все зависимости отнесены к одному моменту времени. Эти модели могут разрабатываться лишь для отдельно взятых периодов, причем в рамках данных моделей не устанавливается связь с предшествующими или последующими периодами. Народнохозяйственная динамика отображается, таким образом, рядом независимо рассчитанных моделей, что очевидно вносит определенные упрощения и сужает возможности анализа.

К числу таких упрощений, прежде всего, следует отнести то, что в статических межотраслевых моделях не анализируются распределение, использование и производственная эффективность капитальных вложений. Капиталовложения вынесены из сферы производства в сферу конечного использования вместе с предметами потребления и непроизводственными затратами, т.е. включены в конечный продукт.

В отличие от статических *динамические* модели призваны отразить не состояние, а процесс развития экономики, установить непосредственную взаимосвязь между предыдущими и последующими этапами развития и тем самым приблизить анализ на основе экономико-математической модели к реальным условиям развития экономической системы.

В рассматриваемой здесь динамической модели, являющейся развитием статической межотраслевой модели, производственные капитальные вложения выделяются из состава конечной продукции, исследуются их структура и влияние на рост объема производства. В основе построения модели в виде динамической системы уравнений лежит математическая зависимость между величиной капитальных вложений и приростом продукции. Решение системы, как и в случае статической модели, приводит к определению уровней производства, но в динамическом варианте в отличие от статического эти искомые уровни зависят от объемов производства в предшествующих периодах.

Принципиальная схема первых двух квадрантов динамического межотраслевого баланса приведена в табл. 6.4.

Модель содержит две матрицы межотраслевых потоков. Матрица текущих производственных затрат с элементами x_{ij} совпадает с соответствующей матрицей статического ба-

Таблица 6.4

Принципиальная схема динамического баланса

Производящие отрасли	Потребляющие отрасли								Конечный продукт	Валовой продукт
	Межотраслевые потоки текущих затрат				Межотраслевые потоки капитальных вложений					
	1	2	...	n	1	2	...	n		
1	x_{11}	x_{12}	...	x_{1n}	$\Delta\Phi_{11}$	$\Delta\Phi_{12}$		$\Delta\Phi_{1n}$	Y'_1	X_1
2	x_{21}	x_{22}	...	x_{2n}	$\Delta\Phi_{21}$	$\Delta\Phi_{22}$		$\Delta\Phi_{2n}$	Y'_2	X_2
...	...	...	...	...	...	...	...	...	...	...
n	x_{n1}	x_{n2}	...	x_{nn}	$\Delta\Phi_{n1}$	$\Delta\Phi_{n2}$		$\Delta\Phi_{nn}$	Y'_n	X_n

ланса. Элементы второй матрицы $\Delta\Phi_{ij}$ показывают, какое количество продукции i-й отрасли направлено в текущем периоде в j-ю отрасль в качестве производственных капитальных вложений в ее основные фонды. Материально это выражается в приросте в потребляющих отраслях производственного оборудования, сооружений, производственных площадей, транспортных средств и др.

В статическом балансе потоки капиталовложений не дифференцируются по отраслям-потребителям и отражаются общей величиной в составе конечной продукции Y_i каждой i-й отрасли. В динамической схеме конечный продукт Y_i включает продукцию i-й отрасли, идущую в личное и общественное потребление, накопление непроизводственной сферы, прирост оборотных фондов, незавершенного строительства, на экспорт. Таким образом, сумма потоков капиталовложений и конечного продукта динамической модели равна конечной продукции статического баланса:

$$\sum_{j=1}^{n} \Delta\Phi_{ij} + Y_i' = Y_i,$$

поэтому уравнение распределения продукции вида (6.2) в динамическом балансе преобразуется в следующее:

$$X_i = \sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} \Delta\Phi_{ij} + Y_i', \quad i = \overline{1, n}. \tag{6.27}$$

Межотраслевые потоки текущих затрат можно выразить, как и в статической модели, через валовую продукцию отраслей с помощью коэффициентов прямых материальных затрат:

$$x_{ij} = a_{ij} X_j.$$

В отличие от потоков текущих затрат межотраслевые потоки капитальных вложений связаны не со всей величиной выпуска продукции, а обусловливают прирост продукции; причем в рассматриваемой модели предполагается, что прирост продукции текущего периода обусловлен вложениями, произведенными в этом же периоде. Если текущий период обозначить через t, то прирост продукции ΔX_j равен разности абсолютных уровней производства в период t и в предшествующий $(t-1)$-й период:

$$\Delta X_j = X_j^{(t)} - X_j^{(t-1)}.$$

Полагая, что прирост продукции пропорционален приросту производственных фондов, можно записать:

$$\Delta \Phi_{ij} = \varphi_{ij} \Delta X_j, \quad i, j = \overline{1, n}. \tag{6.28}$$

Рассмотрим в равенстве (6.28) коэффициенты пропорциональности φ_{ij}. Поскольку

$$\varphi_{ij} = \frac{\Delta \Phi_{ij}}{\Delta X_j},$$

то экономический смысл этих коэффициентов заключается в том, что они показывают, какое количество продукции i-й отрасли должно быть вложено в j-ю отрасль для увеличения производственной мощности j-й отрасли на единицу продукции. Предполагается, что производственные мощности используются полностью и прирост продукции равен приросту мощности. Коэффициенты φ_{ij} называются *коэффициентами вложений*, или *коэффициентами приростной фондоемкости.*

С помощью коэффициентов прямых материальных затрат и коэффициентов вложений φ_{ij} систему уравнений (6.27) можно представить в следующем виде:

$$X_i = \sum_{j=1}^{n} a_{ij} X_j + \sum_{j=1}^{n} \varphi_{ij} \Delta X_j + Y_i', \quad i = \overline{1, n}. \tag{6.29}$$

Система (6.29) представляет собой систему линейных разностных уравнений первого порядка. Ее можно привести к обычной системе линейных уравнений, если учесть, что все объемы валовой и конечной продукции относятся к некоторому периоду t, а прирост валовой продукции определен в сравнении с $(t-1)$-м периодом:

$$X_i^{(t)} = \sum_{j=1}^{n} a_{ij} X_j^{(t)} + \sum_{j=1}^{n} \phi_{ij} \left(X_j^{(t)} - X_j^{(t-1)} \right) + Y_i'^{(t)}.$$

Отсюда можно записать следующие соотношения:

$$X_i^{(t)} = \sum_{j=1}^{n} (a_{ij} + \varphi_{ij}) X_j^{(t)} - \sum_{j=1}^{n} \varphi_{ij} X_j^{(t-1)} + Y_i'^{(t)}, \quad i = \overline{1, n}. \qquad (6.30)$$

Пусть нам известны уровни валовой продукции всех отраслей в предыдущем периоде (величины $X_j^{(t-1)}$) и конечный продукт отраслей в t-м периоде. Тогда очевидно, что соотношения (6.30) представляют собой систему n линейных уравнений с n неизвестными уровнями производства t-го периода. Таким образом, решение динамической системы линейных уравнений позволяет определить выпуск продукции в последующем периоде в зависимости от уровня, достигнутого в предыдущем периоде. Связь между периодами устанавливается через коэффициенты вложений φ_{ij}, характеризующие фондоемкость единицы прироста продукции.

Переходя от дискретного анализа к непрерывному, вместо (6.27) будем иметь:

$$X_i = \sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} \frac{d\Phi_{ij}}{dt} + Y_i'.$$

Выражение (6.28) в пределе дает:

$$\frac{d\Phi_{ij}}{dt} = \varphi_{ij} \frac{dX_j}{dt}.$$

Окончательно для случая непрерывных изменений получим следующую систему соотношений:

$$X_i = \sum_{j=1}^{n} a_{ij} X_j + \sum_{j=1}^{n} \varphi_{ij} \frac{dX_j}{dt} + Y_i', \quad i = \overline{1, n}. \qquad (6.31)$$

Соотношения (6.31) представляют собой систему n линейных дифференциальных уравнений первого порядка с постоянными коэффициентами. Для ее решения, помимо матриц коэффициентов прямых материальных текущих затрат и коэффициентов капитальных затрат (вложений), необходимо знать уровни валового выпуска в начальный момент времени $t = 0$ и закон изменения величины конечного продукта, т.е. вид функций $Y_i'(t)$. На основе этих данных путем решения получившейся задачи Коши для системы дифференциальных уравнений (6.31) можно найти уровни валового выпуска теоретически для любого момента времени. Практически же более или менее достоверное описание валовых и конечных выпусков как функций времени может быть получено лишь для относительно небольших промежутков времени.

В динамической модели особую роль играют коэффициенты приростной фондоемкости φ_{ij}. Они образуют квадратную матрицу n-го порядка

$$(\varphi_{ij}) = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \dots & \varphi_{1n} \\ \varphi_{21} & \varphi_{22} & \dots & \varphi_{2n} \\ \dots & \dots & \dots & \dots \\ \varphi_{n1} & \varphi_{n2} & \dots & \varphi_{nn} \end{pmatrix},$$

каждый столбец которой характеризует для соответствующей j-й отрасли величину и структуру фондов, необходимых для увеличения на единицу ее производственной мощности (выпуска продукции). Матрица коэффициентов приростной фондоемкости дает значительный материал для экономического анализа и планирования капитальных вложений.

Коэффициенты приростной фондоемкости φ_{ij} определенным образом связаны с валовыми коэффициентами прямой фондоемкости продукции f_{kj}, рассмотренными в предыдущем параграфе. Коэффициенты f_{kj} показывают, сколько всего фондов данного вида приходится на единицу валового выпуска продукции, а коэффициенты φ_{ij} отражают прирост фондов на единицу прироста продукции. Если бы технический прогресс в отраслях производства отсутствовал, то на единицу прироста продукции потребовалось бы столько же новых фондов, сколько их уже занято на единицу выпускаемой продукции, т.е. коэффициенты приростной фондоемкости и валовой прямой фондоемкости были бы равны между собой. Так как новые капитальные вложения производятся на новом более высоком техническом уровне по сравнению с объемом и структурой действующих фондов, то на практике коэффициенты приростной фондоемкости и коэффициенты прямой фондоемкости различаются по величине. Однако между этими двумя группами коэффициентов существует вполне определенная связь, и это используется при разработке динамических моделей, особенно в связи с тем, что достоверные данные о фондоемкости продукции получить легче, чем непосредственно рассчитать коэффициенты вложений.

Кроме коэффициентов прямой фондоемкости коэффициенты вложений связаны с другими показателями, например с соответствующими коэффициентами текущих затрат, отражающими износ основных фондов и равными амортизации, приходящейся на единицу продукции.

В рассмотренной динамической модели межотраслевого баланса предполагается, что прирост продукции текущего периода обусловлен капиталовложениями, произведенными в этом же периоде. Для сравнительно коротких периодов это предположение может оказаться нереальным, так как существуют известные, иногда довольно значительные отставания во времени (так называемые *временные лаги*) между вложением средств в производственные фонды и приростом выпуска продукции. Модели, так или иначе учитывающие лаг капитальных вложений, образуют особую группу динамических моделей межотраслевого баланса. Из теоретических моделей данного типа следует назвать, прежде всего, линейную динамическую межотраслевую модель Леонтьева, в которой капитальные вложения представлены в виде так называемого инвестиционного блока в форме Леонтьева. Математическим обобщением этой и ряда других динамических моделей является динамическая модель в матричной форме Неймана, основанная на математической теории равномерного пропорционального роста экономики (так называемая *магистральная теория*).

Вопросы и задания

1. В чем суть балансового метода исследования социально-экономических систем?

2. Поясните принципиальную схему межотраслевого баланса и раскройте экономическое содержание ее разделов.

3. Опишите экономико-математическую модель статического межотраслевого баланса и поясните экономический смысл входящих в нее элементов.

4. Дайте определение коэффициентов прямых и полных материальных затрат и укажите способы их вычисления.

5. Поясните понятие продуктивности матрицы коэффициентов прямых материальных затрат.

6. Раскройте экономический смысл коэффициентов прямой и полной трудоемкости и дайте описание экономико-математической модели межотраслевого баланса затрат труда.

7. В чем заключается экономическое содержание коэффициентов прямой и полной фондоемкости? Поясните порядок их расчета на основе экономико-математической модели МОБ.

8. Раскройте содержательный смысл принципиальной схемы динамического межотраслевого баланса. Дайте характеристику динамической межотраслевой балансовой модели.

Упражнения

1. На основании данных, приведенных в нижеследующих таблицах, рассчитать коэффициенты прямых и полных материальных затрат.

а)

Отрасль	Прямые межотраслевые потоки			Конечная продукция
	1	2	3	
1	50	60	80	60
2	25	90	40	25
3	25	60	40	35

б)

Отрасль	Прямые межотраслевые потоки			Конечная продукция
	1	2	3	
1	40	18	25	21
2	16	9	25	16
3	80	45	50	75

в)

Отрасль	Прямые межотраслевые потоки			Конечная продукция
	1	2	3	
1	18	36	25	1
2	45	90	25	20
3	36	36	50	30

2. В таблицах приведены коэффициенты прямых материальных затрат и объемы конечной продукции в межотраслевом балансе для трех отраслей:

а)

Отрасль	Коэффициенты прямых затрат			Конечная продукция
	1	2	3	
1	0,2	0,2	0,1	50
2	0,5	0,3	0,2	0
3	0,2	0,2	0,4	30

б)

Отрасль	Коэффициенты прямых затрат			Конечная продукция
	1	2	3	
1	0,3	0,4	0,2	40
2	0,2	0,1	0,3	15
3	0,1	0,5	0,2	10

Требуется:

1) проверить продуктивность матрицы коэффициентов, прямых затрат;

2) рассчитать коэффициенты полных материальных затрат;

3) найти объемы валовой продукции отраслей.

3. На основе данных таблиц в упр. 2 восстановить схемы межотраслевого материального баланса.

Глава 7
ЭКОНОМЕТРИЧЕСКИЕ МОДЕЛИ

— Общие понятия эконометрических моделей
— Задачи экономического анализа, решаемые на основе регрессионных эконометрических моделей
— Оценка качества эконометрических регрессионных моделей и прогнозирование на их основе
— Производственные функции

Изучение материалов главы 7 позволит студентам:

***знать*:**

- основные виды эконометрических моделей;
- методы построения одно- и многофакторных регрессионных моделей и использования их в задачах экономического анализа;
- математико-статистические методы оценки качества моделей регрессии и прогнозирования на основе этих моделей;
- экономическое содержание составных элементов производственных функций;

***уметь*:**

- оценивать тесноту и направление связи между экономическими показателями;
- строить однофакторные (линейные и нелинейные) и многофакторные регрессионные модели, оценивать их качество;
- использовать модели регрессии для экономического анализа и прогнозирования;
- применять аппарат производственных функций для оценки ряда показателей микро- и макроэкономики;

***владеть*:**

- понятийным аппаратом эконометрических исследований;
- методами корреляционного и регрессионного анализа;
- основными понятиями производственных функций.

7.1. Общие понятия эконометрических моделей

При анализе экономических явлений на основе экономико-математических методов особое место занимают модели, выявляющие количественные связи между изучаемыми показателями и влияющими на них факторами. Научной дисциплиной, предмет которой составляет изучение этой количественной стороны экономических явлений и процессов средствами математического и статистического анализа, является *эконометрика*, в которой результаты теоретического анализа экономики синтезируются с выводами математики и статистики. Основная задача эконометрики — проверка экономических теорий на фактическом (эмпирическом) материале при помощи методов математической статистики.

Главным инструментом эконометрики служит *эконометрическая модель*, т.е. экономико-математическая модель факторного анализа, параметры которой оцениваются средствами математической статистики. Эта модель выступает в качестве средства анализа и прогнозирования конкретных экономических процессов на основе реальной статистической информации.

Эконометрические модели можно классифицировать по ряду классификационных признаков. Так, по *аналитической форме* модели (уравнения) выделяют линейные, нелинейные, степенные модели, модели Брандона и др. Например, модель Брандона имеет вид

$$\hat{y} = \overline{y} f_1(x_1) \cdot f_2(x_2) \cdot \ldots \cdot f_m(x_m).$$

где y — изучаемый показатель (будем называть его *результативным признаком*), черта над ним означает среднюю величину (математическое ожидание); x_1, x_2, ... x_m — влияющие на изучаемый показатель величины (будем называть их *факторными признаками*).

Одной из основных классификационных рубрик эконометрических моделей является классификация по направлению и сложности причинных связей между показателями, характеризующими экономическую систему. Если пользоваться термином «переменная», то в любой достаточно сложной экономической системе можно выделить внутренние переменные (например, выпуск продукции, численность работников, производительность труда) и внешние переменные (например, поставка ресурсов, климатические условия и др.). Тогда по *направлению* и *сложности* связей между внутренними (*эндогенными*, выходными) переменными и внешними (*экзогенными*, входными) переменными выделяют следующие эконометрические модели: регрессионные модели, взаимозависимые системы, рекурсивные системы.

Регрессионными называют модели, основанные на уравнении регрессии, или системе регрессионных уравнений, связывающих величины эндогенных и экзогенных переменных. Различают уравнения (модели) парной и множественной регрессии. Если для обозначения эндогенных переменных использовать букву y, а для экзогенных переменных букву x, то в случае линейной модели уравнение парной регрессии имеет вид $\hat{y} = a_0 + a_1 x$, а уравнение множественной регрессии: $\hat{y} = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_m x_m$. Подобного типа модели рассматриваются подробно в следующих параграфах. Отметим только, что параметры моделей парной и множественной регрессии находятся на основе метода наименьших квадратов. Одним из часто применяемых видов нелинейных многофакторных регрессионных моделей являются *производственные функции*, ряд общих сведений о которых приведен в параграфе 7.4.

Взаимозависимые системы наиболее полно описывают экономическую систему, содержащую, как правило, множество взаимосвязанных эндогенных и экзогенных переменных. Такие модели задаются системой взаимозависимых уравнений следующего вида (n — число эндогенных переменных, m — число экзогенных переменных):

$$\begin{aligned}
&y_1 = a_{10} + a_{11}x_1 + \ldots + a_{1m}x_m + b_{12}y_2 + b_{13}y_3 + \ldots + b_{1n}y_n;\\
&y_2 = a_{20} + a_{21}x_1 + \ldots + a_{2m}x_m + b_{21}y_1 + b_{23}y_3 + \ldots + b_{2n}y_n;\\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot\\
&y_n = a_{n0} + a_{n1}x_1 + \ldots + a_{nm}x_m + b_{n1}y_1 + b_{n2}y_2 + \ldots + b_{nn-1}y_{n-1}.
\end{aligned}$$

Для нахождения параметров системы взаимозависимых уравнений используются более сложные методы: двух- и трехшаговый метод наименьших квадратов, методы максимального правдоподобия с полной и неполной информацией и др.

На практике стремятся упростить взаимозависимые системы и привести их к так называемому рекурсивному виду. Для этого сначала выбирают эндогенную переменную (внутренний показатель), зависящую только от экзогенных переменных (внешних факторов), обозначают ее y_1. Затем выбирается внутренний показатель, который зависит только от внешних факторов и от y_1 и т.д.; таким образом, каждый последующий показатель зависит только от внешних факторов и от внутренних предыдущих. Такие системы называются *рекурсивными*. Параметры первого уравнения рекурсивных систем находят методом наименьших квадратов, их подставляют во второе уравнение, и опять применяется метод наименьших квадратов, и т.д.

Процесс построения и использования эконометрических моделей является достаточно сложным и включает в себя следующие основные этапы: определение цели исследования, построение системы показателей и логический отбор факторов, наиболее влияющих на каждый показатель; выбор формы связи изучаемых показателей между собой и отобранными факторами, другими словами, выбор типа эконометрической модели; сбор исходных данных и анализ информации; построение эконометрической модели, т.е. определение ее параметров; проверка качества построенной модели, в первую очередь ее адекватности изучаемому экономическому процессу; использование модели для экономического анализа и прогнозирования.

При практической реализации указанных этапов очень важным является построение системы показателей исследуемого экономического процесса и определение перечня факторов, влияющих на каждый показатель.

Укажем основные требования, предъявляемые к включаемым в эконометрическую модель факторам.

— Каждый из факторов должен быть обоснован теоретически.

— В перечень целесообразно включать только важнейшие факторы, оказывающие существенное воздействие на изучаемые показатели; при этом рекомендуется, чтобы количество включаемых в модель факторов не превышало одной трети от числа наблюдений в выборке (длины временного ряда).

— Факторы не должны быть линейно зависимы, поскольку эта зависимость означает, что они характеризуют аналогичные свойства изучаемого явления. Например, заработная плата работников зависит, наряду с другими факторами, от роста производительности труда и от объема выпускаемой продукции. Однако эти факторы могут быть тесно взаимосвязаны, коррелированы и, следовательно, в модель целесообразно включать только один из этих факторов. Включение в модель линейно взаимозависимых факторов приводит к возникновению явления *мультиколлинеарности*, которое отрицательно сказывается на качестве модели; более подробно это явление описано ниже.

— Влияющие на экономический процесс факторы могут быть количественные и качественные. В модель рекомендуется включать только такие факторы, которые могут быть численно измерены.

— В одну модель нельзя включать совокупный фактор и образующие его частные факторы. Одновременное включение таких факторов приводит к неоправданно увеличенному их влиянию на зависимый показатель, к искажению реальной действительности.

При отборе влияющих факторов используются статистические методы отбора. Так, существенного сокращения числа влияющих факторов можно достичь с помощью пошаговых процедур отбора переменных. Ни одна из этих процедур не гарантирует получения оптимального набора переменных. Однако при практическом применении они позволяют получать достаточно хорошие наборы существенно влияющих факторов; кроме того, их можно сочетать с другими подходами к решению данной проблемы, например с экспертными оценками значимости факторов. Среди пошаговых процедур отбора факторов наиболее часто используются процедуры пошагового включения и исключения факторов. Обе эти процедуры хорошо формализованы и потому успешно реализованы в различных машинных программах статистического анализа.

Метод исключения предполагает построение уравнения, включающего всю совокупность переменных, с последующим последовательным (пошаговым) сокращением числа переменных в модели до тех пор, пока не выполнится некоторое наперед заданное условие. Суть *метода включения* — в последовательном включении переменных в модель до тех пор, пока регрессионная модель не будет отвечать заранее установленному критерию качества. Последовательность включения определяется с помощью частных коэффициентов корреляции: переменные, имеющие относительно исследуемого показателя большее значение частного коэффициента корреляции, первыми включаются в регрессионное уравнение.

Выше отмечено, что одной из предпосылок применения методов регрессионного анализа для построения эконометрических моделей является отсутствие среди независимых переменных (факторов) линейно связанных. Если данная предпосылка не выполняется, то возникает, как уже сказано выше, явление мультиколлинеарности, т.е. наличие сильной корреляции между независимыми переменными (включенными в модель факторами). В математическом аспекте мультиколлинеарность приводит к слабой обусловленности матрицы системы нормальных уравнений, т.е. близости ее определителя к нулю, а в содержательном аспекте — к искажению смысла коэффициентов регрессии и затруднению выявления наиболее существенно влияющих факторов.

Основные причины, вызывающие мультиколлинеарность, — независимые переменные, либо характеризующие одно и то же свойство изучаемого явления, либо являющиеся составными частями одного и того же признака.

В настоящее время существует ряд методов, позволяющих оценить наличие мультиколлинеарности в совокупности независимых переменных, измерить ее степень, выявить взаимно коррелированные переменные и устранить или ослабить ее негативное влияние на регрессионную модель. Наиболее распространенным методом выявления мультиколлинеарности является метод корреляции. На практике считают, что две переменные коллинеарны (линейно зависимы), если парный коэффициент корреляции между ними по абсолютной величине превышает 0,8. Устраняют мультиколлинеарность чаще всего путем исключения из модели одного из коррелированных факторов.

7.2. Задачи экономического анализа, решаемые на основе регрессионных эконометрических моделей

Вопросы построения и использования эконометрических моделей рассмотрим более подробно на примере линейных регрессионных моделей как в случае парной регрессии (однофакторная модель), так и в случае множественной регрессии (многофакторная модель); в последнем случае будем рассматривать модели множественной регрессии на примере линейной двухфакторной модели.

Основу математического аппарата для рассматриваемых моделей составляют такие разделы математической статистики, как корреляционный и регрессионный анализ. Для определенности эндогенные переменные в этих моделях будем называть результативными признаками и обозначать их, как и ранее, буквой y, а экзогенные переменные будем называть факторными признаками и обозначать их буквой x. Методы корреляционно-регрессионного анализа позволяют решать три основные задачи: определение формы связи между результативным и факторными признаками, измерение тесноты связи между ними, анализ влияния отдельных факторных признаков. Рассмотрим решение этих задач для указанных видов эконометрических моделей; при этом для наглядности будем иллюстрировать выводы на конкретном примере экономического анализа.

