Е.В. БЕРЕЖНАЯ, В.И. БЕРЕЖНОЙ

МАТЕМАТИЧЕСКИЕ МЕТОДЫ МОДЕЛИРОВАНИЯ ЭКОНОМИЧЕСКИХ СИСТЕМ

Е.В. БЕРЕЖНАЯ, В.И. БЕРЕЖНОЙ

МАТЕМАТИЧЕСКИЕ МЕТОДЫ МОДЕЛИРОВАНИЯ ЭКОНОМИЧЕСКИХ СИСТЕМ

Издание второе, переработанное и дополненное

Рекомендовано
Учебно-методическим объединением (УМО) вузов
по специальностям "Финансы и кредит", "Бухгалтерский учет, анализ и аудит", "Мировая экономика"
в качестве учебного пособия для студентов
высших учебных заведений

Москва
"Финансы и статистика"
2006

УДК 330.45:519.86(075.8)
ББК 65.050в6я73
Б48

РЕЦЕНЗЕНТЫ:

Кафедра прикладной математики и информатики
Ставропольского государственного университета;

Л.Г. Лабскер,
профессор кафедры математического моделирования
экономических процессов
Финансовой академии при Правительстве РФ

Бережная Е.В., Бережной В.И.
Б48 Математические методы моделирования экономических систем: Учеб. пособие. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2006. – 432 с.: ил.
ISBN 5-279-02940-8

Рассматривается моделирование экономических систем с использованием марковских случайных процессов, моделирование систем массового обслуживания, методы и модели корреляционно-регрессионного анализа и прогнозирования временны́х рядов экономических показателей. Приводятся оптимизационные методы и модели в управлении экономическими системами, линейное, динамическое, параметрическое и целочисленное программирование, а также транспортные задачи линейного программирования, теория игр и принятие решений.

Для преподавателей, аспирантов, студентов экономических вузов и факультетов, менеджеров.

Б $\frac{0601000000-107}{010(01)-2006}$ 68 – 2006

УДК 330.45:519.86(075.8)
ББК 65.050в6я73

ISBN 5-279-02940-8

ПРЕДИСЛОВИЕ

Методы экономико-математического моделирования, возможности применения которых существенно расширились благодаря современному программному обеспечению ПЭВМ, представляют собой один из наиболее динамично развивающихся разделов прикладной экономической науки.

Современный экономист должен хорошо разбираться в экономико-математических методах, уметь их практически применять для моделирования реальных экономических ситуаций. Это позволит лучше усвоить теоретические вопросы современной экономики, повысить уровень квалификации и общей профессиональной культуры специалиста.

В учебном пособии систематически излагаются методы экономико-математического моделирования, которые широко используются в различных областях экономики, при принятии управленческих решений в финансовой сфере в силу разработанности математического аппарата и возможности практической реализации.

Пособие включает двенадцать глав, которые объединены в два раздела.

Раздел I посвящен вероятностно-статистическим методам моделирования экономических систем, а также теоретическим основам вероятностных методов. Авторы излагают те вопросы теории вероятностей и математической статистики, знание которых является необходимым минимумом для усвоения материала, рассматриваемого в последующих главах.

Значительное место отведено применению марковских случайных процессов для моделирования экономических систем, а также использованию аппарата теории массового обслуживания для решения финансово-экономических задач. Далее авторы рассматривают возможности применения метода статистического моделирования (метода Монте-Карло).

Достаточно подробно рассмотрены методы и модели корреляционно-регрессионного анализа. Регрессионный и корреляционный анализ находит широкое применение при исследовании зависимостей и взаимосвязей между явлениями в экономике, при прогнозировании и решении задач бизнес-планирования. В настоящее время большинство объективно существующих зависимостей между финансово-экономическими явлениями исследованы и изучены теоретически. Значительно важнее количественно измерить тесноту причинно-следственных связей в экономике и финансах, понять природу исследуемых процессов. Это позволит воздействовать на выявленные факторы, вмешиваться в соответствующий экономический процесс с целью получения нужных результатов. В связи с этим к аппарату корреляционно-регрессионного анализа в ходе своих исследований обращаются как экономисты-практики, так и научные работники.

Внимание к методам корреляционно-регрессионного анализа особенно возросло в связи с появлением современных программных продуктов для компьютеров, реализующих эти и другие математико-статистические методы. Если раньше пакеты прикладных программ по математико-статистическим методам были ориентированы в основном на профессиональных пользователей (математиков-прикладников), то широко распространенные сегодня табличные процессоры Excel, входящие в известный продукт MS Office, не требуют от исследователя подготовки, выходящей за рамки экономического вуза.

Учебное пособие – практическое руководство по корреляционно-регрессионному анализу, которое поможет студентам, аспирантам, менеджерам овладеть этими методами анализа. Рассмотренные методы и модели прогнозирования временных рядов экономических показателей особенно актуальны для современных студентов экономических специальностей.

Раздел II посвящен методам оптимизации: линейному, динамическому, параметрическому, целочисленному программированию, теории игр.

Наряду со сведениями теоретического характера в пособии разбирается большое количество примеров и задач, цель которых – уяснение основных понятий и математических методов. В конце каждой главы приводятся задачи для самостоятельного решения. Задачи подобраны и составлены с особой тщательностью и могут служить для проверки степени усвоения читателем изученного материала. Примеры и задачи предусматривают небольшой объем вычислений и могут быть использованы на практических занятиях при изучении курса «Экономико-математические методы».

Дополнительные теоретические сведения для более глубокого изучения того или иного раздела можно получить из книг, приведенных в списке литературы.

Изучение всех разделов экономико-математических методов, излагаемых авторами в учебном пособии, предусмотрено Государственным образовательным стандартом по экономическим специальностям.

Пособие написано на основе многолетнего опыта преподавания экономико-математических методов и моделей в высших учебных заведениях, а также на основе решения ряда практических задач, которые встречались авторам в научно-исследовательской работе.

Авторы выражают благодарность уважаемым рецензентам и признательны им за ценные замечания, которые улучшили изложение материала, – кафедре прикладной математики и информатики Ставропольского государственного университета и профессору кафедры математического моделирования экономических процессов Финансовой академии при Правительстве РФ Л. Г. Лабскеру.

РАЗДЕЛ I

Вероятностно-статистические методы моделирования экономических систем

Глава 1
Основы вероятностных методов анализа и моделирования экономических систем

1.1. Элементарные понятия о случайных событиях, величинах и функциях

Под событием понимается всякий факт, который может произойти в данных условиях. Теория вероятностей рассматривает события в тесной связи с теми условиями, в которых они наступают. Совокупность условий, в которых рассматривается данное событие, называют *комплексом условий*, а реализацию этого комплекса условий на практике – *испытанием*. В зависимости от связи между событиями и соответствующими комплексами условий различают достоверные, невозможные и случайные события.

Достоверным называется такое событие, которое наступает каждый раз при реализации данного комплекса условий. Достоверное событие обозначим через U.

Невозможным называется событие, которое никогда не наступает при реализации данного комплекса условий. Невозможное событие обозначим символом $\varnothing$.

Случайным называется событие, которое может либо наступить при реализации данного комплекса условий, либо не наступить. Достоверное и невозможное события могут рассматриваться как крайние частные случаи случайных событий. Случайные события обозначим через *A, B, C*...

Согласно теоретико-множественному подходу при рассмотрении понятия «случайное событие» введем понятие «элементарное событие».

Элементарное событие – это один из нескольких возможных, но несовместных исходов того или иного опыта (испытания). Совокупность или множество их составляют пространство элементарных событий.

В общем случае пространство элементарных событий может быть любой природы: конечным и бесконечным, дискретным и непрерывным. Пространство элементарных событий является синонимом достоверного события, так как один из его элементов непременно наступит. Кроме того, существует понятие «пустое множество». Это множество, не содержащее элементарных событий. Очевидно, что пустое множество является синонимом невозможного события. При изучении случайных событий в ходе разработки математических моделей экономических систем используется, как правило, не одно, а группа событий, между которыми существуют определенные соотношения, позволяющие выражать одни события через другие.

Рассмотрим эти соотношения.

1. Событие А содержится в событии В (А ⊂ В). Если при каждом испытании, при котором происходит событие A, непременно происходит и событие B, то говорят, что событие A содержится в событии B, или принадлежит событию B.

2. Тождественные события (А = В). Если событие A содержится в событии B, а событие B содержится в событии A, то говорят, что события A и B тождественны, или равносильны.

3. Произведение событий. Произведением (или пересечением) событий A и B называется событие C, состоящее в совместном наступлении этих событий. Другими словами, множество C содержит элементы, принадлежащие множествам A и B. Произведение событий записывается в виде:

$$C = A \cdot B \text{ или } C = A \cap B, \quad A = A \cdot A, \tag{1.1}$$

где $\cap$ – знак пересечения.

4. Несовместные события. События A и B называются несовместными, если их совместное появление при испытании невозможно. Условие несовместности записывается в виде:

$$A \cdot B = \varnothing. \tag{1.2}$$

5. Сумма событий (объединение событий). Суммой событий A и B называется событие C, состоящее в наступлении хотя бы одного из этих событий. Множество C содержит элементы, принадлежащие хотя бы одному из множеств A или B:

$$C = A + B \text{ или } C = A \cup B, \qquad (1.3)$$
$$A = A + A,$$

где $\cup$ – знак объединения.

*6. **Полная группа событий.*** События A и B составляют полную группу событий, если при реализации заданного комплекса условий непременно появится хотя бы одно из этих событий. Сумма всех таких событий есть событие достоверное:

$$C = A + B = U. \qquad (1.4)$$

*7. **Противоположное событие.*** Два события A и $\bar{A}$ (читается «не A») называются противоположными, если они составляют полную группу несовместных событий, т.е. удовлетворяют условию:

$$A + \bar{A} = U; \; A \cdot \bar{A} = \varnothing. \qquad (1.5)$$

Всякому событию при данном комплексе условий соответствует определенная степень возможности. Более возможные события при многократных испытаниях в среднем наступают чаще, а менее возможные – реже. *Частотой события* называется отношение числа испытаний, в которых появилось данное событие, и общего числа испытаний. Частота события A равна:

$$P^*(A) = \frac{m(A)}{n}, \qquad (1.6)$$

где n – общее число проведенных испытаний;
$m(A)$ – число испытаний, в которых наступило событие A.

Частота достоверного события U равна единице:

$$P^*(U) = \frac{n}{n} = 1.$$

Частота невозможного события равна нулю:

$$P^*(\varnothing) = \frac{0}{n} = 0.$$

Частота случайного события A находится в интервале $[0;1]$:

$$0 \le P^*(A) \le 1.$$

Следует отметить, что частота случайного события обладает устойчивостью, что доказывается и формулируется в теореме Я. Бернулли, относящейся к закону больших чисел.

Свойство устойчивости частоты случайного события отражает связь между комплексом условий и возможностью наступления событий при данном комплексе. Количественной мерой степени возможности появления события для заданного комплекса условий является *вероятность события*. Чем более возможно появление случайного события, тем больше его вероятность. Наоборот, чем менее возможно появление события, тем меньше его вероятность.