В табл. 7.1 представлены статистические данные о расходах на питание, душевом доходе и размере семьи для девяти групп семей. Требуется проанализировать зависимость величины расходов на питание от величины душевого дохода и размера семьи. В соответствии с этим первый показатель будет результативным признаком, который обозначим y, а два других будут факторными признаками, или просто факторами, и мы их обозначим соответственно x_1 и x_2.

Таблица 7.1

Номер группы	Расход на питание (y)	Душевой доход (x_1)	Размер семей (x_2)
1	433	628	1,5
2	616	1577	2,1
3	900	2659	2,7
4	1113	3701	3,2
5	1305	4796	3,4
6	1488	5926	3,6
7	1645	7281	3,7
8	1914	9350	4,0
9	2411	18807	3,7

Рассмотрим сначала однофакторную линейную модель зависимости расходов на питание (y) от величины душевого дохода семей (x_1). Она выражается линейной функцией вида

$$\hat{y} = a_0 + a_1 x_1, \tag{7.1}$$

параметры которой a_0 и a_1 находятся в результате решения системы нормальных уравнений, которая в свою очередь формируется, как уже отмечалось в главе 5, на основе метода наименьших квадратов. Система нормальных уравнений для рассматриваемого случая аналогична системе (5.5) и имеет вид

$$\begin{cases} na_0 + \left(\sum x_1\right)a_1 = \sum y \\ \left(\sum x_1\right)a_0 + \left(\sum x_1^2\right)a_1 = \sum yx_1, \end{cases} \tag{7.2}$$

где суммирование проводится по всем n группам. Используя данные табл. 7.1, получим систему уравнений:

$$\begin{cases} 9a_0 + 54\,725a_1 = 11\,825 \\ 54\,725a_0 + 575\,906\,797a_1 = 98\,049\,159, \end{cases}$$

решением которой являются значения $a_0 = 660{,}03$; $a_1, = 0{,}11$. Таким образом, модель имеет вид

$$\hat{y} = 660{,}03 + 0{,}11x_1. \tag{7.3}$$

Уравнение (7.3) называется *уравнением регрессии*, коэффициент a_1 — *коэффициентом регрессии*. Направление связи между y и x_1 определяет знак коэффициента регрессии a_1; в нашем случае данная связь является прямой. Теснота этой связи определяется *коэффициентом корреляции* (парным):

$$r_{\hat{y}x_1} = \sqrt{1 - \frac{S^2_{\hat{y}x_1}}{S^2_y}}; \tag{7.4}$$

в этой формуле S_y — средняя квадратическая ошибка выборки y из табл. 7.1:

$$S_y = \sqrt{\frac{\sum (y - \overline{y})^2}{n}},$$

где $\overline{y}$ — средняя арифметическая значений y;

$S^2_{\hat{y}x_1}$ — средняя квадратическая ошибка уравнения (7.3) для числа степеней свободы $n - 2$:

$$S_{\hat{y}x_1} = \sqrt{\frac{\sum (y - \overline{y})^2}{n - 2}},$$

где $\hat{y}$ — соответствующее значение расходов на питание, вычисленное по модели (7.3).

В этих формулах, как и ранее, суммирование ведется по всем группам от 1 до n.

Чем ближе значение коэффициента корреляции к единице, тем теснее корреляционная связь. В нашем примере $S^2_{\hat{y}x_1} = 63\,846$, $S^2_y = 454\,070$, следовательно,

$$r_{\hat{y}x_1} = \sqrt{1 - \frac{63\,846}{454\,070}} = 0{,}927.$$

Полученное значение $r_{\hat{y}x_1}$ свидетельствует, что связь между расходами на питание и душевым доходом очень тесная.

Величина $r^2_{\hat{y}x_1}$ называется *коэффициентом детерминации* и показывает долю изменения (вариации) результативного признака под действием факторного признака. В нашем случае $r^2_{\hat{y}x_1}$ = 0,859; это означает, что фактором душевого дохода можно объяснить почти 86% изменения расходов на питание.

Коэффициенты регрессии (в рассматриваемом случае это коэффициент a_1) нельзя использовать для непосредственной оценки влияния факторов на результативный признак из-за различия единиц измерения исследуемых показателей. Для этих целей вычисляются коэффициенты эластичности и бета-коэффициент.

Коэффициент эластичности для рассматриваемой модели парной регрессии рассчитывается по формуле

$$Э_{\hat{y}x_1} = \frac{a_1 \bar{x}_1}{\bar{y}}. \tag{7.5}$$

Он показывает, на сколько процентов изменяется результативный признак y при изменении факторного признака x_1 на один процент.

В нашем примере коэффициент регрессии a_1 равен 0,11, а средние арифметические $\bar{x}_1$ и $\bar{y}$ равны соответственно 6080,6 и 1313,9.

Поэтому коэффициент эластичности расходов на питание в зависимости от душевого дохода будет равен

$$Э_{\hat{y}x_1} = \frac{0{,}11 \cdot 6080{,}6}{1313{,}9} = 0{,}51.$$

Это означает, что при увеличении душевого дохода на 1% расходы на питание увеличатся на 0,51%.

Бета-коэффициент в нашем случае задается формулой

$$\beta_{\hat{y}x_1} = \frac{a_1 S_{x_1}}{S_y}, \tag{7.6}$$

где S_{x_1} и S_y — средние квадратические ошибки выборки величин x_1 и y из табл. 7.1 соответственно.

Величина S_y^2 уже была рассчитана ранее и равна 454 070, поэтому величина S_y равна 673,8; аналогичные расчеты дают значение величины S_{x_1}, равное 4242,0. Бета-коэффициент показывает, на какую часть величины своего среднего квадратического отклонения изменится в среднем значение результативного признака при изменении факторного признака на величину его среднеквадратического отклонения.

В нашем случае получаем следующее значение бета-коэффициента:

$$\beta_{\hat{y}x_1} = \frac{0{,}11 \cdot 4242{,}0}{673{,}8} = 0{,}69,$$

т.е. увеличение душевого дохода на величину среднеквадратического отклонения этого показателя приведет к увеличению среднего значения расходов на питание на 0,69 среднеквадратического отклонения этих расходов.

Рассмотрим теперь двухфакторную линейную модель зависимости расходов на питание (y) от величины душевого дохода семей (x_1) и размера семей (x_2). Как уже отмечено выше, множественный (многофакторный) корреляционно-регрессионный анализ решает три задачи: определяет форму связи результативного признака с факторными, выявляет тесноту этой связи и устанавливает влияние отдельных факторов. В нашем случае эта модель имеет вид

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2. \tag{7.7}$$

Параметры модели a_0, a_1 и a_2 находятся путем решения системы нормальных уравнений:

$$\begin{cases} na_0 + \left(\sum x_1\right)a_1 + \left(\sum x_2\right)a_2 = \sum y \\ \left(\sum x_1\right)a_0 + \left(\sum x_1^2\right)a_1 + \left(\sum x_1 x_2\right)a_2 = \sum yx_1 \\ \left(\sum x_2\right)a_0 + \left(\sum x_1 x_2\right)a_1 + \left(\sum x_2^2\right)a_2 = \sum yx_2. \end{cases} \tag{7.8}$$

Используя данные табл. 7.1, получим систему нормальных уравнений в виде

$$\begin{cases} 9a_0 + 54\,725a_1 + 27{,}9a_2 = 11\,825 \\ 54\,725a_0 + 575\,906\,797a_1 + 194\,841{,}8a_2 = 98\,049\,159 \\ 27{,}9a_0 + 194\,841{,}8a_1 + 92{,}1a_2 = 40\,391{,}7. \end{cases}$$

Решая эту систему (например, методом Гаусса), получим: $a_0 = -192{,}0$; a_1, $= 0{,}072$; $a_2 = 344{,}9$, так что модель (7.7) имеет вид

$$\hat{y} = -192{,}0 + 0{,}072x_1 + 344{,}9x_2.$$

Для определения тесноты связи предварительно вычисляются парные коэффициенты корреляции $r_{yx_1}, r_{yx_2}, r_{x_1x_2}$. Например,

$$r_{yx_1} = \frac{\overline{yx_1} - \overline{y} \cdot \overline{x}_1}{S_y S_{x_1}}, \tag{7.9}$$

где черта над символами означает среднюю арифметическую, а S_y и S_x — средние квадратические ошибки соответствующих выборок из табл. 7.1:

$$S_y = \sqrt{\frac{\sum(y - \overline{y})^2}{n}}; \quad S_{x_1} = \sqrt{\frac{\sum(x_1 - \overline{x}_1)^2}{n}}.$$

Аналогичный вид имеют формулы для r_{yx_2} и $r_{x_1x_2}$.

После этого вычисляют коэффициент множественной корреляции

$$R_{\hat{y}x_1x_2} = \sqrt{\frac{r^2_{\hat{y}x_1} + r^2_{yx_2} - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r^2_{x_1x_2}}}, \tag{7.10}$$

который колеблется в пределах от 0 до 1; чем ближе он к 1, тем в большей степени учтены факторы, влияющие на результативный признак.

В нашем примере расчеты дают следующее значение коэффициента множественной корреляции: $R_{\hat{y}x_1x_2} = 0{,}983$, что выше значения коэффициента корреляции в случае однофакторной модели. Таким образом, степень тесноты связи расходов на питание с факторами душевого дохода и размера семей является очень высокой.

Величина $R^2_{\hat{y}x_1x_2}$ называется *совокупным коэффициентом детерминации* и показывает долю вариации результативного признака под воздействием изучаемых факторных признаков. В нашем примере $R^2_{\hat{y}x_1x_2} = 0{,}966$; это означает, что совместное влияние душевого дохода и размера семей объясняет почти 97% изменения расходов на питание.

Задача анализа тесноты связи между результативным и одним из факторных признаков при неизменных значениях других факторов решается в многофакторных моделях при помощи *частных коэффициентов корреляции*. Так, частный коэффициент корреляции между результативным признаком y и факторным признаком x_1 при неизменном значении факторного признака x_2 рассчитывается по формуле

$$r_{\hat{y}x_1(x_2)} = \frac{r_{yx_1} - r_{yx_2} r_{x_1x_2}}{\sqrt{(1-r_{yx_2}^2)(1-r_{x_1x_2}^2)}}, \tag{7.11}$$

где используются парные коэффициенты корреляции, рассчитываемые по формулам, аналогичным (7.9).

Аналогичная формула имеет место для частного коэффициента корреляции $r_{\hat{y}x_2(x_1)}$ между результативным признаком y и факторным признаком x_2 при неизменном значении факторного признака x_1.

Для рассматриваемого примера частные коэффициенты корреляции расходов на питание от душевого дохода и размера семей составляют

$$r_{\hat{y}x_1(x_2)} = 0{,}927; \quad r_{\hat{y}x_2(x_1)} = 0{,}849,$$

т.е. теснота связи между расходами на питание и одним из исследуемых факторов при неизменном значении другого является весьма значительной.

Если частные коэффициенты корреляции возвести в квадрат, то получим *частные коэффициенты детерминации*, показывающие долю вариации результативного признака под действием одного из факторов при неизменном значении другого фактора. В нашей задаче

$$r_{\hat{y}x_1(x_2)}^2 = 0{,}859; \quad r_{\hat{y}x_2(x_1)}^2 = 0{,}721,$$

следовательно, влиянием душевого дохода при неизменном размере семьи объясняется почти 86% изменения расходов на питание, а изменение размера семьи при неизменном душевом доходе объясняет более 72% изменения расходов на питание.

Влияние отдельных факторов в многофакторных моделях может быть охарактеризовано с помощью *частных коэффициентов эластичности*, которые в случае линейной двухфакторной модели (7.7) рассчитываются по формулам

$$Э_{\hat{y}x_1(x_2)} = \frac{a_1\overline{x}_1}{\overline{y}}; \quad Э_{\hat{y}x_2(x_1)} = \frac{a_2\overline{x}_2}{\overline{y}}. \tag{7.12}$$

Черта над символом, как и ранее, означает среднюю арифметическую. Частные коэффициенты эластичности показывают, на сколько процентов изменится результативный признак, если значение одного из факторных признаков изменится на 1%, а значение другого факторного признака останется неизменным.

В рассматриваемом примере $a_1 = 0{,}072$; $a_2 = 344{,}9$; $\overline{y} = 1313{,}9$; $\overline{x}_1 = 6080{,}6$; $\overline{x}_2 = 3{,}1$, следовательно, по формулам (7.12) получим:

$$Э_{\hat{y}x_1(x_2)} = \frac{0{,}072 \cdot 6080{,}6}{1313{,}9} = 0{,}333; \quad Э_{\hat{y}x_2(x_1)} = \frac{344{,}9 \cdot 3{,}1}{1313{,}9} = 0{,}790.$$

Это означает, что при увеличении душевого дохода на 1% и неизменном размере семьи расходы на питание увеличатся на 0,333%, а увеличение (условное) на 1% размера семьи при неизменном душевом доходе приведет к возрастанию расходов на питание на 0,790%.

Определенные выводы о влиянии отдельных факторов на результативный признак в случае линейной модели множественной регрессии можно сделать на основе расчета *частных бета-коэффициентов*, которые для двухфакторной модели (7.7) задаются формулами

$$\beta_{\hat{y}x_1(x_2)} = \frac{a_1 S_{x_1}}{S_y}; \quad \beta_{\hat{y}x_2(x_1)} = \frac{a_2 S_{x_2}}{S_y}. \tag{7.13}$$

Частные бета-коэффициенты показывают, на какую долю своего среднеквадратического отклонения изменится в среднем результативный признак при изменении одного из факторных признаков на величину его среднеквадратического отклонения и неизменном значении остальных факторов.

В рассматриваемой задаче $a_1 = 0{,}072$; $a_2 = 344{,}9$; $S_y = 673{,}8$; $S_{x_1} = 4242{,}0$; $S_{x_2} = 0{,}79$, так что расчеты по формулам (7.13) дают следующие значения частных бета-коэффициентов:

$$\beta_{\hat{y}x_1(x_2)} = \frac{0{,}072 \cdot 4242{,}0}{673{,}8} = 0{,}45; \quad \beta_{\hat{y}x_2(x_1)} = \frac{344{,}9 \cdot 0{,}79}{673{,}8} = 0{,}40.$$

Это означает, что при неизменном составе семей увеличение на величину своего среднеквадратического отклонения размера душевого дохода приведет к увеличению среднего значения расходов на питание на 0,45 их среднеквадратического отклонения, а при неизменном душевом доходе увеличение размера семей на величину его среднеквадратического отклонения приведет к возрастанию расходов на питание лишь на 0,40 их среднеквадратического отклонения.

7.3. Оценка качества эконометрических регрессионных моделей и прогнозирование на их основе

Будем рассматривать, как и в предыдущем параграфе, линейные эконометрические модели регрессии. Их качество оценивается стандартным для экономико-математических моделей образом: по адекватности и точности. Адекватность регрессионных моделей может быть установлена, как и в случае трендовых моделей, на основе анализа остаточной последовательности; при этом расчетные значения получаются подстановкой в модель фактических значений всех включенных в модель факторов. Остаточная последовательность проверяется на выполнение свойств случайной компоненты временного экономического ряда: близость нулю математического ожидания, случайный характер отклонений, отсутствие автокорреляции и нормальность закона распределения. Эта проверка проводится теми же методами и с использованием тех же статистических критериев, что и для трендовых моделей (см. параграф 5.2).

О качестве моделей регрессии можно судить также по значениям коэффициента корреляции (индекса корреляции) и коэффициента детерминации для однофакторной модели и по значениям коэффициента множественной корреляции и совокупного коэффициента детерминации для моделей множественной регрессии. Формулы расчета этих коэффициентов приведены в параграфе 7.2. Чем ближе абсолютные величины указанных коэффициентов к 1, тем теснее связь между изучаемым признаком и выбранными факторами и, следовательно, с тем большей уверенностью можно судить об адекватности построенной модели, включающей в себя наиболее влияющие факторы.

Для оценки точности регрессионных моделей обычно используются те же статистические критерии точности, что и для трендовых моделей, в частности, средняя относительная ошибка аппроксимации (см. формулу (5.14)). Проверка *значимости модели* регрессии проводится с использованием

F-критерия Фишера, расчетное значение которого находится как отношение дисперсии исходного ряда наблюдений изучаемого показателя и несмещенной оценки дисперсии остаточной последовательности для данной модели. Если расчетное значение этого критерия со степенями свободы $v_1 = n - 1$ и $v_2 = n - m - 1$, где n — количество наблюдений и m — число включенных в модель факторов, больше табличного значения критерия Фишера при заданном уровне значимости, то модель признается статистически значимой.

При проверке качества регрессионной модели целесообразно оценить также *значимость коэффициентов регрессии*. Эта оценка проводится по t-статистике Стьюдента путем проверки гипотезы о равенстве нулю k-го коэффициента регрессии ($k = 1, 2, ..., m$). Расчетное значение t критерия с числом степеней свободы ($n - m - 1$) находят путем деления k-го коэффициента регрессии на среднеквадратическое отклонение этого коэффициента, которое в свою очередь вычисляется как квадратный корень из произведения несмещенной оценки дисперсии остаточной компоненты и k-го диагонального элемента матрицы, обратной матрице системы нормальных уравнений относительно параметров модели. Это расчетное значение сравнивается с табличным значением критерия Стьюдента при заданном уровне значимости, и если оно больше табличного значения, коэффициент регрессии считается статистически значимым. В противном случае соответствующий данному коэффициенту регрессии фактор следует исключить из модели, при этом качество модели не ухудшится.

Перейдем к вопросу экономического прогнозирования на основе модели регрессии, при этом будем предполагать, что модель, построенная на базе временных рядов изучаемого показателя и включенных в модель факторов, является адекватной и достаточно точной. При использовании построенной модели для прогнозирования делается также предположение о сохранении существовавших ранее взаимосвязей переменных и на период упреждения.

Для прогнозирования зависимой переменной (результативного признака) на L шагов вперед необходимо знать прогнозные значения всех входящих в модель факторов. Эти значения могут быть получены на основе экстраполяционных методов, например, с использованием средних абсолютных приростов факторных признаков; они могут быть также определены методами экспертных оценок или непосредственно заданы исследователем экономического процесса. Прогнозные значения факторов подставляют в модель и получают точечные прогнозные оценки изучаемого показателя.

Для определения области возможных значений результативного показателя при известных значениях факторов, т.е. доверительного интервала прогноза, необходимо учитывать два возможных источника ошибок. Ошибки первого рода вызываются рассеиванием наблюдений относительно линии регрессии, и их можно учесть, в частности, величиной среднеквадратической ошибки аппроксимации изучаемого показателя с помощью регрессионной модели. Обозначим эту величину $S_{\hat{y}}$ и вычислим ее по формуле, аналогичной (5.17).

Ошибки второго рода обусловлены тем, что в действительности жестко заданные в модели коэффициенты регрессии являются случайными величинами, распределенными по нормальному закону. Эти ошибки учитываются вводом поправочного коэффициента при расчете ширины доверительного интервала; формула для его расчета включает табличное значение t-статистики при заданном уровне значимости и зависит от вида регрессионной модели.

Для линейной однофакторной модели, общий вид которой имеет структуру, аналогичную (7.1), величина отклонения от линии регрессии задается выражением (обозначим его R):

$$R(n, L, \alpha) = S_{\hat{y}} t_{\alpha} \sqrt{1 + 1/n + (x_{n+L} - \bar{x})^2 \Big/ \sum_{t=1}^{n} (x_t - \bar{x})^2}. \qquad (7.14)$$

Здесь n — число наблюдений, L — количество шагов вперед, α — уровень значимости прогноза, x_t — наблюдаемое значение факторного признака в момент t, $\bar{x}$ — среднее значение наблюдаемого фактора, x_{n+L} — прогнозное значение фактора на L шагов вперед.

Таким образом, для рассматриваемой модели формула расчета нижней и верхней границ доверительного интервала прогноза имеет вид

$$U_y = \hat{y}_{n+L} \pm R(n, L, \alpha), \qquad (7.15)$$

где $\hat{y}_{n+L}$ означает точечную прогнозную оценку изучаемого результативного показателя по модели на L шагов вперед.

7.4. Производственные функции

Производственными функциями называются экономико-математические модели, связывающие переменные величины затрат с величинами выпуска. Понятия «затраты» и «выпуск» имеют отношение, как правило, к процессу производства продукции; это объясняет происхождение названия данного типа моделей. Если рассматривается экономика региона или страны в целом, то разрабатываются агрегированные производственные функции, в которых выпуском служит показатель совокупного общественного продукта. Частными случаями производственных функций являются *функции выпуска* (зависимость объема производства от наличия или потребления ресурсов), *функции издержек* (связь объема продукции и издержек производства), *функции капитальных затрат* (зависимость капитальных вложений от производственной мощности создаваемых предприятий) и др.

Широко используются мультипликативные формы представления производственных функций. В самом общем виде мультипликативная производственная функция записывается следующим образом:

$$P = A \cdot X_1^{\alpha} \cdot X_2^{\beta} \cdot \ldots \cdot X_n^{\gamma}. \tag{7.16}$$

Здесь коэффициент A определяет размерность величин и зависит от избранных единиц измерения затрат и выпуска. Сомножители X_i представляют влияющие факторы и могут иметь различное экономическое содержание в зависимости от того, какие факторы влияют на величину выпуска P. Степенные параметры α, β, ..., γ показывают ту долю в приросте конечного продукта, которую вносит каждый из факторов-сомножителей; они называются *коэффициентами эластичности производства относительно затрат* соответствующего ресурса и показывают, на сколько процентов возрастает выпуск при увеличении затрат данного ресурса на один процент.

Сумма коэффициентов эластичности $E = \alpha + \beta + ... + \gamma$ имеет важное значение для характеристики свойств производственной функции. Предположим, что затраты всех видов ресурсов возрастают в k раз. Тогда величина выпуска в соответствии с (7.16) составит

$$P_1 = A \cdot (k \cdot X_1)^{\alpha} \cdot (k \cdot X_2)^{\beta} \cdot \ldots \cdot (k \cdot X_n)^{\gamma} = A \cdot k^E \cdot P.$$

Следовательно, если $E = \alpha + \beta + ... + \gamma = 1$, то при увеличении затрат в k раз выпуск возрастает также в k раз; производственная функция в этом случае является линейно однородной. При $E > 1$ такое же увеличение затрат приведет к росту выпуска более чем в k раз, а при $E < 1$ — менее чем в k раз (так называемый эффект масштаба).

В качестве примера мультипликативных производственных функций можно привести широко известную производственную функцию Кобба — Дугласа:

$$N = A \cdot L^{\alpha} \cdot K^{\beta}, \tag{7.17}$$

где:

N — национальный доход;

A — коэффициент размерности;

L, K — объемы приложенного труда и основного капитала соответственно;

α и β — коэффициенты эластичности национального дохода по труду L и капиталу K.

Эта функция применялась американскими исследователями при анализе развития экономики США в 30-х годах прошлого века.

Эффективность использования ресурсов характеризуется двумя основными показателями: *средняя (абсолютная) эффективность* ресурса

$$\mu_i = \frac{P(X)}{X_i} \tag{7.18}$$

и *предельная эффективность* ресурса

$$\nu_i = \frac{\partial P(X)}{\partial X_i}. \tag{7.19}$$

Экономический смысл величины μ_i очевиден; в зависимости от типа ресурса она характеризует такие показатели, как производительность труда, фондоотдача и др. Величина ν_i показывает предельный прирост выпуска продукта при увеличении затрат i-го ресурса на «малую единицу» (на 1 руб., на 1 нормо-час и т.д.).

Множество точек n-мерного пространства факторов производства (ресурсов), удовлетворяющих условию постоянства выпуска $P(X) = C$, называется *изоквантой*. Важнейшими свойствами изоквант являются следующие: изокванты не пересекаются друг с другом; большей величине выпуска соответствует более удаленная от начала координат изокванта; если все ресурсы абсолютно необходимы для производства, то изокванты не имеют общих точек с координатными гиперплоскостями и с осями координат.

В материальном производстве большое значение приобретает понятие *взаимозаменяемости ресурсов*. В теории производственных функций возможности замещения ресурсов характеризуют производственную функцию с точки зрения различных комбинаций затрат ресурсов, приводящих к одному и тому же уровню выпуска продукта. Поясним это на условном примере. Пусть производство определенного количества сельхозпродукции требует 10 работников и 2 т удобрений, а при внесении в почву только 1 т удобрений потребуется уже 12 работников для получения того же урожая. Здесь 1 т удобрений (первый ресурс) заменяется трудом двух работников (второй ресурс).