Вероятность и частота события тесно связаны между собой. Зная частоту, вычисленную при достаточно большом числе испытаний, есть все основания считать ее близкой к соответствующей вероятности и полагать, что

$$P(A) \cong P^*(A) = \frac{m(A)}{n}. \tag{1.7}$$

Такой способ определения вероятности события $P(A)$ называется *статистическим*.

Свойства вероятностей событий

1. Вероятность невозможного события равна нулю, т. е.

$$P(\varnothing) = 0.$$

2. Для любого события A вероятность противоположного события $\bar{A}$ равна:

$$P(\bar{A}) = 1 - P(A). \tag{1.8}$$

3. Если событие A влечет за собой событие B, т. е. $A \subset B$, то

$$P(A) \leq P(B). \tag{1.9}$$

4. Вероятность события A заключена между нулем и единицей, т. е.

$$0 \leq P(A) \leq 1. \tag{1.10}$$

5. Вероятность двух событий A и B равна сумме вероятностей этих событий без вероятности их произведения:

$$P(A + B) = P(A) + P(B) - P(AB). \tag{1.11}$$

Вероятность события определяется при условии реализации некоторой совокупности условий. Если никаких ограничений, кроме упомянутых условий, при вычислении вероятности $P(A)$ не налагается, то такие вероятности называются *безусловными*. Однако в ряде случаев приходится находить вероятности событий при условии, что произошло некоторое событие B, имеющее положительную вероятность. Такие вероятности называются *условными* и обозначаются $P(A/B)$.

Событие A называется *независимым* от другого события B, если вероятность события A не изменяется от того, наступает событие B или нет. В противоположном случае событие A называется *зависимым* от события B. Следовательно, если события A и B независимые, то $P(A/B) = P(A)$.

Вероятность произведения двух событий равна произведению вероятности одного из этих событий на условную вероятность другого при условии, что первое произошло:

$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B). \tag{1.12}$$

Вероятность произведения независимых событий равна:

$$P(AB) = P(A) \cdot P(B). \tag{1.13}$$

Вероятность произведения n случайных событий равна произведению вероятности одного из них на условные вероятности остальных, вычисленных при условии, что все предшествующие события произошли.

Правило сложения вероятностей двух событий записывается следующим образом:

$$P(A + B) = P(A) + P(B) - P(AB). \tag{1.14}$$

Читается это правило так: вероятность наступления хотя бы одного из двух событий равна сумме вероятностей этих событий без вероятности их совместного наступления.

Если события несовместны, то правило сложения вероятностей принимает вид:

$$P(A + B) = P(A) + P(B). \tag{1.15}$$

Если несовместные события составляют полную группу, т. е.

$$A_1 + A_2 + \ldots + A_n = U \text{ и } A_iA_j = \varnothing,\ i \neq j,$$

то
$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) = 1. \tag{1.16}$$

Случайные события могут быть представлены через случайные величины. Понятие «случайная величина» расширяет область применения вероятностных методов в решении практических задач, позволяет исследовать более сложные случайные явления. *Случайной* называется такая величина, которая в результате испытания (реализации определенного комплекса условий) может принять то или иное значение, причем до испытания неизвестно, какое именно. Если повторять испытания, то результатом каждого будет какое-либо одно значение случайной величины из множества возможных.

Случайные величины подразделяются на дискретные и непрерывные.

Множество значений *дискретной случайной величины* конечно или счетно, например:

• количество отказов автомобилей автопредприятия в течение рабочей смены;

• число рабочих, пришедших в бухгалтерию завода в течение одного часа получать заработную плату и т. д.

Множество значений *непрерывной случайной величины* представляет собой множество всех точек, принадлежащих какому-либо интервалу числовой оси, например:

• расход топлива на километр пробега;

• время безотказной работы автомобиля и т. д.

Кроме дискретной и непрерывной случайных величин встречаются *случайные величины смешанного типа*, для которых наряду с участками непрерывных значений имеются отдельные, изолированные значения.

Для того чтобы задать случайную величину, необходимо задать множество значений, которые она может принимать. Однако одного перечня значений случайной величины еще недостаточно для каких-либо существенных выводов. Нужно еще знать, как часто, т.е. с какой вероятностью, она принимает эти значения. Ответ на поставленный вопрос дает исчерпывающая характеристика случайной величины – закон ее распределения.

Закон распределения представляет собой соотношение, позволяющее определить вероятность появления случайной величины в любом интервале (и, в частности, вероятности любых значений случайной величины).

Основными формами закона распределения являются: ряд распределения, функция распределения и плотность распределения.

Ряд распределения представляет собой таблицу, в которой перечислены возможные значения случайной величины и соответствующие им вероятности:

Таблица 1.1

Ряд распределения

x_i	x_1	x_2	...	x_n
p_i	p_1	p_2	...	p_n

В табл. 1.1 x_i – i-е значение случайной величины X;

p_i – вероятность появления i-го значения случайной величины X. При этом $\sum_i p_i = 1$.

Эмпирический ряд распределения представляет собой таблицу, в которой перечислены наблюдаемые значения (фактические реализации) случайной величины и соответствующие им частоты (табл. 1.2).

Таблица 1.2

Эмпирический ряд распределения

x_i	x_1	x_2	x_3	...	x_n
m_i	m_1	m_2	m_3	...	m_n

В табл. 1.2 x_i – i-я фактическая (наблюдаемая) реализация случайной величины X;

m_i – количество появлений (частота) величины x_i.

Ряды распределения, образованные из значений случайной величины, характеризующей качественный признак, называются *атрибутивными*. Ряды распределений, образованные из значений случайной величины, характеризующей количественный признак явления (события), называются *вариационными*.

Ряд распределения не может служить характеристикой непрерывной случайной величины, поскольку значения этой величины нельзя перечислить, так как множество их несчетно. Кроме того, вероятность отдельного значения непрерывной случайной величины равна нулю.

Для характеристики непрерывной случайной величины определяют вероятность появления значения случайной величины меньшего x, где x – текущая переменная, т. е. определяют вероятность события $X < x$. Вероятность этого события зависит от x, т. е. является функцией x. Эта функция называется ***функцией распределения случайной величины*** X и обозначается $F(x)$:

$$F(x) = P(X < x). \tag{1.17}$$

Таким образом, функцией распределения случайной величины X называется функция аргумента x, равная вероятности того, что случайная величина X примет любое значение, меньшее x.

Вероятность попадания случайной величины в полузамкнутый интервал $[a, b)$ равна разности значений функции распределения в точках b и a:

$$P(a \leq x < b) = F(b) - F(a). \tag{1.18}$$

Функция распределения есть неубывающая функция, значения которой начинаются с нуля и доходят до единицы, причем в отдельных случаях функция может иметь скачки – разрывы. Функцию распределения дискретной случайной величины можно определить, зная ее ряд распределения, по формуле

$$F(x) = \sum_{x_i < x} P(x_i), \tag{1.19}$$

где суммирование распространяется на значения x_i, которые меньше x.

Следует отметить, что функция распределения дискретной случайной величины увеличивается скачками каждый раз, когда X при своем изменении проходит через какое-нибудь из возможных значений x, причем величина скачка равна вероятности этого значения. Между двумя соседними значениями величины X функция $F(x)$ постоянна.

Поскольку для непрерывной случайной величины нельзя использовать в качестве характеристики вероятность появления ее отдельных значений, то определяют вероятность появления случайной величины в пределах малого интервала $[x, x + \Delta x)$, примыкающего к x. Разделив эту вероятность на длину интервала Δx, находят среднюю плотность вероятности и при неограниченном уменьшении длины интервала переходят к пределу, который является плотностью распределения в точке x:

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x \leq X < x + \Delta x)}{\Delta x}. \tag{1.20}$$

Плотность распределения $f(x)$ есть предел отношения вероятности попадания случайной величины на малый участок и длины этого участка при ее неограниченном уменьшении.

Вероятность попадания случайной величины на произвольный участок $[a, b)$ равна:

$$P(a \leq X < b) = \int_a^b f(x) \cdot dx. \tag{1.21}$$

Интеграл в бесконечных пределах от плотности распределения равен единице, т. е. $\int_{-\infty}^{\infty} f(x) \cdot dx = 1.$ Это очевидно, так как указанный интеграл выражает вероятность достоверного события – попадания случайной величины на участок от $-\infty$ до ∞, а значит, равен единице.

График плотности распределения называется кривой распределения, лежащей в верхней полуплоскости. Кривая распределения совместно с осью абсцисс ограничивает площадь, равную единице (рис. 1.1).

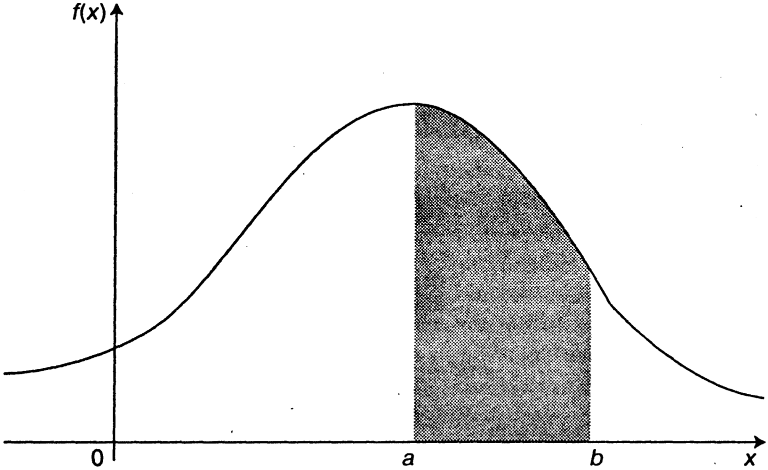


Рис. 1.1. График плотности распределения (кривая распределения)

Вероятность попадания на участок $[a, b)$ равна площади ограниченной кривой распределения, опирающейся на участок $[a, b)$ (на рис. 1.1 – заштрихованная площадь).

Плотность распределения есть производная функции распределения. С другой стороны,

$$F(x) = P(X < x) = P(-\infty < X < x),$$

откуда

$$F(x) = \int_{-\infty}^{x} f(x) \cdot dx. \quad (1.22)$$

Величину $F(x)$ называют *интегральной функцией распределения величины* X. Величина $f(x)$ – дифференциальная функция распределения случайной величины X. Для оценки особенностей законов распределения случайных величин определяют числовые характеристики этих величин.

1.2. Числовые характеристики случайных величин

При решении многих практических задач часто достаточно указать отдельные числовые характеристики, определяющие особенности того или иного распределения случайной величины. Это прежде всего *среднее значение*, которое принадлежит к характеристикам положения случайной величины, т. е. представляет такую величину, относительно которой каким-то образом группируются, рассеиваются всевозможные значения случайной величины.

Среднее значение, или математическое ожидание дискретной случайной величины, вычисляется по формуле

$$M[x] = m_x = a = \sum_{i=1}^{n} x_i p_i, \qquad (1.23)$$

где x_i – возможные значения случайной величины X;

p_i – вероятность появления i-го возможного значения случайной величины X.

Математическое ожидание является теоретической характеристикой случайной величины.