Условия эквивалентной взаимозаменяемости ресурсов в некоторой точке $X^0 = (X_1^0, X_2^0, \ldots, X_n^0)$ вытекают из равенства $dP = 0$:

$$\sum_{i=1}^{n} \nu_i(X^0)dX_i = 0.$$

Отсюда *предельная норма замещения* (эквивалентной заменяемости) каких-либо двух ресурсов k и l задается формулой

$$\sigma_{kl} = \frac{dX_k}{dX_l} = -\frac{\nu_k(X^0)}{\nu_l(X^0)}. \tag{7.20}$$

Предельная норма замещения как показатель производственной функции характеризует относительную эффективность допускающих взаимную замену факторов производства при движении вдоль изокванты. Например, для функции Кобба — Дугласа предельная норма замещения затрат труда затратами капитала, т.е. производственными фондами, имеет вид

$$\sigma_{KL} = \frac{dK}{dL} = -\frac{\alpha \cdot K}{\beta \cdot L}. \tag{7.21}$$

Знак минус в правых частях формул (7.20) и (7.21) означает, что при фиксированном объеме производства увеличению одного из взаимозаменяемых ресурсов соответствует уменьшение другого.

Пример 7.1. Рассмотрим пример производственной функции Кобба — Дугласа, для которой известны коэффициенты эластичности выпуска по труду и капиталу: $\alpha = 0{,}3$; $\beta = 0{,}7$, а также затраты труда и капитала: $L = 30$ тыс. чел.; $K = 490$ млн руб. В этих условиях предельная норма замещения производственных фондов затратами труда равна

$$\tau_{kl} = -\frac{0{,}3 \cdot 490 \text{ млн руб}}{0{,}7 \cdot 30 \text{ тыс. чел.}} = -7\frac{\text{тыс. руб.}}{\text{чел.}} \qquad (7.22)$$

Таким образом, в этом условном примере в тех точках двухмерного пространства (L, K), где ресурсы труда и капитала взаимозаменяемы, уменьшение производственных фондов на 7 тыс. руб. может быть компенсировано увеличением затрат труда на 1 чел., и наоборот.

С понятием предельной нормы замещения связано понятие *эластичности замещения ресурсов*. Коэффициент эластичности замещения γ_{kl} характеризует отношение относительного изменения соотношения затрат ресурсов k и l к относительному изменению предельной нормы замещения этих ресурсов:

$$\gamma_{kl} = \frac{d(X_k / X_l)}{X_k / X_l} : \frac{\partial\sigma_{kl}}{\sigma_{kl}} = \frac{d(X_k / X_l)}{\partial\sigma_{kl}} \cdot \frac{\sigma_{kl} X_l}{X_k}. \qquad (7.23)$$

Этот коэффициент показывает, на сколько процентов должно измениться отношение между взаимозаменяемыми ресурсами, чтобы предельная норма замещения этих ресурсов изменилась на 1%. Чем выше эластичность замены ресурсов, тем в более широких пределах они могут заменять друг друга. При бесконечной эластичности ($\gamma_{kl} = +\infty$) не существует границ взаимозаменяемости ресурсов. При нулевой эластичности замещения ($\gamma_{kl} = 0$) возможность замены отсутствует; в этом случае ресурсы взаимодополняют друг друга и обязательно должны использоваться в определенном соотношении.

Рассмотрим в дополнение к функции Кобба — Дугласа некоторые другие производственные функции, широко используемые в качестве эконометрических моделей. *Линейная производственная функция* имеет вид

$$P = \sum_{i=1}^{n} A_i \cdot X_i,$$

A_i — оцениваемые параметры модели;
X_i, — факторы производства, взаимозамещаемые в любых пропорциях (эластичность замещения $\gamma_{kl} = +\infty$).

Изокванты этой производственной функции образуют семейство параллельных гиперплоскостей в неотрицательном ортанте n-мерного пространства факторов.

Во многих исследованиях применяются *производственные функции с постоянной эластичностью замещения*:

$$P = A \cdot \left(\sum_{i=1}^{n} A_i \cdot X_i^{-\rho} \right)^{-n/\rho}. \tag{7.23}$$

Производственная функция (7.23) является однородной функцией степени n. Все эластичности замещения ресурсов равны между собой:

$$\gamma_{kl} = \gamma = \frac{1}{1+\rho},$$

вследствие этого данная функция называется *функцией с постоянной эластичностью замещения* (*функцией CES*). Если $\rho > 0$, эластичность замещения γ меньше единицы; если $-1 < \rho < 0$, величина γ больше единицы; при $\rho = 0$ функция CES преобразуется в мультипликативную степенную производственную функцию (7.16).

Двухфакторная функция *CES* имеет вид

$$P = A \cdot \left(\alpha L^{-\rho} + \beta K^{-\rho} \right)^{-n/\rho}.$$

При $n = 1$ и $\rho = 0$ эта функция преобразуется в функцию типа функции Кобба — Дугласа (7.17).

Кроме производственных функций с постоянными коэффициентами эластичности выпуска от ресурсов и постоянной эластичностью замещения ресурсов в экономическом анализе и прогнозировании применяются и функции более общего вида. В качестве примера можно привести функцию

$$P = A \cdot L^{\alpha} \cdot K^{\beta} \cdot e^{kz}.$$

Эта функция отличается от функции Кобба — Дугласа множителем e^{kz}, где $z = K/L$ — фондовооруженность (капиталовооруженность) труда, и в ней эластичность замещения принимает различные значения в зависимости от уровня капиталовооруженности труда. В связи с этим данная функция относится к типу *производственных функций с переменной эластичностью замещения* (*функции VES*).

Перейдем к рассмотрению ряда вопросов практического использования производственных функций в экономическом анализе. Макроэкономические производственные функции применяются как инструмент прогнозирования объемов валовой продукции, конечного продукта и национального дохода, для анализа сравнительной эффективности факторов производства. Так, важным условием роста производства и производительности труда является увеличение фондовооруженности труда. Если для функции Кобба — Дугласа

$$P = A \cdot L^{\alpha} \cdot K^{\beta}$$

задать условие линейной однородности $\alpha + \beta = 1$, то из соотношения между производительностью труда (P/L) и фондовооруженностью труда (K/L)

$$\frac{P}{L} = A \cdot L^{\alpha-1} \cdot K^{\beta} = A\left(\frac{K}{L}\right)^{\beta} \tag{7.24}$$

следует, что производительность труда растет медленнее фондовооруженности, так как $0 < \beta < 1$. Этот вывод, как и многие другие результаты анализа на основе производственных функций, всегда справедлив для статических производственных функций, не учитывающих совершенствования технических средств труда и качественных характеристик используемых ресурсов, т.е. без учета технического прогресса. Для оценки параметров модели (7.24) ее линеаризируют путем логарифмирования:

$$\ln(P / L) = \ln A + \beta \cdot \ln(K / L).$$

Наряду с количественным увеличением используемых объемов ресурсов (трудовых ресурсов, производственных фондов и т.д.) важнейшим фактором роста производства служит научно-технический прогресс, заключающийся в совершенствовании технических средств и технологии, повышении квалификации работающих, улучшении организации управления производством. Статические эконометрические модели, в том числе и статические производственные функции, не учитывают фактор технического прогресса, поэтому используются динамические макроэкономические производственные функции, параметры которых определяются путем обработки временных рядов. Технический прогресс обычно отражают в производственных функциях в виде тенденции развития производства, зависящей от времени.

Например, функция Кобба — Дугласа с учетом фактора технического прогресса приобретает следующий вид:

$$P = A \cdot L^{\alpha} \cdot K^{\beta} \cdot e^{\lambda t}. \qquad (7.25)$$

В модели (7.25) множитель $e^{\lambda t}$ отражает тенденцию развития производства, связанную с научно-техническим прогрессом. В этом множителе t — время, а λ — темп прироста выпуска продукции благодаря техническому прогрессу. При практическом использовании модели (7.25) для оценки ее параметров проводится линеаризация путем логарифмирования, аналогично модели (7.24):

$$\ln P = \ln A + \alpha \cdot \ln L + \beta \ln K + \lambda t.$$

Следует особо отметить, что при построении производственных функций, как и для всех многофакторных эконометрических моделей, весьма важным моментом является правильный отбор влияющих факторов X_j. В частности, необходимо избавляться от явлений мультиколлинеарности факторов и явлений автокорреляции внутри каждого из них. Этот вопрос детально описан в параграфе 7.1 данной главы. При оценке параметров производственных функций на основе статистических наблюдений, включая временные ряды, основным методом является метод наименьших квадратов.

Рассмотрим применение производственных функций для экономического анализа и прогнозирования на условном примере из области экономики труда.

Пример 7.2. Пусть объем выпуска продукции отрасли характеризуется производственной функцией типа функции Кобба — Дугласа:

$$P = A \cdot T^{\alpha} \cdot \Phi^{\beta},$$

где:
P — объем выпуска продукции (млн руб.);
T — численность работников отрасли (тыс. чел.);
Φ — среднегодовая стоимость основных производственных фондов (млн руб.).

Допустим, параметры этой производственной функции известны и равны: $\alpha = 0{,}3$; $\beta = 0{,}7$; коэффициент размерности $A = 0{,}6$ (тыс. руб./чел.)^{0,3}. Известна также величина среднегодовой стоимости основных производственных фондов $\Phi = 900$ млн руб. В этих условиях требуется:

1) определить количество работников отрасли, необходимое для выпуска продукции в объеме 300 млн руб.;

2) выяснить, как изменится выпуск продукции при увеличении численности работающих на 1% и тех же объемах производственных фондов;

3) оценить взаимозаменяемость материальных и трудовых ресурсов.

Чтобы ответить на вопрос первого задания, линеаризируем эту производственную функцию путем логарифмирования по натуральному основанию:

$$\ln P = \ln A + \alpha \cdot \ln T + \beta \cdot \ln \Phi,$$

откуда следует, что

$$\ln T = \frac{\ln P - \ln A - \beta \cdot \ln \Phi}{\alpha} = \frac{\ln(P / A) - \beta \cdot \ln \Phi}{\alpha}.$$

Подставляя исходные данные, получим

$$\ln T = \frac{\ln(300 / 0,6) - 0,7 \cdot \ln 900}{0,3} = 4,83.$$

Отсюда $T = e^{4,83} = 125,2$ (тыс. чел.).

Рассмотрим второе задание. Так как $\alpha + \beta = 0,3 + 0,7 = 1$, данная производственная функция является линейно однородной; в соответствии с этим коэффициенты α и β являются коэффициентами эластичности выпуска по труду и фондам соответственно. Следовательно, увеличение числа работающих отрасли на 1% при неизменном объеме производственных фондов приведет к росту выпуска продукции на 0,3%, т.е. выпуск составит 300,9 млн руб.

Переходя к третьему заданию, рассчитаем предельную норму замещения производственных фондов трудовыми ресурсами. В соответствии с формулой (7.21)

$$\sigma_{\Phi T} = -\frac{\alpha \cdot \Phi}{\beta \cdot T} = -\frac{0,3 \cdot 900}{0,7 \cdot 125,2} = -3,08\,(\text{тыс. руб.} / \text{чел.});$$

Таким образом, при условии взаимозаменяемости ресурсов для обеспечения постоянства выпуска (т.е. при движении по изокванте) уменьшение производственных фондов отрасли на 3,08 тыс. руб. может быть возмещено увеличением трудовых ресурсов на 1 чел., и наоборот.

Вопросы и задания

1. Дайте общее понятие эконометрической модели. Какие виды эконометрических моделей вы знаете?
2. Чем вызывается явление мультиколлинеарности в многофакторных эконометрических моделях? Как это явление сказывается на качестве моделей и как оно устраняется?
3. Какие задачи экономического анализа решаются на основе эконометрических моделей регрессии?
4. Раскройте экономическую интерпретацию коэффициентов парной и множественной корреляции, коэффициентов детерминации, совокупных коэффициентов детерминации.
5. Поясните экономический смысл коэффициента эластичности и бета-коэффициента.
6. На основании каких коэффициентов можно проанализировать влияние отдельных факторов в линейных моделях множественной регрессии?
7. Каким образом может быть оценено качество линейных моделей регрессии?
8. Раскройте суть получения точечных и интервальных прогнозных значений результативного показателя на основе регрессионных моделей.
9. Поясните экономическое содержание составных элементов производственных функций и объясните происхождение названия этих функций. Какие виды производственных функций вы знаете? Приведите примеры производственных функций.

Упражнения

1. Данные опроса восьми групп семей о расходах на продукты питания в зависимости от уровня доходов семьи приведены в таблице (числа относительные в расчете на 100 руб. дохода и расхода):

Доходы семьи (x)	1,4	3,3	5,5	7,6	9,8	12,0	14,7	18,9
Расходы на продукты питания (y)	1,1	1,4	2,0	2,4	2,8	3,1	3,5	4,0

Требуется:

1) рассчитать коэффициент корреляции и оценить тесноту связи между доходами семьи и расходами на продукты питания;

2) построить линейную однофакторную модель зависимости расходов на питание от дохода семьи;

3) рассчитать коэффициент детерминации, коэффициент эластичности и бета-коэффициент и пояснить их экономический смысл;

4) найти среднюю по модулю относительную ошибку аппроксимации и оценить точность построенной регрессионной модели.

2. Результаты обследования 10 статистически однородных филиалов фирмы приведены в таблице (числа условные):

Номер филиала	Производительность труда (y)	Фондовооруженность (x_1)	Энерговооруженность (x_2)
1	74	33	56
2	84	34	58
3	73	36	67
4	93	35	70
5	56	33	73
6	71	37	77
7	117	39	78
8	111	42	99
9	135	43	93
10	125	44	96

Требуется:

1) рассчитать парные коэффициенты корреляции и пояснить их экономический смысл;

2) найти коэффициент множественной корреляции и совокупный коэффициент детерминации и охарактеризовать степень совместного влияния факторов фондовооруженности и энерговооруженности на производительность труда;

3) построить модель множественной линейной регрессии производительности труда от факторов фондо- и энерговооруженности;

4) рассчитать частные коэффициенты корреляции, детерминации, эластичности и частные бета-коэффициенты и с их помощью оценить влияние отдельных факторов (при неизменном значении других).

Глава 8
НЕКОТОРЫЕ ПРИКЛАДНЫЕ И ТЕОРЕТИЧЕСКИЕ МОДЕЛИ МИКРО- И МАКРОЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

— Моделирование спроса и потребления
— Модели управления запасами
— Моделирование систем массового обслуживания
— Элементы теории игр в задачах моделирования экономических процессов
— Динамические модели макроэкономики

После изучения материала этой главы студенты должны:

***знать*:**

- общие подходы к моделированию потребления и спроса и основные типы функций покупательского спроса;
- принципиальные системы и модели управления запасами;
- основные типы систем массового обслуживания (СМО) и методы расчета их характеристик;
- элементы теории игр и особенности применения этой теории в экономических исследованиях;
- структуру основных динамических макроэкономических моделей и системы предположений, лежащих в их основе;

***уметь*:**

- использовать конкретные функции покупательского спроса для анализа и прогнозирования этого спроса;
- решать задачи управления запасами при различных предположениях о размере заказа и его периодичности;
- находить основные характеристики СМО различных типов;
- применять элементы теории игр для принятия решения в условиях неопределенности и риска;
- формулировать основные виды макроэкономических моделей (Кейнса, Самуэльсона — Хикса, Солоу);

***владеть*:**

- понятийными аппаратами моделей спроса и потребления, управления запасами, систем массового обслуживания, а также основными элементами теории игр;
- навыками моделирования макроэкономических процессов.

8.1. Моделирование спроса и потребления

Целевая функция потребления и моделирование поведения потребителей

В условиях рыночной системы управления производственной и сбытовой деятельностью предприятий и фирм в основе принятия хозяйственных решений лежит рыночная информация, а обоснованность решений проверяется рынком в ходе реализации товаров и услуг. При таком подходе начальным пунктом всего цикла предпринимательской деятельности становится изучение потребительского спроса. Рассмотрим некоторые вопросы моделирования спроса и потребления.

Уровень удовлетворения материальных потребностей общества (уровень потребления) можно выразить целевой функцией потребления $U = U(Y)$, где вектор переменных $Y \geq 0$ включает разнообразные виды товаров и услуг. Ряд свойств этой функции удобно изучать, используя геометрическую интерпретацию уравнений $U(Y) = C$, где C — меняющийся параметр, характеризующий значение (уровень) целевой функции потребления; в качестве величины C может выступать, например, доход или уровень материального благосостояния.

В пространстве потребительских благ каждому уравнению $U(Y) = C$ соответствует определенная поверхность равноценных, или безразличных, наборов благ, которая называется *поверхностью безразличия*. Для наглядности рассмотрим пространство двух благ, например, в виде двух агрегированных групп товаров: продукты питания (y_1) и непродовольственные товары, включая услуги (y_2). Тогда уровни целевой функции потребления можно изобразить на плоскости в виде *кривых безразличия*, соответствующих различным значениям C (рис. 8.1, где $C_1 < C_2 < C_3$).

Будем далее пользоваться термином «кривые безразличия» вне зависимости от размерности пространства потребительских благ (количества групп товаров).

Из основных свойств целевой функции потребления отметим следующие:

1) функция $U(Y)$ является возрастающей функцией всех своих аргументов, т.е. увеличение потребления любого блага при неизменней уровне потребления всех других благ увеличивает значение данной функции. Поэтому более удаленная от начала координат кривая безразличия соответствует большему значению целевой функции потребления,

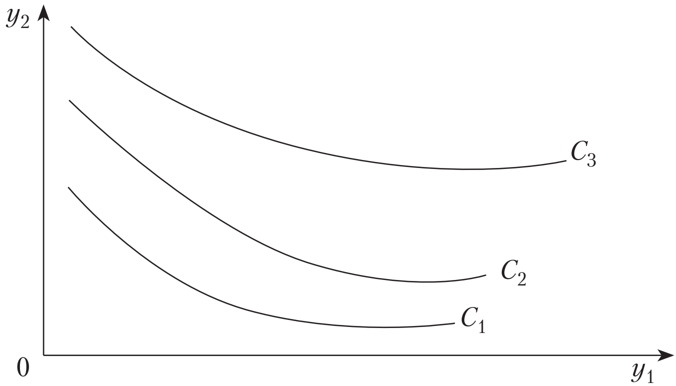


Рис. 8.1

а сам процесс максимизации этой функции на некотором ограниченном множестве допустимых векторов Y можно интерпретировать как нахождение допустимой точки, принадлежащей кривой безразличия, максимально удаленной от начала координат;

2) кривые безразличия не могут пересекаться, т.е. через одну точку пространства благ (товаров, услуг) можно провести только одну поверхность безразличия. В противном случае один и тот же набор благ одновременно соответствовал бы нескольким разным уровням материального благосостояния;

3) кривые безразличия имеют отрицательный наклон к каждой оси координат, при этом абсолютный наклон кривых уменьшается при движении в положительном направлении по каждой оси, т.е. кривые безразличия являются выпуклыми кривыми.

Методы построения целевой функции потребления основаны на обобщении опыта поведения потребителей и тенденций покупательского спроса в зависимости от уровня благосостояния. В качестве примера приведем квадратичную целевую функцию потребления для трех агрегированных групп товаров, построенную на основе обработки данных бюджетной статистики:

$$U(Y) = (1-1{,}841a)y_1 + (1-2{,}054a)y_2 + (1-2{,}11a)y_3 +$$
$$+0{,}668\cdot 10^{-4} y_1^2 + 1{,}230\cdot 10^{-4} y_1 y_2 + 1{,}234\cdot 10^{-4} y_1 y_3 +$$
$$+0{,}506\cdot 10^{-4} y_2^2 + 1{,}104\cdot 10^{-4} y_2 y_3 + 0{,}492\cdot 10^{-4} y_3^2,$$

где:

параметр a означает число детей в семье;

y_1 — потребление продуктов питания;

y_2 — потребление промышленных товаров;

y_3 — потребление платных услуг (в стоимостном выражении).

Перейдем к вопросу моделирования поведения потребителей в условиях товарно-денежных отношений на базе целевой функции потребления. В основе модели поведения потребителей лежит гипотеза, что потребители, осуществляя выбор товаров при установленных ценах и имеющемся доходе, стремятся максимизировать уровень удовлетворения своих потребностей.

Пусть в пространстве n видов товаров исследуется поведение совокупности потребителей. Обозначим спрос потребителей через вектор $Y = (y_1, y_2, ..., y_n)$, а цены на различные товары — через вектор $P = (p_1, p_2, ..., p_n)$. При величине дохода D потребители могут выбирать только такие комбинации товаров, которые удовлетворяют *бюджетному ограничению* $\sum_{i=1}^{n} p_i y_i \le D$. Предположим, что предпочтение потребителей на множестве товаров выражается целевой функцией потребления $U(Y)$. Тогда простейшая модель поведения потребителей в векторной форме записи будет иметь вид:

$$\begin{aligned} &U(Y) \to \max; \\ &PY \le D; \\ &Y \ge 0. \end{aligned} \tag{8.1}$$

Геометрическая интерпретация модели (8.1) для двух агрегированных групп товаров представлена на рис. 8.2.

Линия AB (в других вариантах A_1B_1, A_2B_2) соответствует бюджетному ограничению и называется *бюджетной линией* (см. более подробно в параграфе 8.2). Выбор потребителей ограничен треугольником AOB (A_1OB_1, A_2OB_2). Набор товаров M, соответствующий точке касания прямой AB с наиболее отдаленной кривой безразличия, является оптимальным решением (в других вариантах это точки K и L). Легко заметить, что линии AB и A_1B_1 соответствуют одному и тому же размеру дохода и разным ценам на товары y_1 и y_2; линия A_2B_2 соответствует большему размеру дохода.

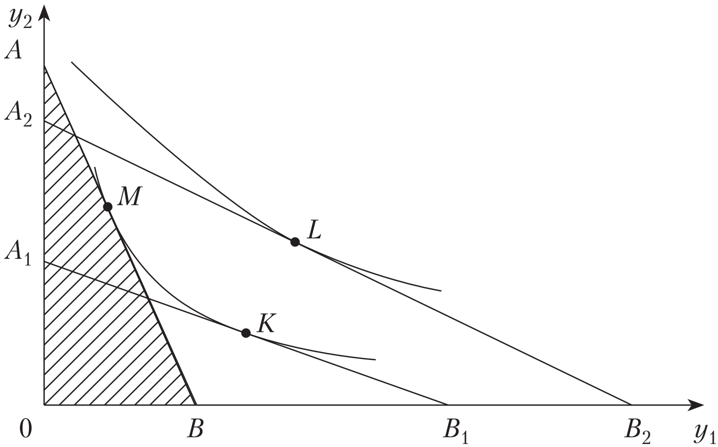


Рис. 8.2

Опираясь на некоторые выводы теории нелинейного программирования, можно определить математические условия оптимальности решений для модели (8.1). С задачей нелинейного программирования связывается так называемая функция Лагранжа, которая для задачи (8.1) имеет вид

$$L(Y,\lambda)=U(Y)+\lambda(D-PY),$$

где множитель Лагранжа λ является оптимальной оценкой дохода.

Обозначим частные производные функции $U(Y)$ через U_i: $U_i=\dfrac{\partial U(Y)}{\partial y_i}$. Эти производные интерпретируются как предельные полезные эффекты (*предельные полезности*) соответствующих потребительских благ и характеризуют прирост целевой функции потребления при увеличении использования i-го блага (товара) на некоторую условную «малую единицу».

Необходимыми условиями того, что вектор Y^0 будет оптимальным решением, являются условия Куна — Таккера:

$$U_i(Y^0)\le\lambda^0 p_i,\quad i=\overline{1,n},$$

при этом

$$\begin{aligned}&U_i(Y^0)=\lambda^0 p_i,\ \text{если}\ y_i^0>0\ \text{(товар приобретается)},\\&U_i(Y^0)<\lambda^0 p_i,\ \text{если}\ y_i^0=0\ \text{(товар не приобретается)},\\&PY^0=D.\end{aligned}\qquad(8.2)$$

Последнее из соотношений (8.2) соответствует полному использованию дохода, и для этого случая очевидно неравенство $\lambda^0 > 0$. Из условий оптимальности (8.2) следует, что

$$\frac{U_i(Y^0)}{p_i} = \lambda^0, \quad y_i^0 > 0.$$

Это означает, что потребители должны выбирать товары таким образом, чтобы отношение предельной полезности к цене товара было одинаковым для всех приобретаемых товаров. Другими словами, в оптимальном наборе предельные полезности выбираемых товаров должны быть пропорциональны ценам.

Функции покупательского спроса

Функциями покупательского спроса (далее будем называть их просто *функциями спроса*) называются функции, отражающие зависимость объема спроса на отдельные товары и услуги от комплекса факторов, влияющих на него. Такие функции применяются в аналитических моделях спроса и потребления и строятся на основе информации о структуре доходов населения, ценах на товары, составе семей и других факторах. Рассмотрим построение функций спроса в зависимости от двух факторов — дохода и цен.

Пусть в модели (8.1) цены и доход рассматриваются как меняющиеся параметры. Переменную дохода будем обозначать Z. Тогда решением оптимизационной задачи (8.1) будет векторная функция $Y^0 = Y^0(P, Z)$, компонентами которой являются функции спроса на определенный товар от цен и дохода:

$$y_i^0 = f_i(p, Z).$$

Рассмотрим частный случай, когда вектор цен остается неизменным, а изменяется только доход. Для двух товаров этот случай представлен на рис. 8.3. Если по оси абсцисс отложить количество единиц товара y_1, которое можно приобрести на имеющийся доход Z (точка B), а по оси ординат — то же самое для товара y_2 (точка A), то прямая линия AB, называемая бюджетной линией, показывает любую комбинацию количеств этих двух товаров, которую можно купить за сумму денег Z. При увеличении дохода бюджетные линии перемещаются параллельно самим себе, удаляясь от начала координат. Вместе с ними перемещаются соответствующие кривые безразличия. Точками оптимума спроса потребите-

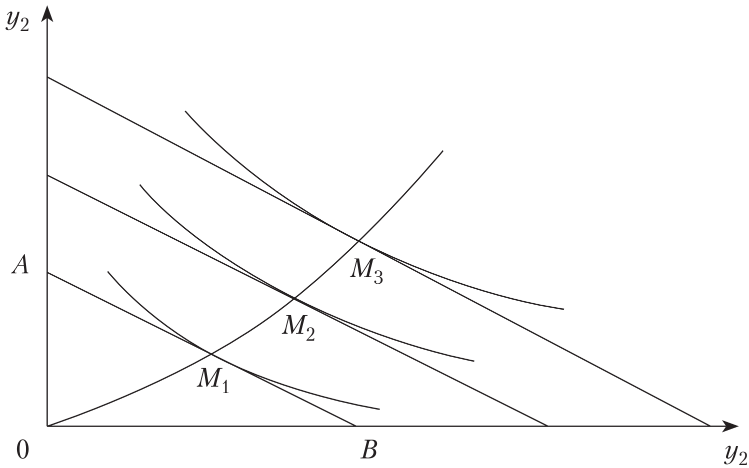


Рис. 8.3

лей для соответствующих размеров дохода будут в данном случае точки касания M_1, M_2, M_3. При нулевом доходе спрос на оба товара нулевой. Кривая, соединяющая точки O, M_1, M_2, M_3, является графическим отображением векторной функции спроса от дохода при заданном векторе цен.

Пример 8.1. Рассмотрим процесс аналитического построения функций спроса от дохода на основе модели (8.1) на конкретном условном примере. Пусть для двух товаров целевая функция потребления имеет вид $U(Y) = y_1 y_2^3;$ вектор цен равен $P = (3; 6)$; величина дохода равна Z. Так как в данной случае предельные полезности имеют вид:

$$U_1 = \frac{\partial U(Y)}{\partial y_1} = y_2^3; \quad U_2 = \frac{\partial U(Y)}{\partial y_2} = 3y_1 y_2^2; \quad D = Z,$$

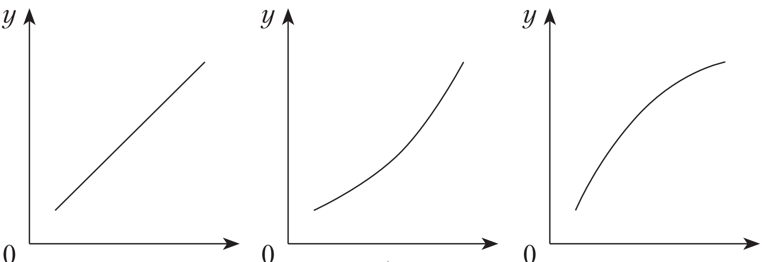
то необходимые условия оптимума (8.2) дают следующую систему уравнений (λ — множитель Лагранжа):

$$y_2^3 = 3\lambda;$$
$$3y_1 y_2^2 = 6\lambda;$$
$$3y_1 + 6y_2 = Z.$$

После подстановки первого уравнения во второе получим $3y_1 y_2^2 = 2y_2^3$. Выразив из третьего уравнения $3y_1$ и подставив в последнее равенство, будем иметь $(Z - 6y_2)y_2^2 = 2y_2^3$, откуда можно получить, что $y_2 = 1/8Z$. Подставив этот результат в третье уравнение, получим $y_1 = 1/12Z$. Таким образом, для данного примера функции спроса на товары y_1 и y_2 от дохода Z имеют вид:

$$y_1 = 1/12\,Z;\; y_2 = 1/8\,Z.$$

Однофакторные функции спроса от дохода широко применяются при анализе покупательского спроса. Соответствующие этим функциям кривые $y_i = f_i\,(Z)$ называются *кривыми Энгеля* (по имени изучавшего их немецкого экономиста). Формы этих кривых для различных товаров могут быть различны. Если спрос на данный товар возрастает примерно пропорционально доходу, то функция будет линейной, как в рассмотренном выше примере. Такой характер имеет, например, спрос на одежду, фрукты и др. Кривая Энгеля для этого случая представлена на рис. 8.4, *а*.



а) *б*) *в*)

Рис. 8.4

Если по мере роста дохода спрос на данную группу товаров возрастает все более высокими темпами, то кривая Энгеля будет выпуклой (рис. 8.4, *б*). Так ведет себя спрос на предметы роскоши.

Если рост значений спроса, начиная с определенного момента, по мере насыщения спроса отстает от роста дохода, то кривая Энгеля имеет вид вогнутой кривой (рис. 8.4, *в*). Например, такой характер имеет спрос на товары первой необходимости.

Тот же принцип разграничения групп товаров по типам функций спроса от дохода использовал шведский экономист Л. Торнквист, который предложил специальные виды функции спроса (*функции Торнквиста*) для трех групп товаров: первой необходимости, второй необходимости, предметов роскоши.

Функция Торнквиста для *товаров первой необходимости* имеет вид

$$y = \frac{a_i Z}{Z + C_1}$$

и отражает тот факт, что рост спроса на эти первоочередные товары с ростом дохода постепенно замедляется и имеет предел a_1 (кривая спроса асимптотически приближается к прямой линии $y = a_1$); график функции является вогнутой кривой I на рис. 8.5.

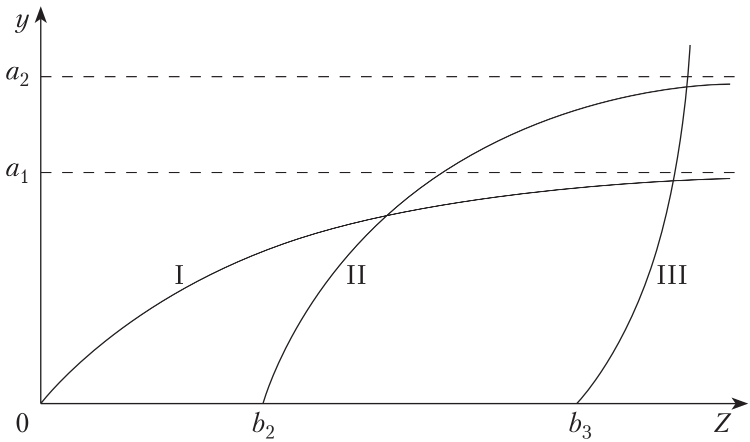


Рис. 8.5

Функция спроса по Торнквисту на *товары второй необходимости* выражается формулой

$$y = \frac{a_2(Z - b_2)}{Z + C_2}, \quad \text{где } Z \geq b_2.$$

Эта функция также имеет предел a_2, но более высокого уровня; при этом спрос на эту группу товаров появляется лишь после того, как доход достигнет величины b_2; график функции — вогнутая кривая II на рис. 8.5.

Наконец, функция Торнквиста *для предметов роскоши* имеет вид

$$y = \frac{a_3 Z(Z - b_3)}{Z + C_3}, \quad \text{где } Z \geq b_3.$$

Эта функция не имеет предела. Спрос на эти товары возникает только после того, как доход превысит величину b_3 и далее быстро возрастает, так что график функции — выпуклая кривая III на рис. 8.5.

Кроме указанных функций, в аналитических моделях покупательского спроса используются также другие функции: степенные, S-образные и т.д.

Важную роль в анализе изменения спроса при небольших изменениях дохода играют коэффициенты эластичности. *Коэффициент эластичности спроса от дохода* показывает относительное изменение спроса при изменении дохода (при прочих не изменяющихся факторах). Вычисляется по формуле

$$E_i^Z = \frac{\partial y_i}{\partial Z} \cdot \frac{Z}{y_i}, \qquad (8.3)$$

где:

E_i^Z — коэффициент эластичности для i-го товара (группы товаров) по доходу Z;

y_i — спрос на этот товар, являющийся функцией дохода: $y_i = f(Z)$.

Например, если спрос на товар описывается функцией Торнквиста для товаров первой необходимости, то формула (8.3) дает следующее выражение для коэффициента эластичности спроса от дохода:

$$E_i^Z = \frac{C_1}{Z + C_1}.$$

Во многих экономико-математических моделях эластичность функций относят к проценту прироста независимой переменной. Таким образом, коэффициент эластичности спроса от дохода показывает, на сколько процентов изменится спрос на товар при изменении дохода на 1%.

Коэффициенты эластичности спроса от дохода различны по величине для разных товаров, вплоть до отрицательных значений, когда с ростом доходов потребление уменьшается. Принято выделять четыре группы товаров в зависимости от коэффициента эластичности спроса на них от дохода:

— малоценные товары ($E_i^Z < 0$);

— товары с малой эластичностью ($0 < E_i^Z < 1$);

— товары со средней эластичностью (E_i^Z близки к единице);

— товары с высокой эластичностью ($E_i^Z > 1$).

К малоценным товарам, т.е. товарам с отрицательной эластичностью спроса от дохода, относятся такие, как хлеб, низкосортные товары. По результатам обследований, коэффициенты эластичности для основных продуктов питания находятся в интервале от 0,4 до 0,8, по одежде, тканям, обуви — в интервале от 1,1 до 1,3 и т.д. По мере увеличения дохода спрос перемещается с товаров первой и второй групп на товары третьей и четвертой групп, при этом потребление товаров первой группы по абсолютным размерам сокращается.

Перейдем к рассмотрению и анализу функций покупательского спроса от цен на товары. Из модели поведения потребителей (8.1) следует, что спрос на каждый товар в общем случае зависит от цен на все товары (вектора P), однако построить функции общего вида $y_i = \varphi_i(P)$ очень сложно. Поэтому в практических исследованиях ограничиваются построением и анализом функций спроса для отдельных товаров в зависимости от изменения цен на этот же товар или группу взаимозаменяемых товаров: $y_i = \varphi_i(p_i)$.

Рассмотрим вопросы построения и использования в задачах оптимизации функций покупательского спроса от цены при неизменном доходе для конкретной группы товаров $y = \varphi(p)$ более подробно. Основным критерием функционирования предприятий и фирм, производящих тот или иной товар, является прибыль. Сформулируем экономико-математическую модель на максимум прибыли при указанной форме зависимости спроса от цены. Конечную прибыль П можно представить в виде

$$\Pi = p \cdot y - C(y), \tag{8.4}$$

где $C(y)$ — совокупные затраты на производство и реализацию продукции; при этом предполагается, что объем выпускаемой продукции целиком совпадает с объемом спроса y, т.е. отсутствуют дефицит и затоваривание продукции.

Если известен вид функции спроса от цены $y = \varphi(p)$, то, подставив эту функцию в (8.4), получим:

$$\Pi(p) = p - \varphi(p) - C(\varphi(p)). \tag{8.5}$$

Тогда оптимизационная задача ставится следующим образом: определить значение цены $p = p_0$, при котором прибыль достигает своего максимума $\max \Pi(p)$. В такой постановке эта задача решается классическими методами оптимизации. Вид функций $\varphi(p)$ и $C(\varphi((p))$, а также их параметры могут быть определены методами математической статистики в результате обработки соответствующих статистических данных.

Рассмотрим случай, когда функция издержек $C(y)$ является линейной и имеет вид

$$C(y) = h + v \cdot y. \tag{8.6}$$

Здесь h есть постоянная величина затрат, не зависящая от объема выпуска спроса продукции, т.е. h — это накладные расходы; величина v задает удельные затраты на единицу продукции. Будем считать, что зависимость спроса от цены также является линейной:

$$y = k + E \cdot p, \tag{8.7}$$

где k характеризует максимально возможный объем спроса, а E представляет собой некоторую условную эластичность спроса от цены; очевидно, что $k > 0$, а $E < 0$.

Подставляя (8.6) и (8.7) в выражение для прибыли (8.5), будем иметь:

$$\begin{aligned}\Pi(p) &= p(k+Ep) - h - v(k+Ep) = Ep^2 + (k - vE)p - (h + vk) = \\ &= E\left(p - \frac{vE-k}{2E}\right)^2 - \frac{(vE-k)^2}{4E} - h - vk.\end{aligned} \tag{8.8}$$

С учетом того что $E < 0$, можно сделать вывод о том, что максимум прибыли достигается при

$$p_0 = \frac{vE-k}{2E}, \tag{8.9}$$

и этот максимум равен

$$\max \Pi(p) = \Pi(p_0) = \frac{(vE-k)^2 + (h+vk)4E}{-4E}. \tag{8.10}$$

График функции (8.8) представляет собой параболу, ветви которой направлены вниз. С учетом требования рентабельности выпуска продукции $\max\,\Pi(p) > 0$ этот график представлен на рис. 8.6.

Условием того, что эта рентабельность может быть достигнута, является условие положительности дискриминанта квадратного трехчлена от p (8.8):

$$(k - vE)^2 + 4E(h + vk) > 0.$$

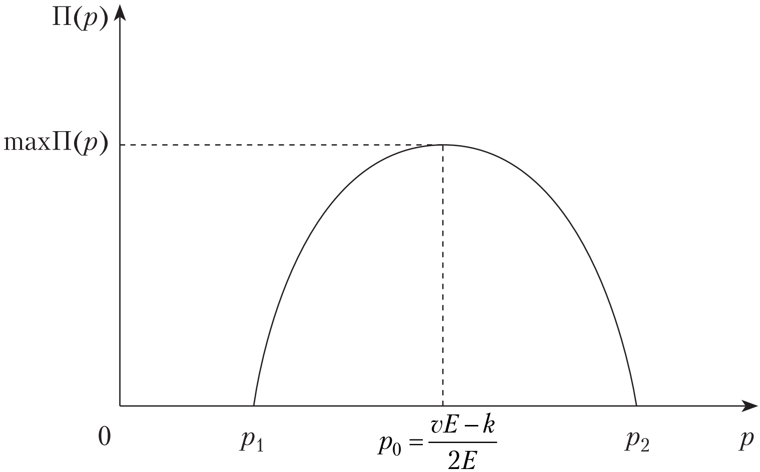


Рис. 8.6

Из этого неравенства следует, что

$$(k+vE)^2 > -4Eh,$$

вследствие чего условие рентабельности можно записать в виде

$$k+vE > 2\sqrt{-Eh}. \tag{8.11}$$

С учетом (8.9) оптимальный объем выпуска продукции будет равен

$$y_0 = k + Ep_0 = \frac{vE+k}{2}; \tag{8.12}$$

в соответствии с этим требование рентабельности (8.11) может быть записано в виде

$$y_0 > \sqrt{-Eh}. \tag{8.13}$$

Как видно из графика функции $\Pi(p)$ на рис. 8.6, существуют две «мертвые точки» p_1 и p_2, в которых прибыль равна нулю. Решение соответствующего уравнения дает для этих точек:

$$p_{1,2} = \frac{(vE-k) \pm \sqrt{(vE+k)^2 + 4Eh}}{2E}. \tag{8.14}$$

Таким образом, если установление оптимальной цены затруднено, то цена в любом случае должна лежать в интервале (p_1, p_2), этим обеспечивается положительность величины прибыли от производства и реализации данного вида товара.

Пример 8.2. Рассмотрим числовой пример из области маркетинга книгоиздательства.

Пусть зависимость объема спроса на книжную продукцию от цены имеет вид (8.7), а функция издержек представлена в виде (8.6). Зададим следующие условные значения величин, участвующих в этих зависимостях:

— накладные расходы на производство и реализацию книжной продукции h =36 тыс. руб.;

— удельные затраты на единицу (один экземпляр) этой продукции v = 20 руб./экз.;

— максимально возможный объем тиража k = 4000 экз.;

— условная эластичность спроса от цены E = –40 экз./руб.

Тогда требование к объему тиража для обеспечения рентабельности определяется условием (8.13):

$$y_0 > \sqrt{-(-40)\cdot 36\,000} = 1200 \text{ экз.}$$

Оптимальная цена задается формулой (8.9):

$$p_0 = \frac{20\cdot(-40)-4000}{2\cdot(-40)} = 60 \text{ руб./экз.}$$

Наиболее оптимальный объем тиража, обеспечивающий максимум прибыли, в соответствии с формулой (8.12) равен

$$y_0 = \frac{20\cdot(-40)+4000}{2} = 1600 \text{ экз.}$$

Интервал для цены, в пределах которого обеспечивается рентабельность данного издания, определяется выражением (8.14):

$$p_{1,2} = 60 \pm \frac{\sqrt{(20\cdot(-40)+4000)^2+4\cdot(-40)\cdot 36\,000}}{2\cdot(-40)} = 60 \pm 26{,}5 \text{ руб./экз.}$$

Таким образом, цена должна лежать в пределах от 33,5 до 86,5 руб. за один экземпляр; оптимальная цена равна 60 руб./экз.

Максимум прибыли, достигаемый при этом, рассчитывается в соответствии с формулой (8.10) и равен

$$\max \Pi(p) = \frac{(20\cdot(-40)-4000)^2+(36\,000+20\cdot 4000)\cdot 4\cdot(-40)}{-4\cdot(-40)} =$$

$$= 28 \text{ тыс. руб.}$$

Представление зависимости спроса от цены в виде линейной функции (8.7) является весьма упрощенным, так как показатель эластичности спроса от цены здесь принимается равным постоянной величине E. Более соответствует действительности предположение о зависимости коэффициента эластичности спроса от цены от величины самой цены. Проведенные конкретные исследования, учитывающие эту зависимость, позволяют с большой долей вероятности принять зависимость спроса y от цены p в виде одной из S-образных убывающих кривых, а именно в виде логистической кривой (функция Перла – Рида; см. параграф 5.1):

$$y = \frac{k}{1 + b \cdot e^{\alpha p}}. \qquad (8.15)$$

График этой функции представлен на рис. 8.7; он имеет точку симметрии, совпадающую с точкой перегиба функции (8.15).

Тогда, предполагая зависимость затрат от спроса $C(y)$ по-прежнему линейной в виде функции (8.6), получим выражение для прибыли:

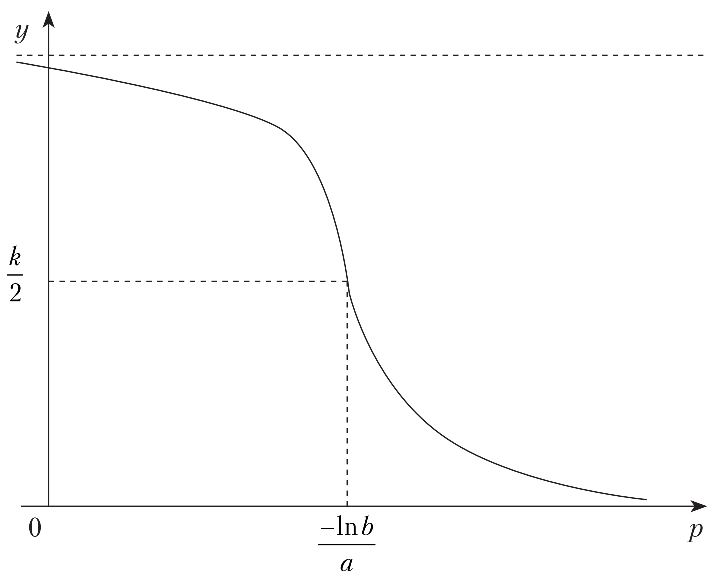


Рис. 8.7

$$\Pi(p) = p \cdot \frac{k}{1+b \cdot e^{\alpha p}} - h - v \cdot \frac{k}{1+b \cdot e^{\alpha p}} = \frac{k(p-v)}{1+b \cdot e^{\alpha p}} - h. \quad (8.16)$$

Вычисляя производную этой функции по переменной p и приравнивая ее нулю, получим уравнение для нахождения критических точек:

$$1 + b \cdot e^{\alpha p} - (p - v) \cdot ab \cdot e^{\alpha p} = 0.$$

Разделив обе части этого уравнения на $e^{\alpha p} \neq 0$, получим:

$$e^{-\alpha p} + b(av+1) = abp. \quad (8.17)$$

Уравнение (8.17) имеет единственный корень p_0, так как кривые функций $y = abp$ и $y = e^{-\alpha p} + b(av+1)$ имеют только одну точку пересечения (рис. 8.8). Трансцендентное уравнение (8.17) можно решить различными численными методами вычислительной математики. Решение этого уравнения p_0 можно оценить:

$$av + \frac{b}{a} < p_0 < v + \frac{b}{a} + \frac{1}{ab}. \quad (8.18)$$

Легко показать, что вторая производная функции прибыли от цены (8.16) в точке $p = p_0$ отрицательна; отсюда

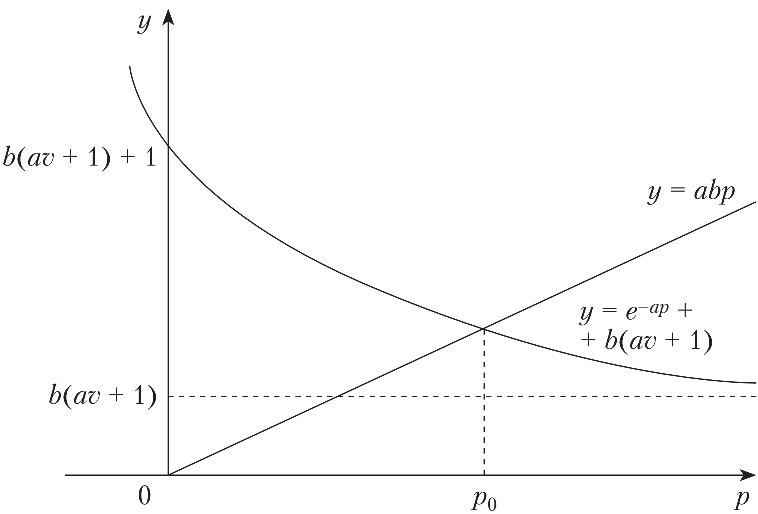


Рис. 8.8

следует, что точка p_0 действительно является точкой максимума этой функции.

Из соотношения (8.18) можно получить интервал для оптимального объема выпуска товара:

$$\frac{k}{1+b\cdot e^{\alpha v+b+\frac{1}{b}}}<y_0<\frac{k}{1+b\cdot e^{\alpha v+b}}. \tag{8.19}$$

Положительные значения параметров a и b логистической кривой (8.15), если значение асимптоты k известно, определяются путем предварительного преобразования и логарифмирования, после чего на основе метода наименьших квадратов составляется и решается система нормальных уравнений относительно логарифмов параметров, а затем находятся и сами параметры. Если значение величины k (максимум выпуска) неизвестно, то для нахождения параметров k, a и b кривой (8.15) применяют приближенные методы: метод трех точек, метод трех сумм и др. Исходные данные для проведения указанных выше расчетов формируются на основе статистических наблюдений за процессом производства и реализации товарной продукции.

Для большинства товаров действует зависимость: чем выше цена, тем ниже спрос, и наоборот. Здесь также возможны разные типы зависимости и, следовательно, разные формы кривых. В практических задачах изучения спроса важно различать действительное увеличение спроса, когда сама кривая сдвигается вверх и вправо (происходит переход с кривой I на кривую II на рис. 8.9), и увеличение объема приобретаемых товаров в результате снижения цен при неизменной сумме затрат (переход от точки A к точке B по одной и той же кривой I на рис. 8.9). Как уже отмечено выше, в общем случае спрос на отдельный товар при прочих равных условиях зависит от уровня цен всех товаров. Относительное изменение объема спроса при изменении цены данного товара или цен других связанных с ним товаров характеризует *коэффициент эластичности спроса от цен*. Этот коэффициент эластичности удобно трактовать как величину изменения спроса в процентах при изменении цены на 1%.