Эмпирической характеристикой случайной величины является *эмпирическая средняя*, вычисляемая по формуле

$$\overline{X} = M^*[X] = \frac{x_1 \cdot m_1 + x_2 \cdot m_2 + \ldots + x_n \cdot m_n}{N} = \sum_{i=1}^{n} x_i \cdot \frac{m_i}{N}, \qquad (1.24)$$

или

$$M^*[X] = \sum_{i=1}^{n} x_i \cdot p^*(x_i),$$

где $p^*(x_i) = \frac{m_i}{N}$ – частота значений x_i при N наблюдениях (испытаниях);

$N = \sum_{i=1}^{n} m_i$;

m_i – количество появлений значений x_i при N наблюдениях.

Эмпирическая средняя случайной величины по мере увеличения испытаний (наблюдений) приобретает тенденцию стабилизироваться относительно постоянной величины – математического ожидания.

Для непрерывной случайной величины X математическое ожидание определяется интегралом:

$$M[X] = a = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx. \qquad (1.25)$$

Кроме математического ожидания на практике иногда применяются и другие характеристики положения, в частности медиана и мода случайной величины.

Медианой Me случайной величины называется такая величина, относительно которой равновероятно получение большего или меньшего значения случайной величины:

$$P(X > \mathrm{Me}) = P(X < \mathrm{Me}). \qquad (1.26)$$

Медиану применяют в качестве характеристики ряда распределения в тех случаях, когда имеются очень большие колебания случайной величины. В этом случае на эмпирическую среднюю $M^*[X]$ будут оказывать сильное влияние крайние значения случайной величины, а медиана менее чувствительна к крайним значениям случайной величины. На медиану влияет не столько колебание в значениях случайной величины X, сколько колебания в частоте появления того или иного значения случайной величины. Медиану необходимо вычислять в дополнение к математическому ожиданию в случае распределений, имеющих большую скошенность[1].

Модой Мо дискретной случайной величины называется ее значение, обладающее наибольшей вероятностью. Для непрерывной случайной величины мода есть такое значение, которое отвечает максимальной плотности распределения.

В общем случае математическое ожидание, медиана и мода не совпадают. В частном случае при симметричном распределении все три характеристики положения случайной величины совпадают.

Для оценки степени разброса, рассеивания значений случайной величины относительно среднего вычисляют следующие характеристики:

- дисперсию;
- среднее квадратическое отклонение;
- коэффициент вариации.

Дисперсией называется математическое ожидание квадрата отклонений случайной величины от своего математического ожидания:

$$D_x = \sigma_x^2 = M[(X - m_x)^2]. \qquad (1.27)$$

[1] О типах распределений см. подразд. 1.3.

Чем больше дисперсия, тем в среднем больше отклонение значений случайной величины относительно математического ожидания, т. е. будет больше рассеивание случайной величины.

Дисперсия дискретной случайной величины вычисляется по формуле

$$\sigma_x^2 = \sum_{i=1}^{n} (x_i - m_x)^2 \cdot p(x_i). \tag{1.28}$$

Дисперсия непрерывной случайной величины равна:

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 \cdot f(x) \cdot dx. \tag{1.29}$$

Наряду с дисперсией случайной величины, в качестве характеристики рассеивания случайной величины используется *среднее квадратическое отклонение*, которое равно положительному значению корня квадратного из дисперсии.

Среднее квадратическое отклонение имеет одинаковую размерность со случайной величиной, в этом состоит ее преимущество относительно дисперсии.

Эмпирические значения характеристик рассеивания вычисляют по формулам:

дисперсия

$$\sigma_x^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \frac{m_i}{N}; \tag{1.30}$$

среднее квадратическое отклонение

$$\sigma_x = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \frac{m_i}{N}}. \tag{1.31}$$

Для малых выборок, если число испытаний (наблюдений) не превышает $N \leq 30$, то характеристики рассеивания вычисляются по формулам:

дисперсия

$$\sigma_x^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \frac{m_i}{N-1}; \tag{1.32}$$

среднее квадратическое отклонение

$$\sigma_x = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \frac{m_i}{N-1}}. \tag{1.33}$$

Величины σ_x^2 и σ_x показывают абсолютное отклонение от среднего значения случайной величины, что недостаточно характеризует уровень ее рассеивания. Относительной характеристикой рассеивания является *коэффициент вариации*, вычисляемый как отношение среднего квадратического отклонения и эмпирической средней

$$V = \frac{\sigma_x}{\bar{x}} \cdot 100\%, \tag{1.34}$$

или

$$V = \frac{\sigma_x}{\bar{x}}. \tag{1.35}$$

Коэффициент вариации может использоваться для сравнения меры рассеивания (колеблемости) случайных величин, имеющих различную размерность.

1.3. Статистическая оценка законов распределения случайных величин

Эмпирические ряды распределения, получаемые при обработке первичных статистических данных, оформляются в таблицах или изображаются графически посредством геометрических образов – точек, линий и фигур в различных сочетаниях. Построение эмпирических графиков и диаграмм позволяет установить на первом этапе исследования к какому типу теоретических распределений ближе всего полученное эмпирическое распределение, что облегчает выбор конкретных технических приемов обработки исходных данных.

Для применения графического метода анализа распределений необходимо знать, как строить графики распределения, какие существуют типы распределений и какими свойствами обладают теоретические распределения.

Покажем, каким образом производится обработка статистического материала для нахождения законов распределения случайной величины. Для этого будем рассматривать некоторую случайную величину X. При функционировании экономической системы или ее элемента в течение некоторого времени t случайная величина X может принять n определенных значений. Совокупность этих случайных значений случайной величины в математической статистике называется *статистической выборкой* объема n. Если расположить отдельные значения случайной величины X в возрастающем или убывающем порядке и указать относительно каждого значения, как часто оно встречалось в данной совокупности, то получится эмпирическое распределение случайной величины, или вариационный ряд, на основании которого определяются аналитическая форма неизвестной плотности вероятности $f(x)$, функция распределения $F(x)$ и оцениваются входящие в нее параметры.

Рассмотрим подробнее процедуру построения вариационного ряда.

Весь диапазон значений непрерывной случайной величины X разбивается на интервалы. Далее подсчитывается количество значений m_i случайной величины X, приходящейся на каждый интервал, и определяется частота ее попадания в данный интервал по формуле

$$p_i^* = \frac{m_i}{n}. \tag{1.36}$$

Если случайная величина X принимает значение, попадающее на границу i-го и $(i + 1)$-го интервалов, то это значение учитывается в числе попаданий в $(i + 1)$-й интервал.

Определив таким образом частоты попадания случайной величины X в каждый интервал, получим вариационный (статистический) ряд, который представлен в табл. 1.3.

Таблица 1.3

Вариационный ряд

Интервал	$t_1 - t_2$	$t_2 - t_3$	...	$t_i - t_{i+1}$	...	$t_k - t_{k+1}$
Частота p_i^*	p_1^*	p_2^*	...	p_i^*	...	p_k^*

Оптимальная длина интервала определяется по формуле

$$\Delta x = \frac{x_{\max} - x_{\min}}{1 + 3{,}21 \cdot \lg n}, \tag{1.37}$$

где $x_{\max} - x_{\min}$ – размах вариации случайной величины X.

Число интервалов будет равно:

$$k = \frac{x_{\max} - x_{\min}}{\Delta x}. \tag{1.38}$$

Если k не целое число, то в качестве числа интервалов надо взять ближайшее к k целое число, не меньшее k.

Вариационные ряды могут быть изображены графически в виде полигона распределения и гистограммы.

Полигон распределения представляет собой многоугольник, который строится на прямоугольной координатной сетке следующим образом. В выбранных масштабах на оси абсцисс наносится шкала для фактических значений случайной величины X, на оси ординат — для частот $p^* = \frac{m}{n}$ (рис. 1.2). Пользуясь этими шкалами, наносят точки M_i с координатами x_i и $\frac{m_i}{n}$. Точки $M_1\left(x_1, \frac{m_1}{n}\right)$, $M_2\left(x_2, \frac{m_2}{n}\right)$, ..., $M_k\left(x_k, \frac{m_k}{n}\right)$ соединяют ломаной линией M_1 M_2 M_3 ... M_i ... M_k. Крайние точки M_1 и M_k, если они не лежат на оси $0x$, соединяют также со смежными точками соответственно $M_0(x_0, 0)$ и $M_{k+1}(x_{k+1}, 0)$ на оси абсцисс. Полученный таким образом многоугольник M_0 M_1 M_2 ... M_i ... M_k M_{k+1} является полигоном распределения.

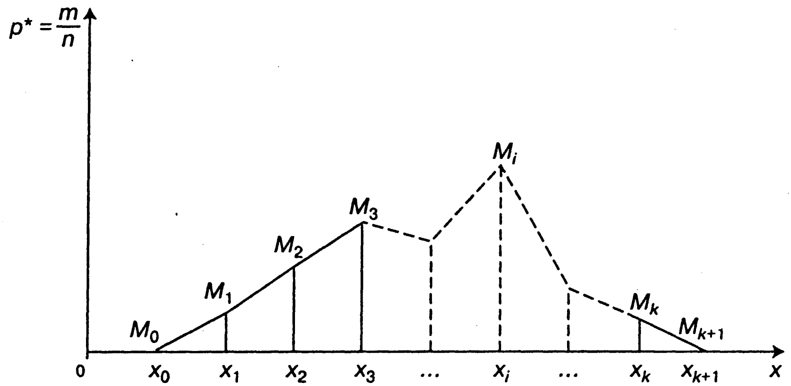


Рис. 1.2. Полигон распределения реализаций случайной величины X

Полигоны распределения чаще всего применяются для изображения дискретных вариационных рядов.

Гистограмма распределения реализаций случайной величины применяется для графического изображения интервальных рядов распределения. Она представляет собой многоугольник, построенный с помощью смежных прямоугольников. В случае непрерывных равных интервалов с шириной интервала Δx гистограмма строится следующим образом (рис. 1.3).

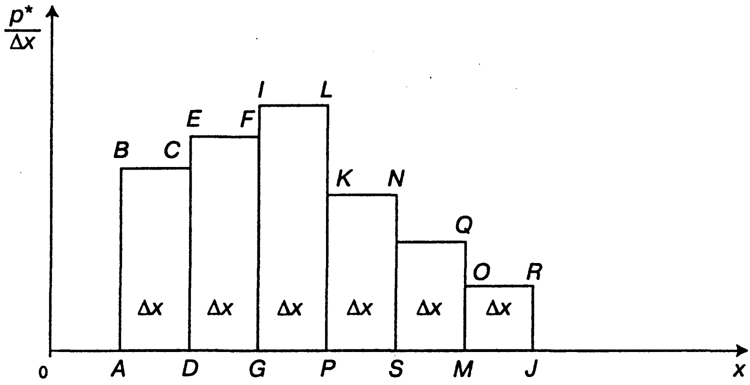


Рис. 1.3. Гистограмма распределения

В выбранных масштабах на оси абсцисс наносится шкала для реализаций случайной величины X, на оси ординат – величины $\frac{p^*}{\Delta x}$. Пользуясь этими шкалами, строят прямоугольники $ABCD$, $DEFG$, ..., основания которых соответствуют ширине интервала Δx, а высоты равны отношениям $\frac{p_1^*}{\Delta x}, \frac{p_2^*}{\Delta x}, ..., \frac{p_k^*}{\Delta x}$. Многоугольник $ABCEF$ ... $QORJA$ и является гистограммой распределения.