Для спроса y_i на i-й товар относительно его собственной цены p_i коэффициент эластичности исчисляется по формуле

$$E_{ii}^{p}=\frac{\partial y_i}{\partial p_i}\cdot\frac{p_i}{y_i}. \tag{8.20}$$

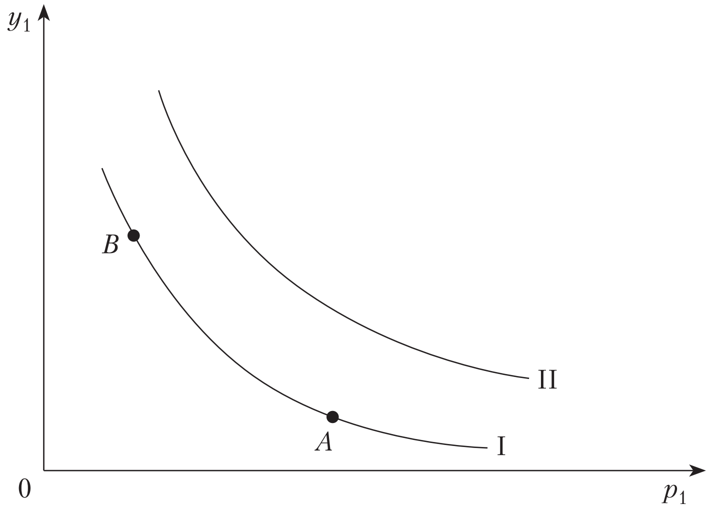


Рис. 8.9

Значения коэффициентов эластичности спроса от цен практически всегда отрицательны. Однако по абсолютным значениям этих коэффициентов товары могут существенно различаться друг от друга. Их можно разделить на три группы:

— товары с неэластичным спросом в отношении цены $E_{ii}^{p} > -1$;

— товары со средней эластичностью спроса от цены (E_{ii}^{p} близки к -1);

— товар с высокой эластичностью спроса ($E_{ii}^{p} < -1$;).

В товарах *эластичного спроса* повышение цены на 1% приводит к снижению спроса более чем на 1% и, наоборот, понижение цены на 1% приводит к росту покупок более чем на 1%. Если повышение цены на 1% влечет за собой понижение спроса менее чем на 1%, то говорят, что этот товар *неэластичного спроса.*

Рассмотрим влияние на спрос на какой-либо товар изменения цен на другие товары. Коэффициент, показывающий, на сколько процентов изменится спрос на данный товар при изменении на 1% цены на другой товар при условии, что другие цены и доходы покупателей остаются прежними, называется *перекрестным коэффициентом эластичности*. Для спроса y_i на i-й товар относительно цены p_j на j-й товар ($i \neq j$) перекрестный коэффициент эластичности рассчитывается по формуле

$$E_j^p = \frac{\partial y_i}{\partial p_j} \cdot \frac{p_j}{y_i}. \tag{8.21}$$

По знаку перекрестных коэффициентов эластичности товары можно разделить на *взаимозаменяемые* и *взаимодополняемые*. Если $E_{ij}^p > 0$, это означает, что i-й товар заменяет в потреблении товар j, т.е. на товар i переключается спрос при увеличении цены на товар j. Примером взаимозаменяемых товаров могут служить многие продукты питания.

Если $E_{ij}^p < 0$, это служит признаком того, что i-й товар в процессе потребления дополняет товар j, т.е. увеличение цены на товар j приводит к уменьшению спроса на товар i. В качестве примера можно привести такие взаимодополняемые товары, как автомобили и бензин.

В качестве иллюстрации в табл. 8.1 приведены значения прямых и перекрестных коэффициентов эластичности потребления от цен для одной из категорий семей. На основании этих данных по значениям прямых коэффициентов эластичности (по диагонали таблицы) можно сделать вывод, что продукты питания в целом мало эластичны по отношению к цене; одежда, ткани и обувь имеют среднюю эластичность; две последние группы товаров — товары с высокой эластичностью спроса по отношению к цене.

Таблица 8.1

Прямые и перекрестные коэффициенты эластичности

Группы товаров	Продукты питания	Одежда, ткани, обувь	Мебель, хозтовары	Культ-товары
Продукты питания	−0,7296	0,0012	0,0043	0,0045
Одежда, ткани, обувь	−0,1991	−1,0000	0,0071	0,0074
Мебель, хозтовары	−0,2458	0,0024	−1,2368	0,0092
Культтовары	−0,2494	0,0024	0,0089	−1,2542

На основании значений внедиагональных элементов этой таблицы можно сделать вывод, что все промышленные товары (вторая, третья и четвертая группы) — взаимозаменяемы. То, что перекрестные коэффициенты эластичности по строке «Продукты питания» положительные, означает, что повышение цен на промышленные товары увеличивает спрос на продукты питания (уменьшение спроса на промышленные товары освободит средства для продуктов питания). Отрицательные значения перекрестных коэффициентов эластичности по столбцу (графе) «Продукты питания» означает, что при росте цен на продукты питания спрос на промышленные товары сокращается (повышение цен на продукты питания уменьшает размер средств на приобретение других товаров).

Моделирование и прогнозирование покупательского спроса

Очевидно, что спрос во многом определяет стратегию и тактику организации производства и сбыта товаров и услуг. Учет спроса, обоснованное прогнозирование его на краткосрочную и долгосрочную перспективу — одна из важнейших задач служб маркетинга различных организаций и фирм.

Состав и уровень спроса на тот или иной товар зависят от многих факторов, как экономических, так и естественных. К экономическим факторам относятся уровень производства (предложения) товаров и услуг (обозначим этот фактор в общем виде Π), уровень денежных доходов отдельных групп населения (D), уровень и соотношение цен (P). К естественным факторам относятся демографический состав населения, в первую очередь размер и состав семьи (S), а также привычки и традиции, уровень культуры, природно-климатические условия и т.д.

Экономические факторы очень мобильны, особенно распределение населения по уровню денежных доходов. Естественные же факторы меняются сравнительно медленно и в течение небольшого периода (до 3—5 лет) не оказывают заметного влияния на спрос. Исключение составляет демографический состав населения. Поэтому в текущих и перспективных прогнозах спроса все естественные факторы, кроме демографических, целесообразно учитывать сообща, введя фактор под названием «время» (t).

Таким образом, в общем виде спрос определяется в виде функции перечисленных выше факторов:

$$y = f(\Pi, D, P, S, t). \tag{8.22}$$

Поскольку наибольшее влияние на спрос оказывает фактор дохода (известно выражение «спрос всегда платежеспособен»), многие расчеты спроса и потребления осуществляются в виде функции от душевого денежного дохода: $y = f(D)$.

Наиболее простой подход к прогнозированию спроса на небольшой период времени связан с использованием так называемых *структурных моделей спроса*. При построении модели исходят из того, что для каждой экономической группы населения по статистическим бюджетным данным может быть рассчитана присущая ей структура потребления. При этом предполагается, что на изучаемом отрезке времени заметные изменения претерпевает лишь доход, а цены, размер семьи и прочие факторы принимаются неизменными. Изменение дохода, например его рост, можно рассматривать как перемещение определенного количества семей из низших доходных групп в высшие. Другими словами, изменяются частоты в различных интервалах дохода: они уменьшаются в нижних и увеличиваются в верхних интервалах. Семьи, которые попадают в новый интервал, будут иметь ту же структуру потребления и спроса, что и у семей с таким же доходом к настоящему времени.

Таким образом, структурные модели рассматривают спрос как функцию только распределения потребителей по уровню дохода. Имея соответствующие структуры спроса, рассчитанные по данным статистики бюджетов, и частоты распределения потребителей по уровню дохода, можно рассчитать общую структуру спроса. Если обозначить структуру спроса в группе семей со средним доходом D_i через $r(D_i)$, а частоты семей с доходом D_i через $w(D_i)$, то общая структура спроса R может быть рассчитана по формуле

$$R = \sum_{i=1}^{n} r(D_i) \cdot w(D_i), \tag{8.23}$$

где n — количество интервалов дохода семей.

Структурные модели спроса — один из основных видов экономико-математических моделей планирования и прогнозирования спроса и потребления. В частности, широко распространены так называемые *компаративные* (сравнительные) *структурные модели*, в которых сопоставляются структуры спроса данного исследуемого объекта и некоторого аналогового объекта. Аналогом обычно считаются регион или группа населения с оптимальными потребительскими характеристиками.

Наряду со структурными моделями в планировании и прогнозировании спроса используются *конструктивные модели спроса*. В основе их лежат уравнения бюджета населения, т.е. такие уравнения, которые выражают очевидное равенство общего денежного расхода (другими словами, объема потребления) и суммы произведений количества каждого потребленного товара на его цену. Если Z — объем потребления, m — количество разных видов благ, q_i — размер потребления i-го блага, p_i — цена i-го блага, то конструктивная модель спроса может быть записана следующим образом:

$$Z = \sum_{i=1}^{m} q_i p_i. \tag{8.24}$$

Эти модели, называемые также *моделями бюджетов потребителей*, играют важную роль в планировании потребления. Одной из таких моделей является, например, всем известный прожиточный минимум. К таким моделям относятся также рациональные бюджеты, основанные на научных нормах потребления, прежде всего продуктов питания, перспективные бюджеты (например, так называемый бюджет достатка) и др.

В практике планирования и прогнозирования спроса кроме структурных и конструктивных моделей применяются также *аналитические модели спроса и потребления*, которые строятся в виде уравнений, характеризующих зависимость потребления товаров и услуг от тех или иных факторов. В аналитических моделях функциональная зависимость (8.22) принимает вполне определенный вид. Такие модели могут быть однофакторными и многофакторными. Аналитические модели спроса на примере линейных корреляционно-регрессионных статических моделей рассмотрены в параграфе 7.2.

8.2. Модели управления запасами

Классическая задача управления запасами

Под задачей управления товарными запасами понимается такая оптимизационная задача, в которой задана информация:

— о поставках товара;
— о спросе на товар;
— об издержках и условиях хранения товарных запасов;
— критерий оптимизации.

Рассмотрим задачу управления запасами в так называемой *классической постановке*. Выберем за единичный интервал времени один день. Пусть к концу дня $t-1$ на складе находится запас товара в объеме $x_{t-1} \geq 0$. Склад делает заказ на пополнение запаса товара в объеме h_t. Это пополнение поступает к началу следующего дня t, так что запас товара в начале следующего дня составляет $x_{t-1} + h_t$. Пусть s_t – объем товара, запрашиваемый потребителем (или потребителями) в день t (объем заявки).

Если $s_t \leq x_{t-1} + h_t$, то склад удовлетворяет заявку потребителя полностью, а остатки товара $x_t = x_{t-1} + h_t - s_t$ переносятся на следующий день $t+1$, причем издержки хранения этого запаса пропорциональны его объему и равны

$$cx_t = c(x_{t-1} + h_t - s_t).$$

Если объем заявки $s_t > x_{t-1} + h_t$, то склад полностью отдает свой запас, а за недопоставленный товар несет потери (например, штрафуется за дефицит), пропорциональные объему дефицита и равные

$$k(s_t - x_{t-1} - h_t) = -k\,(x_{t-1} + h_t - s_t).$$

Таким образом, полные издержки $\varphi\,(x_{t-1}, h_t, s_t)$ склада можно записать в виде:

$$\varphi(x_{t-1}, h_t, s_t) = \max\left\{c(x_{t-1} + h_t - s_t);\quad -k(x_{t-1} + h_t - s_t)\right\}, \quad (8.25)$$

при этом остаток товара $x_t = \max\{0; x_{t-1} + h_t - s_t)$.

Из (8.25) следует:
если $x_t > 0$, то $\varphi\,(x_t) = cx_t$;
если $x_t < 0$, то $\varphi\,(x_t) = -kx_t$;
если $x_t = 0$, то $\varphi\,(x_t) = 0$.

В классической постановке задачи управления запасами предполагается, что сама величина спроса s_t неизвестна, однако известно, что она является независимой случайной величиной, имеющей заданный закон распределения. Пусть распределение вероятностей величины s_t задается непрерывной функцией распределения $F(x)$ с плотностью распределения $f(x)$. Тогда средние полные издержки $\Phi(x_{t-1} + h_t)$ задаются формулой (M — математическое ожидание)

$$\Phi(x_{t-1}+h_t)=M\varphi(x_{t-1},h_t,s_t)=\int\limits_{-\infty}^{+\infty}\varphi(x_{t-1},h_t,s_t)\alpha F(s_t).$$

Задача ставится таким образом: определить объем заказа на пополнение h_t, минимизирующий средние полные издержки, т.е.

$\Phi(x_{t-1}+h_t) \to \min$, где $h_t \geq 0$.

Рассмотрим решение классической задачи управления товарными запасами в статической постановке. Обозначим $y = (x_{t-1} + h_t)$ и опустим ввиду статичности задачи индекс t в записи объема спроса и пополнения. Рассмотрим следующую задачу:

$$\Phi(y)=\int\limits_{-\infty}^{+\infty}\max\left\{c(y-s);-k(y-s)\right\}\alpha F(s)\to\min. \qquad (8.26)$$

Перепишем функцию $\Phi(y)$ в более удобном виде:

$$\Phi(y)=c\int\limits_{-\infty}^{y}(y-s)\,\alpha F(s)+k\int\limits_{y}^{+\infty}(s-y)\,\alpha F(s) \qquad (8.26')$$

и вычислим ее первую производную:

$$\Phi'(y)=cF(y)-k(1-F(y))=(c+k)F(y)-k.$$

Заметим, что вторая производная этой функции неотрицательна (т.е. функция выпукла вниз):

$$\Phi''(y)=(c+k)F'(y)=(c+k)f(y)\geq 0,$$

поэтому, приравняв первую производную нулю, получим уравнение для минимизирующего запаса y^*:

$$F(y^*)=\frac{k}{c+k}. \qquad (8.27)$$

Решение (8.27) задачи (8.26) определяет стратегию оптимального пополнения запасов. Величина пополнения запасов h_t^*, минимизирующая средние полные издержки, задается следующим правилом:

$$h_t^*=\begin{cases}0, & \text{если } x_{t-1}\geq y^*,\\ y^*-x_{t-1}, & \text{если } x_{t-1}\leq y^*.\end{cases} \qquad (8.28)$$

Конкретные числовые характеристики системы управления запасами зависят от вида функции плотности распределения $f(x)$ случайной величины спроса. В качестве примера рассмотрим случай симметричного «треугольного» распределения спроса, при котором функция плотности распределения имеет график, представленный на рис. 8.10, *а*. Очевидно, этот график получается параллельным переносом вправо (заменой x на $x - a$) графика, изображенного на рис. 8.10, *б*, функции

$$f(x)=\begin{cases} 0 & \text{при } x \le -b \\ \dfrac{x+b}{b^2} & \text{при } -b \le x \le 0 \\ -\dfrac{x+b}{b^2} & 0 \le x \le b \\ 0 & x \ge b. \end{cases} \tag{8.29}$$

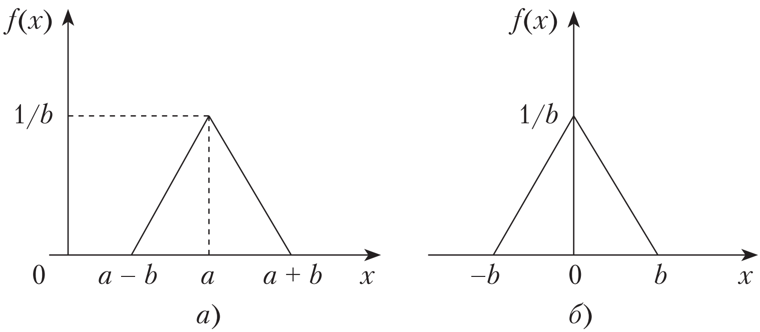


Рис. 8.10

Вычислим числовые характеристики для функции плотности распределения, заданной в (8.29) (рис 8.10, *б*). В этом случае функция распределения $F(x)$ задается формулой

$$F(x)=\begin{cases} 0 & \text{при } x \le -b \\ \dfrac{(x+b)^2}{2b^2} & \text{при } -b \le x \le 0 \\ 1-\dfrac{(x+b)^2}{2b^2} & 0 \le x \le b \\ 1 & x \ge b. \end{cases} \tag{8.30}$$

Случайная величина спроса s имеет следующие числовые характеристики: математическое ожидание $Ms = 0$; математическое ожидание квадрата $Ms^2 = b^2/6$; дисперсия $Ds = b^2/6$, среднее квадратическое отклонение $\sigma(s) = b\big/\sqrt{6}$.

Непосредственные вычисления с использованием функции средних полных издержек в виде (8.26′) показывают, что при $x \le -b$

$$\Phi(x) = -k\int_{-\infty}^{+\infty}(x-s)f(s)ds = -kx\int_{-\infty}^{+\infty}f(s)ds + k\int_{-\infty}^{+\infty}sf(s)ds = \\ = -kx + kMs = -kx. \tag{8.31}$$

Используя выражение для первой производной $\Phi'(x)$ и выражение (8.30) для функции распределения $F(x)$, можно определить, что при $-b \le x \le 0$

$$\Phi(x) = (c+k)\frac{(x+b)^3}{6b^2} - kx + C_1. \tag{8.32}$$

Здесь константа интегрирования C_1 определяется путем приравнивания выражения (8.31) и (8.32) при $x = -b$: $kb = kb + C_1$, т.е. $C_1 = 0$.

Аналогично при $0 \le x \le b$ можно получить:

$$\Phi(x) = -(c+k)\frac{(x-b)^3}{6b^2} + cx, \tag{8.33}$$

а при $x \ge b$:

$$\Phi(x) = cx. \tag{8.34}$$

Выражения (8.31)—(8.34) задают в разных интервалах искомую функцию средних полных издержек. Заменяя в ней x на $x - a$, получим функцию средних полных издержек, если функция плотности распределения спроса имеет вид, изображенный на рис. 8.10, *а*. Для иллюстрации график функции средних полных издержек для такой функции спроса в случае $k > c$ представлен на рис. 8.11, где оптимальный уровень запаса $y^* = a + b - \sqrt{\frac{2c}{c+k}} \cdot b.$

В общем виде для данной функции плотности распределения спроса оптимальный уровень запаса задается формулами

$$y^* = \begin{cases} a - b + \sqrt{\dfrac{2k}{c+k}} \cdot b & \text{при } c > k \\ a & \text{при } c = k \\ a + b - \sqrt{\dfrac{2c}{c+k}} \cdot b & \text{при } c < k, \end{cases} \tag{8.35}$$

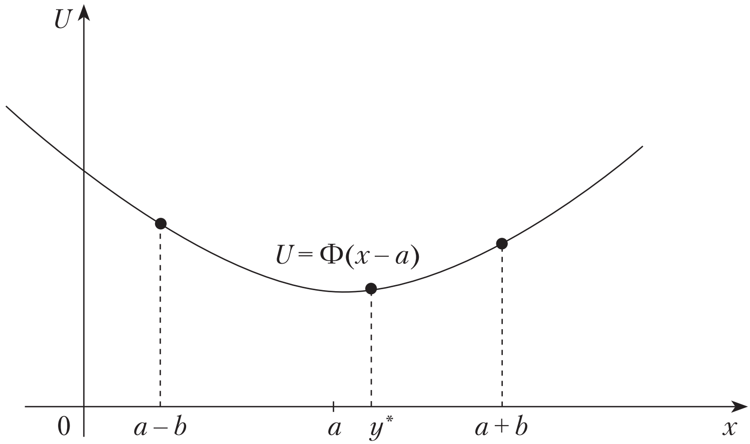


Рис. 8.11

а значение $\Phi^* = \Phi(y^*)$ минимума средних полных издержек имеет вид

$$\Phi^* = \begin{cases} bk\left(1-\dfrac{2}{3}\sqrt{\dfrac{2k}{c+k}}\right) & \text{при } c > k \\ b\cdot k/3 = b\cdot c/3 & \text{при } c = k \\ bc\left(1-\dfrac{2}{3}\sqrt{\dfrac{2c}{c+k}}\right) & \text{при } c < k. \end{cases} \tag{8.36}$$

Из формул (8.35) и (8.36) для данной модели следует, что оптимальный уровень запаса при $c \neq k$ и минимум средних полных издержек при всех c, k линейно зависят от величины b, т.е. от длины интервала разброса значений величины спроса на товар.

Напомним, что стратегия оптимального пополнения запасов задается формулами (8.28).

Пример 8.3. Пусть некоторая фирма в соответствии с договором реализует со склада по заявкам холодильники, причем ежедневный спрос является случайной величиной, функция плотности распределения которой представлена графически на рис. 8.10, *а* и колеблется от 20 до 80 холодильников в день. Средние издержки хранения одного холодильника в день составляют 8 руб., а штраф за дефицит (недопоставку) одного холодильника в день равен 17 руб. Требуется определить стратегию оптимального пополнения запаса холодильников и минимальные средние полные издержки.

В условиях рассматриваемой задачи $b = (80 - 20)/2 = 30$ (хол.); $a = (20 + 80/2 = 50$ (хол.); $c = 8$ руб.; $k = 17$ руб.

В соответствии с формулой (8.35) оптимальный уровень запаса ($c < k$) составляет

$$y^* = 50 + 30 - \sqrt{2 \cdot 8 / (8 + 17)} \cdot 30 = 80 - 0{,}8 \cdot 30 = 56 \text{ (хол.)}.$$

Тогда величина h_t^* пополнения запаса холодильников фирмой, при которой средние полные издержки будут минимальны, задается в соответствии с формулой (8.28) правилом:

$$h_t^* = \begin{cases} 0, & \text{если } x_{t-1} \geq 56 \\ 56 - x_{t-1}, & \text{если } x_{t-1} \leq 56, \end{cases}$$

где x_{t-1} — запас холодильников на складе фирмы на конец предыдущего дня. Так, если на конец предыдущего дня на складе фирмы было 60 холодильников, то пополнять запас не следует, а если на конец предыдущего дня на складе фирмы оставалось 25 холодильников, то следует реализовать заказ на пополнение запаса в количестве $56 - 25 = 31$ холодильника.

Если придерживаться этой стратегии пополнения запаса холодильников, то минимальный уровень средних полных издержек в расчете на один день в соответствии с формулой (8.36) составит

$$\Phi^* = 30 \cdot 8\left(1 - 2 / 3\sqrt{2 \cdot 8 / (8 + 17)}\right) = 240 \cdot 7 / 15 = 112 \text{ руб.}$$

Принципиальные системы регулирования товарных запасов

Рассмотренная в предыдущем параграфе классическая задача управления запасами иллюстрирует общий теоретический подход к задаче регулирования запасов. В практической деятельности организаций и служб маркетинга используются более простые принципиальные системы регулирования товарных запасов, основанные на различных стратегиях пополнения запасов, т.е. на определенных правилах этого пополнения, выраженных в достаточно общей форме. В качестве параметров в этих системах принимаются величина имеющихся на складе запасов, допустимые колебания уровня запасов, размеры заказа на пополнение запасов, его периодичность и др. Системы различаются между собой в зависимости от того, какие из параметров выбраны в качестве регулирующих. Принципиальные системы регулирования запасов, используемые в практике маркетинга, подробно описаны во многих учебниках и пособиях[1]. Поэтому дадим здесь лишь краткий обзор этих систем.

Система с фиксированным размером заказа. Это наиболее распространенная система, в которой размер заказа на пополнение запасов — постоянная величина, а поставка очередной партии товара осуществляется при уменьшении наличных запасов до определенного критического уровня, называемого *точкой заказа*. Поэтому регулирующими параметрами системы с фиксированным размером заказа являются: 1) точка заказа, т.е. фиксированный уровень запаса, при снижении до которого организуется заготовка очередной партии товара, и 2) размер заказа, т.е. величина партии поставки. Данную систему часто называют «двухбункерной», так как запас хранится как бы в двух бункерах: в первом бункере для удовлетворения спроса в течение периода между фактическим пополнением запаса и датой следующего ближайшего заказа, а во втором — для удовлетворения спроса в течение периода от момента подачи заказа до поступления очередной партии товара, т.е. во втором бункере хранится запас на уровне точки заказа.

Система с фиксированной периодичностью заказа. При этой системе заказы на очередную поставку товарного запаса повторяются через равные промежутки времени. В конце каждого периода проверяется уровень запасов и исходя из этого определяется размер заказываемой партии; при этом запас пополняется каждый раз до определенного уровня, не превышающего максимальный запас. Таким образом, регулирующие параметры этой системы: 1) максимальный уровень запасов, до которого осуществляется их пополнение, и 2) продолжительность периода повторения заказов. Система с фиксированной периодичностью заказа эффективна, когда имеется возможность пополнять запас в различных размерах, причем затраты на оформление заказа любого размера невелики. Одним из достоинств этой системы можно считать возможность периодической проверки остатков на складе и отсутствие необходимости вести систематический учет движения остатков. К недостаткам системы относится то, что она не исключает возможность нехватки товарных запасов.

---

[1] См., например, *Мельник М. М.* Экономико-математические методы и модели в планировании и управлении материально-техническим снабжением : учеб. М. : Высшая школа, 1990.

Система с двумя фиксированными уровнями запасов и с фиксированной периодичностью заказа. В этой системе допустимый уровень запасов регламентируется как сверху, так и снизу. Кроме максимального верхнего уровня запаса устанавливается нижний уровень (точка заказа). Если размер запаса снижается до нижнего уровня еще до наступления фиксированного времени пополнения запаса, то делается внеочередной заказ. В остальных случаях система функционирует как система с фиксированной периодичностью заказа. В данной системе имеется три регулирующих параметра: 1) максимальный уровень запаса; 2) нижний уровень запаса (точка заказа) и 3) длительность периода между заказами. Первые два параметра постоянны, третий — частично переменный. Рассматриваемая система сложнее предыдущей, однако она позволяет исключить возможность нехватки товарного запаса. Недостатком системы является то, что пополнение запасов до максимального уровня не может производиться независимо от фактического расходования запасов.

Система с двумя фиксированными уровнями запасов без постоянной периодичности заказа, или **(s, S)-стратегия управления запасами.** Эту систему называют также (S—s)-системой, или системой «максимум-минимум». Рассмотрим (s, S)-стратегию управления запасами более подробно. Она устраняет недостаток предыдущей системы и является ее модификацией. В этой системе два регулирующих параметра: 1) нижний (критический) уровень запаса s и 2) верхний уровень запаса S.

Если через x обозначить величину запасов до принятия решения о их пополнении, через p — величину пополнения, а через $y = x + p$ – величину запасов после пополнения, то (s, S)-стратегия управления запасами задается функцией

$$y(x) = \begin{cases} x & \text{при } x > s \\ S & \text{при } x \le s, \end{cases} \qquad (8.37)$$

т.е. пополнения не происходит, если имеющийся уровень запасов больше критического уровня s; если имеющийся уровень меньше или равен s, то принимается решение о пополнении запаса обязательно до верхнего уровня S, так что величина пополнения равна $p = S - x$.

Пример 8.4. Рассмотрим типовую задачу. Пусть при пополнении запасов автомобилей на складе служба маркетинга магазина

«Автомобили» придерживается (s, S)-стратегии при $s = 50$, $S = 300$. Требуется определить, на какое количество автомобилей надо оформить заказ, если в момент принятия решения о заказе на складе: а) 40, б) 70, в) 150, г) 10, д) 290 автомобилей (временем доставки заказанных автомобилей пренебречь).

В соответствии с формулой (8.37) величина пополнения p в каждом из рассматриваемых случаев будет равна: а) 260, б) 0, в) 0, г) 290, д) 0 автомобилей, т.е. в случаях б), в) и д) заказ на пополнение запаса автомобилей не оформляется.

Саморегулирующиеся системы. Рассмотренные выше системы регулирования запасов предполагают относительную неизменность условий их функционирования. На практике такое постоянство условий встречается редко, что вызвано изменениями потребности в товарных запасах, условиями их поставки и т.д. В связи с этим возникает необходимость создания комбинированных систем с возможностью саморегулирования (адаптации к изменившимся условиям). Создаются системы с изменяющимися периодичностью и размером заказов, учитывающие стохастические (недетерминированные) условия. В каждой такой системе устанавливается определенная целевая функция, служащая критерием оптимальности функционирования системы в рамках соответствующей экономико-математической модели управления запасами.

В качестве целевой функции в моделях управления запасами чаще всего используется минимум затрат, связанных с заготовкой и хранением запасов, а также потери от дефицита. Пример подобной целевой функции в общем виде рассмотрен выше при изучении классической задачи управления запасами.

Одним из элементов целевой функции при построении саморегулирующихся систем управления запасами являются затраты, связанные с организацией заказа и его реализацией, начиная с поиска поставщика и кончая оплатой всех услуг по доставке товарных запасов на склад. Часть расходов, связанных с организацией заказов, не зависит от размера заказа, но зависит от количества этих организаций в год. Расходы, связанные с реализацией заказа, зависят от размера заказанной партии, причем расходы в расчете на единицу товара уменьшаются при увеличении размера партии.

Другой элемент целевой функции — затраты, связанные с хранением запаса. Часть издержек хранения носит суточный характер (плата за аренду помещений, за отопление и др.), другая часть прямо зависит от уровня запасов (расходы на складскую переработку товарных запасов, потери от порчи, издержки учета и др.). При расчетах на основе экономико-математических моделей управления запасами обычно пользуются удельной величиной издержек хранения, равной издержкам на единицу хранимого товара в единицу времени. При этом предполагают, что издержки хранения за календарный период пропорциональны размеру запасов и длительности периода между заказами и обратно пропорциональны количеству заказов за этот период.

Наконец, третьим элементом рассматриваемой целевой функции являются потери из-за дефицита, когда снабженческо-сбытовая организация несет материальную ответственность за неудовлетворение потребности потребителей из-за отсутствия запасов.

Например, при неудовлетворенном спросе снабженческо-сбытовая организация может нести убытки в виде штрафа за срыв поставки. Вероятность дефицита — это ожидаемая относительная частота появления случаев нехватки товарной продукции в течение более или менее продолжительного интервала времени. Часто вероятность дефицита определяется как частное от деления числа дней, когда товар на складе отсутствует, на общее число рабочих дней, например, в году.

Модель экономически выгодных размеров заказываемых партий

Рассмотрим работу склада, на котором хранятся товарные запасы, расходуемые на снабжение потребителей. Работа реального склада сопровождается множеством отклонений от идеального режима: заказана партия одного объема, а прибыла партия другого объема; по плану партия должна прибыть через две недели, а она пришла через 10 дней; при норме разгрузки одни сутки разгрузка партии длилась трое суток и т.д. Учесть все эти отклонения практически невозможно, поэтому при моделировании работы склада обычно делаются следующие предположения:

— скорость расходования запасов со склада — постоянная величина, которую обозначим M (единиц товарных запасов в единицу времени); в соответствии с этим график изменения величины запасов в части расходования является отрезком прямой;

— объем партии пополнения Q есть постоянная величина, так что система управления запасами — это система с фиксированным размером заказа;

— время разгрузки прибывшей партии пополнения запасов мало, будем считать его равным нулю;

— время от принятия решения о пополнении до прихода заказанной партии есть постоянная величина Δt, так что можно считать, что заказанная партия приходит как бы мгновенно: если нужно, чтобы она пришла точно в определенный момент, то ее следует заказать в момент времени на Δt ранее;

— на складе не происходит систематического накопления или перерасхода запасов. Если через T обозначить время между двумя последовательными поставками, то обязательно выполнение равенства: $Q = MT$. Из сказанного выше следует, что работа склада происходит одинаковыми циклами длительностью T и за время цикла величина запаса изменяется от максимального уровня S до минимального уровня s;

— наконец, будем считать обязательным выполнение требования, чтобы отсутствие запасов на складе было недопустимым, т.е. выполняется неравенство $s \geq 0$. С точки зрения уменьшения издержек склада на хранение отсюда вытекает, что $s = 0$ и, следовательно, $S = Q$.

Окончательно график «идеальной» работы склада в форме зависимости величины запасов y от времени t будет иметь вид, представленный на рис. 8.12.

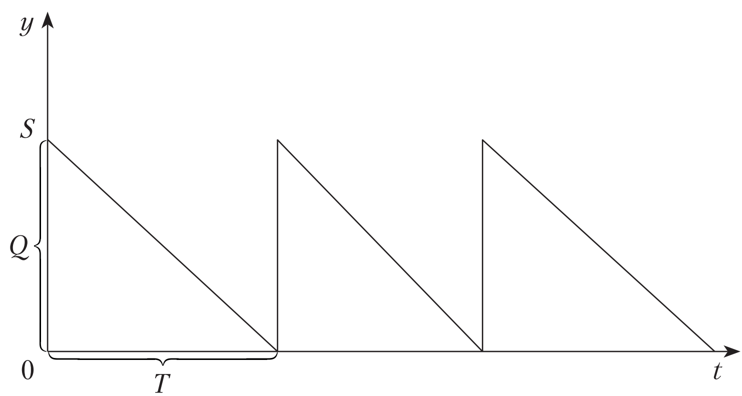


Рис. 8.12

В предыдущем параграфе отмечалось, что эффективность работы склада оценивается по его затратам на пополнение запасов и их хранение. Расходы, не зависящие от объема партии, называют накладными. Сюда входят почтово-телеграфные расходы, командировочные, некоторая часть транспортных расходов и др. Накладные расходы будем обозначать через K. Издержки хранения запасов будем считать пропорциональными величине хранящихся запасов и времени их хранения. Издержки на хранение одной единицы запасов в течение одной единицы времени называются величиной удельных издержек хранения; мы их будем обозначать через h.

При изменяющейся величине хранящихся запасов издержки хранения за некоторое время T получают путем умножения величины $h \cdot T$ на среднее значение величины запасов в течение этого времени T. Таким образом, затраты склада за время T при размере партии пополнения Q в случае идеального режима работы склада, представленного на рис. 8.12, равны

$$Z_T(Q) = K + h \cdot T \cdot Q / 2.$$

После деления этой функции на постоянную величину T с учетом равенства $Q = MT$ получим выражение для величины затрат на пополнение и хранение запасов, приходящихся на единицу времени:

$$Z_1(Q) = Z_T(Q) / T = K / T + h \cdot Q / 2 = K \cdot M / Q + h \cdot Q / 2. \quad (8.38)$$

Это и будет целевой функцией, минимизация которой позволит указать оптимальный режим работы склада.

Найдем объем заказываемой партии Q, при котором минимизируется функция средних затрат склада за единицу времени, т.е. функция $Z_1(Q)$. На практике величина Q часто принимает дискретные значения, в частности, из-за использования транспортных средств определенной грузоподъемности; в этом случае оптимальное значение Q находят перебором допустимых значений Q. Мы будем считать, что ограничений на принимаемые значения Q нет, тогда задачу на минимум функции $Z_1(Q)$ (легко показать, что она является выпуклой (рис. 8.13)) можно решить методами дифференциального исчисления:

$$\frac{dZ_1}{dQ} = -\frac{KM}{Q^2} + \frac{h}{2} = 0,$$

откуда находим точку минимума $Q_{\text{опт}}$:

$$Q_{\text{опт}} = \sqrt{\frac{2KM}{h}}. \tag{8.39}$$

Эта формула называется *формулой Уилсона* (по имени английского ученого-экономиста, получившего ее в 20-х годах прошлого столетия).

Оптимальный размер партии, рассчитываемый по формуле Уилсона, обладает *характеристическим свойством*: размер партии Q оптимален тогда и только тогда, когда издержки хранения за время цикла T равны накладным расходам K.

Действительно, если $Q = \sqrt{2KM / h}$, то издержки хранения за цикл равны

$$h\frac{Q}{2}T = h\frac{Q}{2}\frac{Q}{M} = h\frac{2KM}{2Mh} = K.$$

Если же издержки хранения за цикл равны накладным расходам, т.е.

$$h\frac{Q}{2}T = h\frac{Q}{2}\frac{Q}{M} = K,$$

то

$$Q = \sqrt{2KM / h}.$$

Проиллюстрируем характеристическое свойство оптимального размера партии графически (рис. 8.13).

На рис. 8.13 видно, что минимальное значение функции $Z_1(Q)$ достигается при том значении Q, при котором равны значения двух других функций, ее составляющих.

Используя формулу Уилсона (8.39), в сделанных ранее предположениях об идеальной работе склада можно получить ряд расчетных характеристик работы склада в оптимальном режиме: оптимальный средний уровень запаса

$$\bar{Q}_{\text{опт}} = Q_{\text{опт}} / 2 = \sqrt{KM / 2h}; \tag{8.40}$$

оптимальная периодичность пополнения запасов

$$T_{\text{опт}} = Q_{\text{опт}} / M = \sqrt{2K / Mh}; \tag{8.41}$$

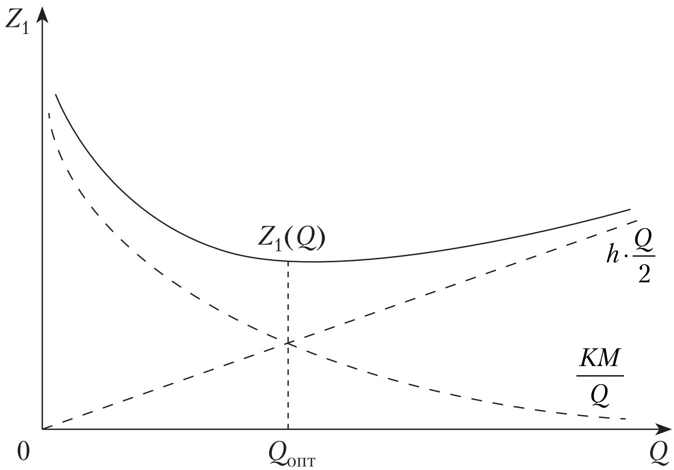


Рис. 8.13

оптимальные средние издержки хранения запасов в единицу времени

$$\bar{H}_1 = Q_{опт}h = \sqrt{KMh / 2}. \qquad (8.42)$$

Пример 8.5. Рассмотрим типовую задачу. На склад доставляют цемент на барже по 1500 т. В сутки со склада потребители забирают 50 т цемента. Накладные расходы по доставке партии цемента равны 2 тыс. руб. Издержки хранения 1 т цемента в течение суток равны 0,1 руб. Требуется определить: 1) длительность цикла, среднесуточные накладные расходы и среднесуточные издержки хранения; 2) эти же величины для размеров партии в 500 т и в 3000 т; 3) каковы оптимальный размер заказываемой партии и расчетные характеристики работы склада в оптимальном режиме.

Решение. Параметры работы склада: M = 50 т/сут.; K = 2 тыс. руб.; h = 0,1 руб./т · сут.; Q = 1500 т.

1. Длительность цикла:

$T = Q : M$ = 1500 т : 50 т/сут. = 30 сут.;

среднесуточные накладные расходы:

$K : T$= 2 тыс. руб. : 30 сут. ≈ 67 руб./сут.;

среднесуточные издержки хранения:

$h \cdot Q/2$ = 0,1 руб./т · сут. · 1500 т/2 = 75 руб./сут.

2. Аналогичные расчеты проведем для Q_1 = 500 т:

$T_1 = Q_1 : M$= 500 т : 50 т/сут. = 10 сут.;

$K : T_1$ = 2 тыс. руб. : 10 сут. = 200 руб./сут.;

$h \cdot Q_1/2$ = 0,1 руб./т · сут · 500 т/2 = 25 руб./сут.

и для $Q_2 = 3000$ т:

$T_2 = Q_2 : M = 3000$ т : 50 т/сут. = 60 сут.

$K : T_2 = 2$ тыс. руб. : 60 сут. ≈ 33 руб./сут.;

$h \cdot Q_2/2 = 0{,}1$ руб./т · сут. · 3000 т/2 = 150 руб./сут.

3. Найдем оптимальный размер заказываемой партии по формуле Уилсона (8.39):

$$Q_{\text{опт}} = \sqrt{\frac{2KM}{h}} = \sqrt{\frac{2 \cdot 2000 \cdot 50}{0{,}1}} \approx 1400 \text{ т};$$

оптимальный средний уровень запаса по формуле (8.40):

$$\bar{Q}_{\text{опт}} = Q_{\text{опт}} / 2 = 1400 \text{ т} / 2 = 700 \text{ т};$$

оптимальную периодичность пополнения запасов по формуле (8.41):

$$T_{\text{опт}} = \frac{Q_{\text{опт}}}{M} = \frac{1400}{50 \text{ т} / \text{сут.}} = 28 \text{ сут.};$$

оптимальные средние издержки хранения запасов в единицу времени по формуле (8.42):

$$\bar{H}_1 = \bar{Q}_{\text{опт}} h = 700 \text{ т} \cdot 0{,}1 \text{ руб.} / \text{т} \cdot \text{сут.} = 70 \text{ руб./сут.}$$

8.3. Моделирование систем массового обслуживания

Многие экономические задачи связаны с *системами массового обслуживания* (СМО), т.е. такими системами, в которых, с одной стороны, возникают массовые запросы (требования) на выполнение каких-либо услуг, с другой — происходит удовлетворение этих запросов. СМО включает в себя следующие элементы: источник требований, входящий поток требований, очередь, обслуживающие устройства (каналы обслуживания), выходящий поток требований. Исследованием таких систем занимается *теория массового обслуживания*.

Методами теории массового обслуживания могут быть решены многие задачи исследования процессов, происходящих в экономике. Так, в организации торговли эти методы позволяют определить оптимальное количество торговых точек данного профиля, численность продавцов, частоту завоза товаров и другие параметры. Другим характерным примером систем массового обслуживания могут служить склады или базы снабженческо-сбытовых организаций, и задача теории массового обслуживания в данном случае сводится к тому, чтобы установить оптимальное соотношение между числом поступающих на базу требований на обслуживание и числом обслуживающих устройств, при котором суммарные расходы на обслуживание и убытки от простоя транспорта были бы минимальными. Теория массового обслуживания может найти применение и при расчете площади складских помещений, при этом складская площадь рассматривается как обслуживающее устройство, а прибытие транспортных средств под выгрузку — как требование. Модели теории массового обслуживания применяются также при решении ряда задач организации и нормирования труда, других социально-экономических проблем.

Системы массового обслуживания могут быть классифицированы по ряду признаков.

1. В зависимости от *условий ожидания* начала обслуживания различают:

— СМО *с потерями* (отказами);

— СМО *с ожиданием*.

В СМО с отказами требования, поступающие в момент, когда все каналы обслуживания заняты, получают отказ и теряются. Классическим примером системы с отказами является телефонная станция. Если вызываемый абонент занят, то требование на соединение с ним получает отказ и теряется.

В СМО с ожиданием требование, застав все обслуживающие каналы занятыми, становится в очередь и ожидает, пока не освободится один из обслуживающих каналов.

СМО, допускающие очередь, но с ограниченным числом требований в ней, называются *системами с ограниченной длиной очереди*.

СМО, допускающие очередь, но с ограниченным сроком пребывания каждого требования в ней, называются *системами с ограниченным временем ожидания*.

2. По *числу каналов обслуживания* СМО делятся на:

— *одноканальные*;

— *многоканальные*.

3. По месту нахождения источника требований СМО делятся на:

— *разомкнутые*, когда источник требования находится вне системы;

— *замкнутые*, когда источник находится в самой системе.

Примером разомкнутой системы может служить ателье по ремонту телевизоров. Здесь неисправные телевизоры — это источник требований на их обслуживание, находятся вне самой системы, число требований можно считать неограниченным. К замкнутым СМО относится, например, станочный участок, в котором станки являются источником неисправностей, а следовательно, источником требований на их обслуживание, например, бригадой наладчиков.

Возможны и другие признаки классификации СМО, например *по дисциплине обслуживания, однофазные и многофазные* СМО и др.

Методы и модели, применяющиеся в теории массового обслуживания, можно условно разделить на аналитические и имитационные.

Аналитические методы теории массового обслуживания позволяют получить характеристики системы как некоторые функции параметров ее функционирования. Благодаря этому появляется возможность проводить качественный анализ влияния отдельных факторов на эффективность работы СМО. *Имитационные методы* основаны на моделировании процессов массового обслуживания на ЭВМ и применяются, если невозможно применение аналитических моделей; ряд основных понятий имитационного моделирования рассмотрен в параграфе 3.5. Далее будем рассматривать аналитические методы моделирования СМО.

В настоящее время теоретически наиболее разработаны и удобны в практических приложениях методы решения таких задач массового обслуживания, в которых входящий поток требований является *простейшим* (*пуассоновским*).

Для простейшего потока частота поступления требований в систему подчиняется закону Пуассона, т.е. вероятность поступления за время t ровно k требований задается формулой

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}. \tag{8.43}$$

Простейший поток обладает тремя основными свойствами: ординарности, стационарности и отсутствием последействия.

Ординарность потока означает практическую невозможность одновременного поступления двух и более требований. Например, достаточно малой является вероятность того, что из группы станков, обслуживаемых бригадой ремонтников, одновременно выйдут из строя сразу несколько станков.

Стационарным называется поток, для которого математическое ожидание числа требований, поступающих в систему в единицу времени (обозначим λ), не меняется во времени. Таким образом, вероятность поступления в систему определенного количества требований в течение заданного промежутка времени Δt зависит от его величины и не зависит от начала его отсчета на оси времени.

Отсутствие последействия означает, что число требований, поступивших в систему до момента t, не определяет того, сколько требований поступит в систему за промежуток времени от t до $t + \Delta t$.

Например, если на ткацком станке в данный момент произошел обрыв нити и он устранен ткачихой, то это не определяет, произойдет новый обрыв на данном станке в следующий момент или нет, тем более это не влияет на вероятность возникновения обрыва на других станках.

Важная характеристика СМО — *время обслуживания* требований в системе. Время обслуживания одного требования является, как правило, случайной величиной и, следовательно, может быть описано законом распределения. Наибольшее распространение в теории и особенно в практических приложениях получил *экспоненциальный закон распределения времени обслуживания*. Функция распределения для этого закона имеет вид

$$F(t) = 1 - e^{-\mu t}, \tag{8.44}$$

т.е. вероятность того, что время обслуживания не превосходит некоторой величины t, определяется формулой (8.44), где μ — параметр экспоненциального закона распределения времени обслуживания требований в системе, т.е. величина, обратная среднему времени обслуживания $\overline{t}_{об}$:

$$\mu = 1/\overline{t}_{об}. \tag{8.45}$$

Рассмотрим аналитические модели наиболее распространенных СМО с ожиданием, т.е. таких СМО, в которых требования, поступившие в момент, когда все обслуживающие каналы заняты, ставятся в очередь и обслуживаются по мере освобождения каналов.

Общая постановка задачи состоит в следующем. Система имеет n обслуживающих каналов, каждый из которых может одновременно обслуживать только одно требование.

В систему поступает простейший (пуассоновский) поток требований с параметром λ. Если в момент поступления очередного требования в системе на обслуживании уже находится не меньше n требований (т.е. все каналы заняты), то это требование становится в очередь и ждет начала обслуживания.

Время обслуживания каждого требования $t_{\text{об}}$ — случайная величина, которая подчиняется экспоненциальному закону распределения с параметром μ.

СМО с ожиданием можно разбить на две большие группы: замкнутые и разомкнутые. К *замкнутым* относятся системы, в которых поступающий поток требований возникает в самой системе и ограничен. Например, мастер, задачей которого является наладка станков в цехе, должен периодически их обслуживать. Каждый налаженный станок становится потенциальным источником требований на накладку. В подобных системах общее число циркулирующих требований конечно и чаще всего постоянно.

Если питающий источник обладает бесконечным числом требований, то системы называются *разомкнутыми*. Примерами подобных систем могут служить магазины, кассы вокзалов, портов и др. Для этих систем поступающий поток требований можно считать неограниченным.

Отмеченные особенности функционирования систем этих двух видов накладывают определенные условия на используемый математический аппарат. Расчет характеристик работы СМО различного вида может быть проведен на основе расчета вероятностей состояний СМО (так называемые *формулы Эрланга*).

Рассмотрим алгоритмы расчета показателей качества функционирования разомкнутой системы массового обслуживания с ожиданием.

При изучении таких систем рассчитывают различные показатели эффективности обслуживающей системы. В качестве основных показателей могут быть вероятность того, что все каналы свободны или заняты, математическое ожидание длины очереди (средняя длина очереди), коэффициенты занятости и простоя каналов обслуживания и др.