Гистограммы чаще всего применяются для изображения вариационных рядов с непрерывными значениями случайной величины X. При уменьшении величины каждого интервала гистограмма будет приближаться к некоторой плавной кривой, соответствующей графику функции плотности распределения случайной величины X. Следовательно, в результате построения гистограммы можно получить представление о дифференциальном законе распределения случайной величины X.

Эмпирическая (статистическая) функция распределения строится следующим образом. Над каждым отрезком оси абсцисс (Δx), изображающим расстояние между концами интервалов, проводится отрезок горизонтальной прямой на уровне ординаты, равной величине накопленной частоты; концы горизонтальных отрезков соединяются вертикальными линиями.

Статистическая функция распределения $F^*(X)$ представляет собой частоту событий $X < x$ в данной выборке:

$$F^*(x) = P^*(X < x) = \sum_{x_i < x} p^*(X < x_i), \qquad (1.39)$$

где x – текущая переменная;

p^* – частота, или статистическая вероятность, события.

Неравенство $x_i < x$ под знаком суммы указывает, что суммирование распространяется на все те значения x_i, которые меньше x.

Значения $F^*(x_i)$ при данном значении x_i определяется по формуле

$$F^*(x_i) = \frac{n_i}{n}, \qquad (1.40)$$

где n_i – число опытов, при которых $X < x_i$.

При неограниченном увеличении числа опытов (наблюдений) n согласно теореме Я. Бернулли при любом x_i частота события $p^*(X < x_i)$ приближается (сходится по вероятности) к вероятности этого события. Следовательно, если X – непрерывная величина, то при увеличении n график функции $F^*(x)$ приближается к плавной кривой $F(x)$ – интегральной функции распределения величины X.

Таким образом, графическое изображение рядов распределения дает возможность наглядно представить эмпирическое распределение реализаций случайной величины и выразить закономерность ее распределения путем построения статистической интегральной функции распределения.

Пример 1.1. *Построить* гистограмму и статистическую функцию распределения часовой выработки подвижного состава автопредприятия.

Значения часовой выработки получены в ходе наблюдения за работой автомобилей-самосвалов КамАЗ-5511 в течение календарного года. Объем выборки составил $n = 100$ наблюдений. Размах вариации равен:

$$R = x_{max} - x_{min} = 15{,}13 - 4{,}0 = 11{,}13.$$

Величина интервала вариационного ряда определена по формуле (1.37)

$$\Delta x = \frac{x_{max} - x_{min}}{1 + 3{,}21 \cdot \lg n} = \frac{15{,}13 - 4{,}0}{1 + 3{,}21 \cdot \lg 100} = 1{,}5.$$

Количество интервалов вариационного ряда равно:

$$k = \frac{x_{max} - x_{min}}{\Delta x} = \frac{15{,}13 - 4{,}0}{1{,}5} = 7{,}42 \approx 8.$$

Вариационный ряд часовой выработки автомобиля представлен в табл. 1.4.

Таблица 1.4

Вариационный ряд часовой выработки автомобиля

Интервал Δx_i, m	4–5,5	5,5–7,0	7,0–8,5	8,5–10	10–11,5	11,5–13,0	13,0–14,5	14,5–16
Частота $\overset{*}{p_i}$	0,07	0,14	0,17	0,17	0,15	0,14	0,11	0,05

Решение

Для построения гистограммы определим ее ординаты из выражения:

$$\frac{m_i}{n \cdot \Delta x} = \frac{\overset{*}{p_i}}{\Delta x} = a_i.$$

Отсюда находим:

1) $\frac{\overset{*}{p_1}}{\Delta x} = \frac{0,07}{1,5} = 0,047;$

2) $\frac{\overset{*}{p_2}}{\Delta x} = \frac{0,14}{1,5} = 0,093;$

3) $\frac{\overset{*}{p_3}}{\Delta x} = \frac{0,17}{1,5} = 0,113;$

4) $\frac{\overset{*}{p_4}}{\Delta x} = \frac{0,17}{1,5} = 0,113;$

5) $\frac{\overset{*}{p_5}}{\Delta x} = \frac{0,15}{1,5} = 0,1;$

6) $\frac{\overset{*}{p_6}}{\Delta x} = \frac{0,14}{1,5} = 0,093;$

7) $\frac{\overset{*}{p_7}}{\Delta x} = \frac{0,11}{1,5} = 0,073;$

8) $\frac{\overset{*}{p_8}}{\Delta x} = \frac{0,05}{1,5} = 0,033.$

Основываясь на данных табл. 1.4 и проведенных расчетах, построим гистограмму (рис. 1.4).

Следует отметить, что при неограниченном увеличении объема выборки n кривая гистограммы частот совпадает с графиком плотности вероятностей.

Построим статистическую функцию распределения часовой выработки автомобиля:

1) при $x \le 4$ $F^*(x_1) = 0;$
2) при $4 < x \le 5,5$ $F^*(x_2) = 0,07;$
3) при $5,5 < x \le 7$ $F^*(x_3) = 0,21;$