1. Введем в рассмотрение параметр $\alpha = \lambda/\mu$. Заметим, что если $\alpha/n < 1$, то очередь не может расти безгранично. Это условие имеет следующий смысл: λ — среднее число требований, поступающих за единицу времени, $1/\mu$ — среднее время обслуживания одним каналом одного требования,

тогда $\alpha = \lambda \cdot 1/\mu$ — среднее число каналов, которое необходимо иметь, чтобы обслуживать в единицу времени все поступающие требования. Поэтому условие $\alpha/n < 1$ означает, что число обслуживающих каналов должно быть больше среднего числа каналов, необходимых для того, чтобы за единицу времени обслужить все поступившие требования. Важнейшие характеристики работы СМО:

$$P_0 = \left[\sum_{k=0}^{n-1} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!(1-\alpha/n)}\right]^{-1}. \tag{8.46}$$

2. Вероятность того, что занято ровно k обслуживающих каналов при условии, что общее число требований, находящихся на обслуживании, не превосходит числа обслуживающих аппаратов:

$$P_k = \frac{\alpha^k}{k!} P_0 \quad \text{при} \quad 1 \le k \le n. \tag{8.47}$$

3. Вероятность того, что в системе находится k требований в случае, когда их число больше числа обслуживающих каналов:

$$P_k = \frac{\alpha^k}{n!n^{k-n}} P_0 \quad \text{при} \quad k \ge n. \tag{8.48}$$

4. Вероятность того, что все обслуживающие каналы заняты:

$$P_n = \frac{\alpha^n}{n!(1-\alpha/n)} P_0; \quad (\alpha/n < 1). \tag{8.49}$$

5. Среднее время ожидания требованием начала обслуживания в системе:

$$\bar{t}_{\text{ож}} = \frac{P_n}{\mu(n-\alpha)}; \quad (\alpha/n < 1). \tag{8.50}$$

6. Средняя длина очереди:

$$\bar{L}_{\text{оч}} = \frac{\alpha P_n}{n(1-\alpha/n)} = \frac{\alpha^{n+1}}{n!n(1-\alpha/n)^2} P_0; \quad (\alpha/n < 1). \tag{8.51}$$

7. Среднее число свободных от обслуживания каналов:

$$\bar{N}_0 = \sum_{k=0}^{n-1} \frac{n-k}{k!} \alpha^k P_0. \tag{8.52}$$

8. Коэффициент простоя каналов:

$$K_{\text{пр}} = \frac{\bar{N}_0}{n}. \tag{8.53}$$

9. Среднее число занятых обслуживанием каналов:

$$\bar{N}_{\text{з}} = n - \bar{N}_0. \tag{8.54}$$

10. Коэффициент загрузки каналов:

$$K_{\text{з}} = \frac{\bar{N}_{\text{з}}}{n}. \tag{8.55}$$

Пример 8.6. Пусть филиал фирмы по ремонту радиоаппаратуры имеет $n = 5$ опытных мастеров. В среднем в течение рабочего дня от населения поступает в ремонт $\lambda = 10$ радиоаппаратов. Общее число радиоаппаратов, находящихся в эксплуатации у населения, очень велико, и они независимо друг от друга в различное время выходят из строя. Поэтому есть все основания полагать, что поток заявок на ремонт аппаратуры является случайным, пуассоновским. В свою очередь каждый аппарат в зависимости от характера неисправности также требует различного случайного времени на ремонт. Время на проведение ремонта зависит во многом от серьезности полученного повреждения, квалификации мастера и множества других причин. Пусть статистика показала, что время ремонта подчиняется экспоненциальному закону; при этом в среднем в течение рабочего дня каждый из мастеров успевает отремонтировать $\mu = 2{,}5$ радиоаппарата. Требуется оценить работу филиала фирмы по ремонту радиоаппаратуры, рассчитав ряд основных характеристик данной СМО.

За единицу времени принимаем 1 рабочий день (7 часов).

1. Определим параметр:

$$\alpha = \lambda \frac{1}{\mu} = 10 \cdot 1 / 2{,}5 = 4,$$

так как $\alpha < n$, то очередь не может расти безгранично.

2. Вероятность того, что все мастера свободны от ремонта аппаратуры, равна согласно (8.46):

$$P_0 = \frac{1}{1 + 4 + 4^2/2! + 4^3/3! + 4^4/4! + 4^5/5!(1 - 4/5)} = 0{,}013.$$

3. Вероятность того, что все мастера заняты ремонтом, находим по (8.49):

$$P_n = \frac{4^5 \cdot 0{,}013}{5!(1 - 4/5)} = 0{,}554.$$

Это означает, что 55,4% времени мастера полностью загружены работой.

4. Среднее время обслуживания (ремонта) одного аппарата согласно (8.45):

$$\bar{t}_{\text{об}} = 1 / \mu = 7 / 2{,}5 = 2{,}8 \text{ ч/аппарат.}$$

(при условии семичасового рабочего дня).

5. В среднем время ожидания каждого неисправного аппарата начала ремонта равно по (8.50):

$$\bar{t}_{\text{ож}} = \frac{0{,}554 \cdot 2{,}8}{5-4} = 1{,}55 \text{ ч.}$$

6. Очень важной характеристикой является средняя длина очереди, которая определяет необходимое место для хранения аппаратуры, требующей ремонта; находим ее по (8.51):

$$\bar{L}_{\text{оч}} = \frac{0{,}554 \cdot 4}{5(1-4/5)} = 2{,}2 \text{ аппарата.}$$

7. Определим среднее число мастеров, свободных от работы, по (8.52):

$$N_0 = 0{,}013\left[\frac{5-0}{1}\cdot 1 + \frac{5-1}{1!}\cdot 4 + \frac{5-2}{2!}\cdot 4^2 + \frac{5-3}{3!}\cdot 4^3 + \frac{5-4}{4!}\cdot 4^4\right] = 0{,}95.$$

Таким образом, в среднем в течение рабочего дня ремонтом заняты четыре мастера из пяти.

Перейдем к рассмотрению алгоритмов расчета характеристик функционирования замкнутых СМО. Поскольку система замкнутая, то к постановке задачи следует добавить условие: поток поступающих требований ограничен, т.е. в системе обслуживания одновременно не может находиться больше m требований (m — число обслуживаемых объектов).

За критерий, характеризующий качество функционирования рассматриваемой системы, выберем отношение средней длины очереди к наибольшему числу требований, находящихся одновременно в обслуживающей системе, — коэффициент простоя обслуживаемого объекта. В качестве другого критерия возьмем отношение среднего числа незанятых обслуживающих каналов к их общему числу — коэффициент простоя обслуживающего канала.

Первый из названных критериев характеризует потери времени из-за ожидания начала обслуживания; второй показывает полноту загрузки обслуживающей системы.

Очевидно, что очередь может возникнуть, лишь когда число каналов меньше наибольшего числа требований, находящихся одновременно в обслуживающей системе ($n < m$).

Приведем последовательность расчетов характеристик замкнутых СМО и необходимые формулы.

1. Определим параметр $\alpha = \lambda/\mu$ — показатель загрузки системы, т.е. математическое ожидание числа требований, поступающих в систему за время, равное средней длительности обслуживания ($1/\mu = \bar{t}_{об}$).

2. Вероятность того, что занято k обслуживающих каналов при условии, что число требований, находящихся в системе, не превосходит числа обслуживающих каналов системы:

$$P_k = \frac{m!}{k!(m-k)!}\alpha^k P_0 \quad (1 \le k \le m). \tag{8.56}$$

3. Вероятность того, что в системе находится k требований для случая, когда их число больше числа обслуживающих каналов:

$$P_k = \frac{m!}{n^{k-n} \cdot n!(m-k)!}\alpha^k P_0 \quad (n \le k \le m). \tag{8.57}$$

4. Вероятность того, что все обслуживающие каналы свободны, определим, используя очевидное условие:

$$\sum_{k=0}^{m} P_k = 1, \quad \text{откуда} \quad P_0 = 1 - \sum_{k=1}^{m} P_k.$$

Величину P_0 можно получить также путем подстановки в равенство $\sum\limits_{k=0}^{m} P_k = 1$ значений $P_1, P_2, ..., P_m$, в которые P_0 входит сомножителем. Подставляя их, получаем следующее уравнение для определения P_0:

$$\begin{gathered} P_0\left[\sum_{k=0}^{n}\frac{m!}{k!(m-k)!}\alpha^k + \sum_{k=n+1}^{m}\frac{m!}{n^{k-n}\cdot n!(m-k)!}\alpha^k\right] = 1, \\ P_0\left[\sum_{k=0}^{n}\frac{m!}{k!(m-k)!}\alpha^k + \sum_{k=n+1}^{m}\frac{m!}{n^{k-n}\,!(m-k)!}\alpha^k\right]^{-1}. \end{gathered} \tag{8.58}$$

5. Среднее число требований, ожидающих начала обслуживания (средняя длина очереди):

$$M_{\text{оч}} = \sum_{k=n+1}^{m} (k-n)P_k,$$

или

$$M_{\text{оч}} = \sum_{k=n+1}^{m} \frac{(k-n)m!}{n^{k-n}n!(m-k)!}\alpha^k P_0. \quad (8.59)$$

6. Коэффициент простоя обслуживаемого требования (объекта):

$$K_{\text{пр.об}} = \frac{M_{\text{оч}}}{m} = \frac{1}{m}\sum_{k=n+1}^{m} (k-n)P_k. \quad (8.60)$$

7. Среднее число требований, находящихся в обслуживающей системе, обслуживаемых и ожидающих обслуживания:

$$M = \sum_{k=1}^{n} kP_k + \sum_{k=n+1}^{n} kP_k, \quad (8.61)$$

или

$$M = \left[\sum_{k=1}^{n} \frac{m!}{(k-1)!(m-k)!}\alpha^k + \sum_{k=n+1}^{m} \frac{km!}{n^{k-n}n!(m-k)!}\alpha^k\right] P_0.$$

8. Среднее число свободных обслуживающих каналов:

$$N_0 = \sum_{k=1}^{n-1} (n-k)P_k = \sum_{k=0}^{n-1} \frac{(n-k)m!}{k!(m-k)!}\alpha^k P_0. \quad (8.62)$$

9. Коэффициент простоя обслуживающего канала:

$$N_{\text{пр.кан}} = \frac{N_0}{n} = \sum_{k=0}^{n-1} P_k - \frac{1}{n}\sum_{k=0}^{n-1} kP_k. \quad (8.63)$$

Рассмотрим пример расчета характеристик замкнутой СМО.

Пример 8.7. Рабочий обслуживает группу автоматов, состоящую из 3 станков. Поток поступающих требований на обслуживание станков пуассоновский с параметром $\lambda = 2$ ст./ч. Обслуживание одного станка занимает у рабочего в среднем 12 мин, а время обслуживания подчинено экспоненциальному закону. Тогда $1/\mu =$ $= 0{,}2$ ч/ст., т.е. $\mu = 5$ ст./ч, а $\alpha = \lambda/\mu = 0{,}4$.

Необходимо определить среднее число автоматов, ожидающих обслуживания, коэффициент простоя автомата, коэффициент простоя рабочего. Обслуживающим каналом здесь является рабочий; так как станки обслуживает один рабочий, то $n = 1$. Общее число требований не может превзойти числа станков, т.е. $m = 3$.

Система может находиться в четырех различных состояниях: 1) все станки работают; 2) один стоит и обслуживается рабочим, а два работают; 3) два стоят, один обслуживается, один ждет обслуживания; 4) три стоят, из них один обслуживается, а два ждут очереди.

Для ответа на поставленные вопросы можно воспользоваться формулами (8.56) и (8.57):

$$P_1 = \frac{3!}{1!(3-1)!} \cdot 0,4 P_0 = 1,2 P_0;$$

$$P_2 = \frac{3!}{1^{2-1} 1!(3-2)!} \cdot 0,4^2 P_0 = 0,96 P_0;$$

$$P_3 = \frac{3!}{1^{3-1} 1!(3-3)!} \cdot 0,4^3 P_0 = 0,384 P_0.$$

Сведем вычисления в табл. 8.2.

Таблица 8.2

k	$k - n$	P_k / P_0	P_k	$(k - n) P_k$	kP_k
1	2	3	4	5	6
0	0	1,0000	0,2822	0	0
1	0	1,2000	0,3386	0	0,3386
2	1	0,9600	0,2709	0,2709	0,5418
3	2	0,3840	0,1083	0,2166	0,3249
	Σ	3,5440	1,0000	0,4875	1,2053

В этой таблице первой вычисляется третья графа, т.е. отношения P_k/P_0 при $k = 0, 1, 2, 3$. Затем, суммируя величины по графе и учитывая, что $\sum_{k=0}^{3} P_k = 1,$, получаем

$$\sum_{k=0}^{3} \frac{P_k}{P_0} = \frac{1}{P_0} \sum_{k=0}^{3} P_k = \frac{1}{P_0} = 3,544,$$

откуда $P_0 = 0,2822$. Умножая величины третьей графы на P_0, получаем четвертую графу. Величина $P_0 = 0,2822$, равная вероятности того, что все автоматы работают, может быть истолкована как вероятность того, что рабочий свободен. Получается, что в рассматриваемом случае рабочий будет свободен более 1/4 всего рабочего времени. Однако это не означает, что «очередь» станков, ожидающих обслуживания, всегда будет отсутствовать. Математическое ожидание числа автоматов, стоящих в очереди, равно

$$M_{\text{оч}} = \sum_{k=2}^{3} (k-1)P_k \quad (\text{так как } n = 1).$$

Суммируя пятую графу, получим $M_{\text{оч}} = 0{,}4875$, следовательно, в среднем из трех станков 0,49 станка будет простаивать в ожидании, пока освободится рабочий.

Суммируя шестую графу, получим математическое ожидание числа простаивающих станков (ремонтируемых и ожидающих ремонта):

$$M = \sum_{k=1}^{3} kP_k = 1{,}2053,$$

т.е. в среднем 1,2 станка не будет выдавать продукцию. Коэффициент простоя станка равен $K_{\text{пр.об}} = M_{\text{оч}}/3 = 0{,}1625$, т.е. каждый станок простаивает примерно 0,16 своего рабочего времени в ожидании, пока рабочий освободится.

Коэффициент простоя рабочего в данном случае совпадает с P_0, так как $n = 1$, поэтому $K_{\text{пр.кан}} = \dfrac{N_0}{n} = 0{,}2822.$

8.4. Элементы теории игр в задачах моделирования экономических процессов

При решении экономических задач часто приходится анализировать ситуации, в которых сталкиваются интересы двух или более конкурирующих сторон, преследующих различные цели; это особенно характерно в условиях рыночной экономики. Такого рода ситуации называются *конфликтными*. Математической теорией конфликтных ситуаций является *теория игр*. В игре могут сталкиваться интересы двух (*игра парная*) или нескольких (*игра множественная*) противников; существуют игры с бесконечным множеством игроков. Если во множественной игре игроки образуют коалиции, то игра называется *коалиционной*; если таких коалиций две, то игра сводится к парной.

На промышленных предприятиях теория игр может применяться для выбора оптимальных решений, например при создании рациональных запасов сырья, материалов, полуфабрикатов, когда противоборствуют две тенденции:

увеличения запасов, гарантирующих бесперебойную работу производства, и сокращения запасов в целях минимизации затрат на их хранение. В сельском хозяйстве теория игр может применяться при решении таких экономических задач, как выбор для посева одной из возможных культур, урожай которых зависит от погоды, если известны цена единицы той или иной культуры и средняя урожайность каждой культуры в зависимости от погоды (например, будет ли лето засушливым, нормальным или дождливым); в этом случае одним из игроков выступает сельскохозяйственное предприятие, стремящееся обеспечить наибольший доход, а другим — природа.

Решение подобных задач требует полной определенности в формулировании их условий (*правил игры*) установления количества игроков, выявления возможных стратегий игроков, возможных выигрышей (проигрыш понимается как отрицательный выигрыш). Важным элементом в условии игровых задач является *стратегия*, т.е. совокупность правил, которые в зависимости от ситуации в игре определяют однозначный выбор действий данного игрока. Если в процессе игры игрок применяет попеременно несколько стратегий, то такая стратегия называется *смешанной*, а ее элементы — *чистыми* стратегиями. Количество стратегий у каждого игрока может быть конечным и бесконечным, в зависимости от этого игры подразделяются на конечные и бесконечные.

Важными являются понятия *оптимальной стратегии, цены игры, среднего выигрыша*. Эти понятия находят отражение в определении решения игры, стратегии P^* и Q^* первого и второго игроков соответственно называются их *оптимальными стратегиями*, а число V — *ценой игры*, если для любых стратегий P первого игрока и любых стратегий Q второго игрока выполняются неравенства

$$M(P,Q^*) \le V \le M(P^*,Q), \tag{8.64}$$

где $M(P, Q)$ означает математическое ожидание выигрыша (средний выигрыш) первого игрока, если первым и вторым игроками избраны соответственно стратегии P и Q.

Из неравенств (8.64) следует, в частности, что $V =$ $= M(P^*, Q^*)$, т.е. цена игры равна математическому ожиданию выигрыша первого игрока, если оба игрока изберут оптимальные для себя стратегии.

Одним из основных видов игр являются *матричные игры*, которыми называются парные игры с *нулевой суммой* (один игрок выигрывает столько, сколько проигрывает другой) при условии, что каждый игрок имеет конечное число стратегий. В этом случае парная игра формально задается матрицей $A = (a_{ij})$, элементы которой a_{ij} определяют выигрыш первого игрока (и соответственно проигрыш второго), если первый игрок выберет i-ю стратегию ($i=\overline{1,m}$), а второй — j-ю стратегию ($j=\overline{1,n}$). Матрица A называется *матрицей игры*, или *платежной матрицей*.

Рассмотрим построение платежной матрицы на примере.

Пример 8.8. На базе торговой фирмы имеется n типов товара ассортиментного минимума. В магазин фирмы должен быть завезен только один из этих типов товара. Если товар типа j ($j=\overline{1,n}$) будет пользоваться спросом, то магазин от его реализации получит прибыль P_j. Если же этот товар не будет пользоваться спросом, то издержки на его хранение принесут магазину убыток q_j. Требуется выбрать тип товара, который целесообразно завезти в магазин.

В условиях неопределенного покупательского спроса конфликтная ситуация товароснабжения формализуется матричной игрой. Пусть первый игрок — магазин, второй игрок — покупательский спрос. Каждый из игроков имеет по n стратегий. Завоз i-го товара — i-я стратегия первого игрока, спрос на j-й товар – j-я стратегия второго игрока. Тогда матрица выигрышей первого игрока имеет вид квадратной матрицы n-го порядка:

$$A=\begin{pmatrix} p_1 & -q_1 & \dots & -q_1 \\ -q_2 & p_2 & \dots & -q_2 \\ \dots & \dots & \dots & \dots \\ -q_n & -q_n & \dots & p_n \end{pmatrix}.$$

Существует ряд методов решения матричных игр. Если матрица игры имеет одну из размерностей, равную двум (у одного из игроков имеется только две стратегии), то решение игры может быть получено графически. Известно несколько методов приближенного решения матричной игры, например метод Брауна. Применяются также методы линейного программирования.

В качестве примера рассмотрим решение игры, когда матрица игры имеет так называемую *седловую точку*.

Пример 8.9. Матрица игры имеет вид

$$A=\begin{pmatrix} 2 & 10 & 3 & 14 & 5 \\ 8 & 9 & 5 & 6 & 7 \\ 10 & 8 & 4 & 8 & 12 \end{pmatrix}.$$

Минимальный элемент первой строки (первой стратегии первого игрока) равен 2, второй — 5, третьей — 4; максимальное значение из этих величин равно 5. Максимальный элемент первого столбца (первой стратегии второго игрока) равен 10, второго — 10; третьего — 5, четвертого — 14, пятого — 12; минимальное значение из них равно 5. Следовательно, данная игра имеет седловую точку (2, 3) и задача разрешима в чистых стратегиях. Придерживаясь чисто второй стратегии, первый игрок обеспечивает себе выигрыш не меньший 5; второй игрок, применяя чистую третью стратегию, проигрывает не более 5. Обе стратегии $i = 2$ и $j = 3$ являются оптимальными для первого и второго игроков, при этом цена игры $V = 5$.

Во многих игровых задачах в сфере экономики неопределенность вызвана не сознательным противодействием противника, а недостаточной осведомленностью об условиях, в которых действуют стороны. Так, в рассматриваемых выше примерах были неизвестны заранее погода в некотором регионе, покупательский спрос на некоторую продукцию.

Подобного рода игры называются *играми с природой*. В этих случаях строки матрицы игры соответствуют стратегии игрока, а столбцы — состояниям «природы». В ряде случаев при решении такой игры рассматривают *матрицу рисков*.

При решении игр с природой используется также ряд критериев: критерий Сэвиджа, критерий Вальда, критерий Гурвица и др.

При *максиминном критерии Вальда* оптимальной считается та стратегия лица, принимающего решение (ЛПР), которая обеспечивает максимум минимального выигрыша; применяя этот критерий, ЛПР в большей степени ориентируется на наихудшие условия (этот критерий иногда называют критерием «крайнего пессимизма»).

Критерий *минимаксного риска Сэвиджа* предполагает, что оптимальной является та стратегия, при которой величина риска в наихудшем случае минимальна.

При использовании *критерия «пессимизм–оптимизм» Гурвица* ЛПР выбирает некоторый так называемый коэффициент пессимизма q; при $q = 1$ критерий Гурвица приводит к критерию Вальда («крайнего пессимизма»), а при $q = 0$ — к критерию «крайнего оптимизма».

Рассмотрим пример использования указанных критериев в играх с природой.

Пример 8.10. Диспетчер автобусного парка (ЛПР) в летние месяцы в конце каждой недели должен принять решение о целесообразности выделения дополнительных автобусов на загородный маршрут. ЛПР имеет три варианта решений: увеличить количество автобусов на 10 (стратегия P_1), увеличить это количество на 5 (стратегия P_2) или оставить без изменения обычное число автобусов на линии (стратегия P_3). Возможны два состояния погоды: Q_1 — плохая погода, Q_2 — хорошая погода, причем в момент принятия решения нет возможности определить ожидаемое состояние погоды. Если в выходные дни будет хорошая погода и много желающих выехать за город, а выделено мало автобусов, то парк понесет убытки, связанные с недополученной прибылью. Если же выделены дополнительные автобусы, а погода окажется плохой, то возникнут потери вследствие эксплуатации незаполненных автобусов.

Пусть на основе анализа статистических данных за определенный период установлена функция потерь для возможных комбинаций состояний природы и решений ЛПР в виде матрицы игры $A(P_iQ_j)$, в которой отрицательные значения показывают дополнительную прибыль, а положительные — потери:

$$A = \begin{array}{c} P_1 \\ P_2 \\ P_3 \end{array} \begin{pmatrix} 4 & -5 \\ 2 & -2 \\ 0 & 6 \end{pmatrix}.$$

Если нет сведений о вероятностях различных состояний погоды, то по критерию Вальда и по критерию Сэвиджа оптимальной является стратегия P_2. По критерию Гурвица при «коэффициенте пессимизма» $q = 1$ оптимальной окажется стратегия P_2, а при $q = 0$ — стратегия P_1.

Рассмотрим в заключение конкретный числовой пример решения задачи принятия решения в экономике методами теории игр.

Пример 8.11. Швейное предприятие, выпускающее детские платья и костюмы, реализует свою продукцию через фирменный магазин. Сбыт продукции зависит от состояния погоды. По данным прошлых наблюдений, предприятие в течение апреля — мая в условиях теплой погоды может реализовать 600 костюмов и 1975 платьев, а при прохладной погоде — 1000 костюмов и 625 платьев. Известно, что затраты на единицу продукции в течение указанных месяцев составили для костюмов 27 руб., дня платьев 8 руб., а цена реализации равна соответственно 48 руб. и 16 руб. (цифры условные).

Задача заключается в максимизации средней величины прибыли от реализации выпущенной продукции с учетом неопределенности погоды в рассматриваемые месяцы. Таким образом, служба маркетинга предприятия должна в этих условиях определить оптимальную стратегию предприятия, обеспечивающую при любой погоде определенный средний доход. Решим эту задачу методами теории игр, игра в этом случае будет относиться к типу игр с природой.

Предприятие располагает в этих условиях двумя чистыми стратегиями: стратегия А — в расчете на теплую погоду и стратегия Б — в расчете на холодную погоду. Природу будем рассматривать как второго игрока также с двумя стратегиями: прохладная погода (стратегия В) и теплая погода (стратегия Г). Если предприятие выберет стратегию А, то в случае прохладной погоды (стратегия природы В) доход составит

$$600(48-27)+625(16-8)-(1975-625)8=6800 \text{ руб.},$$

а в случае теплой погоды (стратегия природы Г) доход будет равен

$$600(48-27)+1975(16-8)=28\,400 \text{ руб.}$$

Если предприятие выберет стратегию Б, то реализация продукции в условиях прохладной погоды даст доход

$$1000(48-27)+625(16-8)=26\,000 \text{ руб.},$$

а в условиях теплой погоды

$$600(48-27)+625(16-8)-(1000-600)27=6800 \text{ руб.}$$

Следовательно, матрица данной игры (платежная матрица) имеет вид

$$A=\begin{pmatrix} 6800 & 28\,400 \\ 26\,000 & 6800 \end{pmatrix}.$$

Первая и вторая строки этой матрицы соответствуют стратегиям А и Б предприятия, а первый и второй столбцы — стратегиям В и Г природы.

По платежной матрице видно, что первый игрок (предприятие) никогда не получит доход меньше 6800 руб. Но если погодные условия совпадают с выбранной стратегией, то выручка (выигрыш) составит 26 000 или 28 400 руб. Отсюда можно сделать вывод, что в условиях неопределенности погоды наибольший гарантированный доход предприятие обеспечит, если будет попеременно применять то стратегию А, то стратегию Б. Такая стратегия, как отмечалось выше, называется смешанной. Оптимизация смешанной стратегии позволит первому игроку всегда получать среднее значение выигрыша независимо от стратегии второго игрока.

Пусть x означает частоту применения первым игроком стратегии А, тогда частота применения им стратегии Б равна $(1 - x)$. В случае оптимальной смешанной стратегии первый игрок (предприятие) получит и при стратегии В (холодная погода), и при стратегии Г (теплая погода) второго игрока одинаковый средний доход:

$$6800x + 26\,000(1 - x) = 28\,400x + 6800(1 - x).$$

Отсюда можно найти, что $x = 8/17$; $1 - x = 9/17$. Следовательно, первый игрок, применяя чистые стратегии А к Б в соотношении 8 : 9, будет иметь оптимальную смешанную стратегию, обеспечивающую ему в любом случае средний доход в сумме $6800 \cdot 8/17 + 26\,000 \cdot 9/17 \approx 16\,965$ руб.; эта величина и будет в данном случае ценой игры.

Легко рассчитать, какое количество костюмов и платьев должно выпускать предприятие при оптимальной стратегии: (600 костюмов + 1975 платьев) · 8/17 + (1000 костюмов + 625 платьев) · 9/17 = 812 костюмов + 1260 платьев. Следовательно, оптимальная стратегия предприятия заключается в выпуске 812 костюмов и 1260 платьев, что обеспечит ему при любой погоде средний доход в сумме 16 965 руб.

8.5. Динамические модели макроэкономики

Макроэкономической моделью (*макромоделью*) называется экономико-математическая модель, отражающая функционирование экономики (региона, страны, мира) как единого целого. Макромодели оперируют такими крупноагрегированными показателями, как национальный доход, валовой внутренний продукт, валовые инвестиции, валовой фонд потребления и др. Эти модели используются для теоретического анализа наиболее общих закономерностей функционирования и развития народного хозяйства. По характеру зависимостей макромодели могут быть детерминированными и вероятностными (стохастическими), по роли временного фактора — статическими и динамическими, по представлению связанных со временем переменных — дискретными и непрерывными. В данном параграфе будем рассматривать некоторые основные динамические макромодели рыночной экономики.

8.5.1. Каноническая модель Кейнса

Основным постулатом экономической теории Кейнса является положение о том, что товаров и услуг производится столько, каков будет спрос на них. Пусть единицей времени является один год, тогда планируемое производство следующего года должно удовлетворить планируемый спрос.

В рассматриваемой модели Кейнса состояние экономики описывается двумя макроэкономическими переменными:

Y_s — валовой внутренний продукт (ВВП);

Y_d — совокупный спрос на товары и услуги.

Первая переменная трактуется как предложение товаров и услуг, а вторая представляет собой сумму двух составляющих спроса C и I:

C — спрос на текущее потребление;

I — спрос на инвестиции.

Существенным допущением модели Кейнса является также то, что текущее потребление C есть возрастающая функция от ВВП:

$$C = C_0 + m \cdot Y_s,$$

при этом предполагается, что спрос меняется медленнее, чем предложение (ВВП), так что величина коэффициента m, называемая *предельной склонностью к потреблению*, удовлетворяет условию $0 < m < 1$.

Пусть до некоторого момента времени t_0 экономика находилась в состоянии равновесия, т.е. совокупный спрос при $t < t_0$ был равен предложению:

$$Y_d(t) = Y_s(t).$$

Предположим, что в момент времени t_0 по какой-либо причине произошло изменение спроса (например, за счет роста спроса на инвестиции). Тогда по принятому в модели допущению «спрос рождает предложение» произойдет соответствующее изменение предложения. Так как предельная склонность к потреблению m меньше единицы, разность $(Y_d - Y_s)$ между спросом и предложением с течением времени сокращается; эту разность принято называть *избыточным спросом на товары и услуги*. Если этот избыточный спрос положителен, то в каждый последующий момент времени происходит рост предложения и избыточный спрос сокращается. При отрицательном избыточном спросе происходит сокращение предложения, что ведет к увеличению избыточного спроса.

При формализации такого процесса в канонической модели Кейнса исходят из того, что в момент времени t_{n+1} ВВП равен совокупному спросу в предыдущий момент t_n: $Y_s(t_{n+1}) = Y_d(t_n)$. Это приводит к следующему конечно-разностному уравнению:

$$Y_s(t_{n+1}) = C_0 + I + m \cdot Y_s(t_n), \tag{8.65}$$

задающему определенный итерационный процесс. Общая теория решения конечно-разностных уравнений позволяет установить, будет ли объем производства неограниченно расти или это уравнения имеет так называемую *стационарную точку*.

В целях упрощения выкладок перейдем к безразмерным переменным:

$$y(t) = Y_s(t)/\ Y_s(t_0); \qquad p = (C_0 + I)/\ Y_s(t_0). \tag{8.66}$$

Тогда в этих переменных первое уравнение можно переписать в виде

$$y(t_{n+1}) = m \cdot y(t_n) + p \tag{8.67}$$

с начальным условием $y(t_0) = 1$.

Из курса дифференциальных и конечно-разностных уравнений известно, что общее решение неоднородного уравнения (8.67) равно сумме частного решения этого уравнения и общего решения соответствующего однородного уравнения $y(t_{n+1}) = m \cdot y(t_n)$. Из этого же математического курса можно получить общее решение рассматриваемого однородного уравнения в виде $y\ (t_n) = A \cdot\ m^{t_n}$, где константа A определяется из начального условия. Частное решение неоднородного уравнения (8.67) определяется из предположения, что это уравнение имеет стационарную точку: $y_c = y(t_{n+1}) = y(t_n)$. Тогда можно найти, что

$$y_c = p/(1 - m) \tag{8.68}$$

и общее решение уравнения (8.67) можно записать в виде

$$y(t_n) = p/(1 - m) + A \cdot\ m^{t_n}.$$

В начальный момент времени $y(t_0) = 1$, поэтому из предыдущего уравнения, выбирая в качестве начала отсчета времени $t_0 = 0$, можно получить, что $A = 1 - p/(1 - m) = 1 - y_c$. С учетом этого результата можно записать полное аналитическое решение уравнения (8.67):

$$y(t_n) = y_c + (1 - y_c\ ) \cdot\ m^{t_n}. \tag{8.69}$$

Из полученного решения следует важный вывод: если выполняется условие $m < 1$, то данный итерационный процесс всегда сходится к стационарной точке независимо от значения $y(t)$, выбранного в качестве начального. Другими словами, это свидетельствует об **устойчивости рынка товаров и услуг** в модели Кейнса.

Существуют графические методы решения уравнений вида (8.67). Один из наиболее простых таких методов известен в экономической теории как *крест Самуэльсона — Хансена*, или *диаграмма Ламерея* (см., например, [1, с. 362]). Численное решение этой задачи можно получить также, используя возможности табличного процессора Excel.

8.5.2. Модель Самуэльсона — Хикса

В модели Кейнса предполагается, что инвестиции в экономику являются постоянной величиной. Это существенно сужает круг реальных макроэкономических процессов, исследуемых с помощью этой модели. В рассматриваемой модели Самуэльсона — Хикса делается попытка освободиться от данного ограничения и предполагается, что инвестиции состоят из постоянной части I_0 и переменной части, пропорциональной приросту ВВП текущего года по сравнению с прошлым:

$$I = I_0 + r \cdot [Y_s(t_n) - Y_s(t_{n-1})],$$

где r — так называемый *коэффициент акселерации инвестиций* $(0< r <1)$. Тогда можно записать уравнение модели Самуэльсона — Хикса, аналогичное уравнению (8.65) модели Кейнса:

$$Y_s(t_{n+1}) = C_0 + I_0 + m \cdot Y_s(t_n) + r \cdot [Y_s(t_n) - Y_s(t_{n-1})]. \quad (8.70)$$

С математической точки зрения модель (8.70) представляет собой линейное конечно-разностное уравнение второго порядка с постоянными коэффициентами. Приведем это уравнение к стандартному виду, для чего перейдем по аналогии с предыдущей моделью к относительным безразмерным переменным

$$y(t) = Y_s(t)/\ Y_s(t_0);\ p = (C_0 + I)/\ Y_s(t_0)$$

и сдвинем начало отсчета времени на один год назад. Кроме того, для упрощения записи дальнейших преобразований обозначим $t_n = n$, тогда уравнение (8.70) примет следующий вид:

$$y_{n+2} - (r + m) \cdot y_{n+1} + r \cdot y_n = p. \quad (8.71)$$

Как и в модели Кейнса, решение данного уравнения представляет собой сумму его частного решения и общего решения соответствующего однородного уравнения. Характер общего решения однородного уравнения зависит от дискриминанта этого уравнения

$$D = (r + m)^2/4 - r.$$

Если $D > 0$, то колебательных решений нет и объем рынка товаров и услуг монотонно стремится к своему асимптотическому значению, равному стационарному решению уравнения (8.71). Это решение может быть найдено из предположения стационарности ($y_{n+2} = y_{n+1} = y_n$):

$$y_n = p/(1 - m), \tag{8.72}$$

что совпадает с соответствующим выводом (8.68) по модели Кейнса.

Если $D < 0$, то в течение некоторого переходного периода решение имеет колебательный затухающий характер, вследствие чего решение уравнения (8.71) с дальнейшим ростом времени стремится к стационарному значению (8.72).

Как отмечалось, в реальных условиях выполняются ограничения на коэффициент акселерации инвестиций и предельную склонность к потреблению: $0 < r < 1$; $0 < m < 1$. Поэтому дискриминант однородного уравнения в рассматриваемой модели может принимать как положительные, так и отрицательные значения. Пример численного расчета по модели Самуэльсона — Хикса с использованием средств пакета Excel при значениях параметров r = 0,495; m = 0,636 (D = − 0,175) осуществлен и представлен графически в [1, с. 365]. Таким образом, можно сделать вывод о том, что модель Самуэльсона — Хикса в состоянии объяснить **наличие колебаний на рынке товаров и услуг**.

8.5.3. Модель Солоу

Динамическое равновесие, описываемое в предыдущих классических моделях, было достаточно устойчивым для ранних стадий развития рыночной экономики, когда эти модели неплохо описывали реальные макроэкономические процессы. Для последующих периодов развития кейнсианские модели роста оказались малопригодны ввиду нестабильности динамического равновесия в этих моделях. Р. Солоу показал, что такая нестабильность является следствием отсутствия взаимозаменяемости факторов производства, поскольку в этих моделях вообще не рассматривались трудовые ресурсы. Рассмотрим основные положения макроэкономической модели Солоу.

Модель Солоу является *односекторной* моделью экономического роста, когда экономическая система рассматривается как единое целое и производит один универсальный продукт. Этот продукт может как потребляться, так и инвестироваться, при этом экспорт и импорт в явном виде не учитываются.

Состояние экономики в модели Солоу задается следующими эндогенными переменными:

Y — валовой внутренний продукт (ВВП);

C — фонд непроизводственного потребления;

I — инвестиции;

L — число занятых в экономике;

K — основные производственные фонды (ОПФ).

Все эндогенные переменные считаются функциями времени; для упрощения записи аргумент времени опущен.

Кроме отмеченных эндогенных переменных, в модели используются также следующие экзогенные параметры:

ν — годовой темп прироста числа занятых в экономике $(-1 < \nu < 1)$;

μ — доля выбывших за год основных производственных фондов $(0 < \mu < 1)$;

ρ — доля валовых инвестиций в ВВП, или *норма накопления* $(0 < \rho < 1)$.

Экзогенные параметры являются управляющими и не зависят от времени.

В модели Солоу предполагается, что годовой выпуск продукции определяется однородной производственной функцией первого порядка Кобба — Дугласа:

$$Y = F(K, L) = A \cdot K^{\alpha} \cdot L^{1-\alpha}, \tag{8.73}$$

где α — коэффициент эластичности ВВП по основным производственным фондам. Определенная таким образом производственная функция удовлетворяет очевидному равенству

$$F(K, L)/L = F(K/L, 1). \tag{8.74}$$

Рассмотрим изменения ресурсных показателей за малый промежуток времени Δt. Согласно определению темпа прироста занятых в экономике

$$\Delta L/L = \nu \cdot \Delta t, \text{ или (при } \Delta t \to 0) \; dL/dt = \nu L. \tag{8.75}$$

Решение дифференциального уравнения (8.75) описывает экспоненциальный рост занятых в экономике:

$$L(t) = L_0 \cdot e^{\nu t}, \tag{8.76}$$

где L_0 — заданное начальное число занятых в экономике. Таким образом, в модели Солоу численность занятых в экономике в любой момент времени t является известной величиной.

Износ ОПФ за время Δt равен $\mu \cdot K \cdot \Delta t$, а прирост инвестиций за это же время составляет $I \cdot \Delta t$. Поэтому прирост ОПФ за время Δt будет равен

$$\Delta K = -\mu \cdot K \cdot \Delta t + I \cdot \Delta t,$$

откуда при $\Delta t \to 0$ можно получить дифференциальное уравнение, описывающее динамику основных производственных фондов:

$$dK/dt = -\mu \cdot K + I \text{ при начальном условии } K(0) = K_0. \tag{8.77}$$

Инвестиции и фонд потребления следующим образом выражаются через ВВП:

$$I = \rho \cdot Y,\ C = (1 - \rho) \cdot Y. \tag{8.78}$$

Таким образом, полную систему уравнений модели Солоу при условии, что ВВП определяется функцией Кобба — Дугласа, образуют уравнения (8.73), (8.76), (8.77) и (8.78):

$$\begin{cases} Y = F(K, L) = A \cdot K^{\alpha} \cdot L^{1-\alpha}; \\ L(t) = L_0 \cdot e^{\nu t}; \\ dK / dt = -\mu \cdot K + I, \quad K(0) = K_0; \\ I = \rho \cdot Y, \quad C = (1-\rho) \cdot Y. \end{cases} \tag{8.79}$$

Систему уравнений (8.79) можно упростить путем деления всех входящих в нее величин на численность занятых в экономике L, т.е. записать эту систему в расчете на одного занятого (см., например, равенство (8.74)).

Систему уравнений модели Солоу можно решить методом последовательных приближений, используя электронные таблицы Excel. Особый интерес представляет поиск стационарных решений (траекторий) модели Солоу, т.е. таких траекторий, на которых фондовооруженность является постоянной величиной и равна, следовательно, своему начальному значению $K(0) = K_0$. Для стационарных решений $dK/dt = 0$, следовательно, на стационарных траекториях выполняются очевидные равенства:

$$I = \mu \cdot K_0,\ Y = \mu \cdot K_0/\rho,\ C = (1 - \rho) \cdot \mu \cdot K_0/\rho. \qquad (8.80)$$

Анализ стационарных решений модели Солоу позволяет сделать вывод, что на стационарной траектории все основные макропоказатели растут по экспоненциальному закону пропорционально трудовым ресурсам $L(t)$.

Если в качестве критерия оптимальности развития экономики принять максимум удельного непроизводственного потребления C/L, то в результате несложных математических выкладок можно получить для оптимальной нормы накопления (доли инвестиций в ВВП) равенство

$$\rho_{\text{опт}} = \alpha. \qquad (8.81)$$

Равенство (8.81) выражает тот факт, что *оптимальная норма накопления в стационарном режиме для производственной функции Кобба — Дугласа равна коэффициенту эластичности по основным производственным фондам*. Этот вывод, полученный в результате анализа стационарных решений модели Солоу, носит название «золотого правила» экономического роста.

Вопросы и задания

1. Раскройте основные понятия целевой функции потребления и кривой безразличия.

2. Что такое «бюджетная линия» и как она связана с кривой безразличия?

3. Укажите наиболее характерные типы кривых Энгеля для различных групп товаров. Поясните характерные свойства функций спроса Торнквиста.

4. Поясните экономический смысл коэффициентов эластичности спроса от дохода, спроса от цен, перекрестных коэффициентов эластичности.

5. В чем суть постановки классической задачи управления запасами?

6. Укажите основные принципиальные системы регулирования запасов и назовите их регулирующие параметры.

7. Перечислите основные предположения и выводы на базе классической модели экономически выгодных размеров заказываемых партий.

8. Приведите примеры систем массового обслуживания в экономике. Из каких элементов состоит СМО?

9. Раскройте суть аналитического и имитационного моделирования СМО. Укажите требования к входящему потоку и времени обслуживания в аналитических моделях СМО.

10. Назовите основные характеристики СМО и укажите методы их расчета для замкнутых и разомкнутых систем.

11. Дайте основные понятия теории игр и приведите примеры экономических задач, которые могут быть решены методами теории игр.

12. Какие парные игры называются матричными? Приведите пример построения платежной матрицы.

13. Поясните принципы использования моделей теории игр в экономических задачах в условиях неопределенности (игры с природой).

14. Дайте определение макроэкономической модели и назовите основные типы таких моделей.

15. В чем заключается основное отличие между макромоделями Кейнса и Самуэльсона — Хикса?

16. Раскройте экономический смысл основных предположений в макромодели Солоу.

17. Сформулируйте «золотое правило» экономического роста, вытекающее из модели Солоу с производственной функцией Кобба — Дугласа.

Упражнения

1. Целевая функция потребления для двух товаров имеет вид $U(Y) = 3y_1^2 y_2^3$, а вектор цен равен $P = (6, 9)$; величину дохода обозначим Z. Построить аналитические функции спроса на товары от дохода $y_1 = f_1(Z)$ и $y_2 = f_2(Z)$.

Указание: вычислить предельные полезности и использовать необходимые условия оптимума целевой функции потребления (соотношения (8.2)).

2. Фирма реализует со склада по заявкам телевизоры, причем ежедневный спрос является случайной величиной с симметричной «треугольной» функцией плотности распределения (см. рис. 8.10, *а*) и колеблется от 30 до 70 телевизоров в день. Средние издержки хранения одного телевизора в день составляют 6 руб., а штраф за недопоставку одного телевизора в день равен 12 руб. Определить стратегию оптимального пополнения запаса телевизоров и минимальные средние полные издержки.

3. Магазин ежедневно продает 100 телевизоров. Накладные расходы на поставку партии телевизоров в магазин оцениваются в 300 руб. Стоимость хранения одного телевизора на складе магазина составляет 6 руб. Определить оптимальный объем партии телевизоров, оптимальные среднесуточные издержки на хранение

и пополнение запасов телевизоров на складе. Чему будут равны эти издержки при объемах партий 50 и 300 телевизоров?

Указание: работу склада принять идеальной и воспользоваться формулой Уилсона (8.39).

4. На АЗС имеются две колонки для заправки автомобилей. Автомобили подъезжают на АЗС в соответствии с пуассоновским распределением со средней частотой два автомобиля за 5 мин. Заправка автомобиля в среднем длится 3 мин, и продолжительность заправки распределена по экспоненциальному закону. Требуется определить:

а) вероятность того, что у АЗС не окажется ни одного автомобиля;

б) вероятность того, что обе колонки будут заняты;

в) среднюю длину очереди в ожидании заправки;

г) среднее время ожидания автомобиля в очереди.

5. Для следующих платежных матриц определить нижнюю и верхнюю цены игры, минимаксные стратегии игроков; найти оптимальное решение игры, если существует седловая точка:

$$\text{а) } A = \begin{pmatrix} 4 & 5 & 3 \\ 6 & 7 & 4 \\ 5 & 2 & 3 \end{pmatrix}; \text{ б) } A = \begin{pmatrix} 8 & 9 & 9 & 4 \\ 6 & 5 & 8 & 7 \\ 3 & 4 & 5 & 6 \end{pmatrix};$$

$$\text{в) } A = \begin{pmatrix} 4 & 5 & 6 & 7 & 9 \\ 3 & 4 & 6 & 5 & 6 \\ 7 & 6 & 10 & 8 & 11 \\ 8 & 5 & 4 & 7 & 3 \end{pmatrix}.$$

Библиографический список

1. *Белый, В. С., Середа, Д. А.* Математика для студентов экономической специальности : учеб. пособие. — М. : Издательство РГСУ, 2010.
2. *Гармаш, А. Н., Орлова, И. В.* Математические методы в управлении : учеб. пособие. — Вузовский учебник, 2012.
3. *Джонстон, Д. Ж.* Эконометрические методы : пер. с англ. — М. : Статистика, 1980.
4. *Дрейпер, И., Смит, Г.* Прикладной регрессионный анализ : пер. с англ. — Кн. 1, 2. — М. : Финансы и статистика, 1986, 1987.
5. *Карасев, А. Л., Кремер, Н. Ш., Савельева, Т. И.* Математические методы и модели в планировании. — М. : Экономика, 1987.
6. *Курицкий, Б. Я.* Поиск оптимальных решений средствами EXCEL 7.0. — СПб. : BHV, 1997.
7. *Лопатников, Л. И.* Экономико-математический словарь. — М. : Наука, 1987.
8. *Орлова, И. В., Половников, В. А., Федосеев, В. В.* Курс лекций по экономико-математическому моделированию. — М. : Экономическое образование, 1993.
9. *Федосеев, В. В.* Экономико-математические модели и прогнозирование рынка труда : учеб. пособие. — 2-е изд., доп. и исправ. — М. : Вузовский учебник, 2010.
10. *Прицкер, А.* Введение в имитационное моделирование и язык СЛАМП : пер. с англ. — М. : Мир, 1987.
11. Статистическое моделирование и прогнозирование / под ред. А. Г. Гранберга. — М. : Финансы и статистика, 1990.
12. Справочник по математике для экономистов / под ред. В. И. Ермакова. — М. : Высшая школа, 1987.
13. *Терехов, Л. Л.* Кибернетика для экономистов. — М. : Финансы и статистика, 1983.
14. *Федосеев, В. В., Эриашвили, Н. Д.* Экономико-математические методы и модели в маркетинге : учеб. пособие /

под ред. В. В. Федосеева. — 2-е изд., перераб. и доп. — М. : ЮНИТИ-ДАНА, 2001.

15. *Френкель, А. А.* Прогнозирование производительности труда: методы и модели. — М. : Экономика, 1989.

16. *Четыркин, Е. М.* Статистические методы прогнозирования. — М. : Финансы и статистика, 1979.

17. Экономико-математические методы и прикладные модели : учеб. пособие / под ред. В. В. Федосеева. — 2-е изд., перераб. и доп. — М. : ЮНИТИ-ДАНА, 2005.

Покупайте наши книги:

В офисе издательства «ЮРАЙТ»:
111123, г. Москва, ул. Плеханова, д. 4,
тел.: (495) 744-00-12, e-mail: sales@urait.ru, www.urait.ru

В логистическом центре «ЮРАЙТ»:
140053, Московская область, г. Котельники, мкр. Ковровый, д. 37,
тел.: (495) 744-00-12, e-mail: sales@urait.ru, www.urait.ru

В интернет-магазине «ЮРАЙТ»: www.urait-book.ru,
e-mail: order@urait-book.ru, тел.: (495) 742-72-12

Для закупок у Единого поставщика в соответствии с Федеральным законом
от 21.07.2005 № 94-ФЗ обращайтесь по тел.: (495) 744-00-12,
e-mail: sales@urait.ru, vuz@urait.ru

Учебное издание

Федосеев Владилен Валентинович,
Гармаш Александр Николаевич,
Орлова Ирина Владленовна

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ
И ПРИКЛАДНЫЕ МОДЕЛИ

Учебник для бакалавров

Под редакцией *В. В. Федосеева*

Формат 84×108 $^1/_{32}$.
Гарнитура «Petersburg». Печать офсетная.
Усл. печ. л. 17,22. Доп. тираж 1000 экз. Заказ №

ООО «Издательство Юрайт»
111123, г. Москва, ул. Плеханова, д. 4.
Тел.: (495) 744-00-12. E-mail: izdat@urait.ru, www.urait.ru