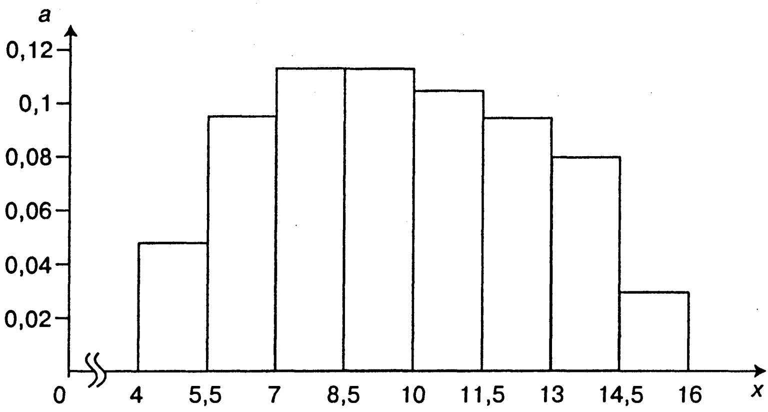


Рис. 1.4. Гистограмма часовой выработки автомобиля

4) при $7 < x \leq 8,5$	$F^*(x_4) = 0,38;$
5) при $8,5 < x \leq 10$	$F^*(x_5) = 0,55;$
6) при $10 < x \leq 11,5$	$F^*(x_6) = 0,70;$
7) при $11,5 < x \leq 13$	$F^*(x_7) = 0,84;$
8) при $13 < x \leq 14,5$	$F^*(x_8) = 0,95;$
9) при $14,5 < x \leq 16,0$	$F^*(x_9) = 1,0.$

График статистической функции распределения представлен на рис. 1.5.

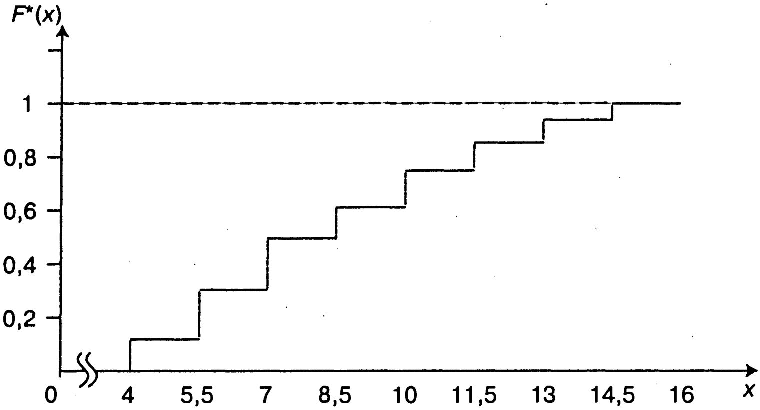


Рис. 1.5. Статистическая функция распределения часовой выработки автомобиля

Статистическая функция распределения случайной величины всегда есть разрывная ступенчатая функция, скачки которой происходят в точках, соответствующих возможным значениям случайной величины, и равны эмпирическим вероятностям этих значений. Сумма всех скачков функции $F^*(x)$ равна единице. По мере увеличения объема выборки и уменьшения интервалов Δx число скачков становится больше, а сами скачки – меньше; ступенчатая кривая становится более плавной; случайная величина постепенно приближается к непрерывной величине, а ее статистическая функция распределения – к непрерывной функции – интегральной функции распределения $F(x)$.

1.4. Основные законы распределения случайных величин

Полигон распределения и гистограмма есть реализация распределения выборочной совокупности при ограниченном числе наблюдений (N), а предельная кривая при $N \to \infty$ является распределением генеральной совокупности. Распределение генеральной совокупности является теоретическим распределением. Отдельные распределения изучены и поддаются точному аналитическому описанию. Приведем некоторые из них.

Дискретные законы распределения

А. Биномиальное распределение. Это распределение числа X появления события A в серии из n независимых испытаний. Вероятность наступления события A в каждом испытании равна p, а вероятность его отсутствия $q = 1 - p$. В каждом испытании возможны два исхода: наступление или ненаступление события A. При сформулированных условиях ряд распределения числа появления события A определяется формулой Бернулли

$$P(X = m) = C_n^m p^m (1 - p)^{n-m} (m = 0,1, \dots n), \tag{1.41}$$

или

$$P(X = m) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} (m = 0,1, \dots, n), \tag{1.42}$$

где $P(X = m)$ – вероятность появления события A равна m раз в серии из n испытаний.

Характер биномиального распределения определяется двумя параметрами p и n. На рис. 1.6 показаны многоугольники биномиального распределения для некоторых значений этих величин.

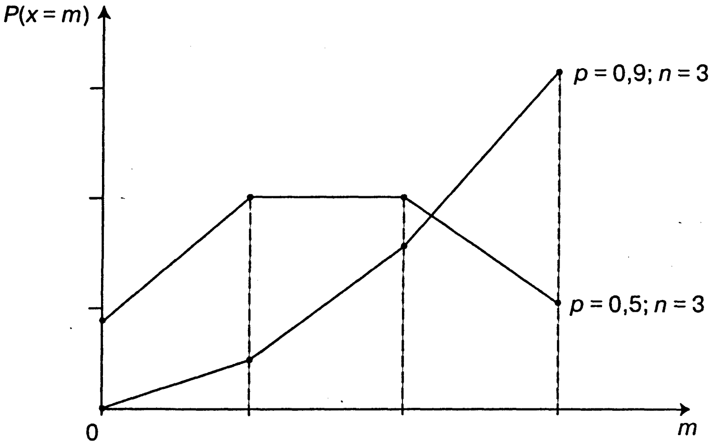


Рис. 1.6. Примеры кривых биномиального распределения

Определим числовые характеристики биномиального распределения случайной величины X:

математическое ожидание

$$M[X] = np; \tag{1.43}$$

дисперсию

$$D_x = npq = np(1 - p); \tag{1.44}$$

коэффициент асимметрии (скошенности) распределения

$$a_x = \frac{q - p}{\sqrt{npq}}; \tag{1.45}$$

коэффициент эксцесса (мера крутости) распределения

$$c_x = \frac{1 - 6pq}{npq}. \tag{1.46}$$

Из формул (1.45) и (1.46) следует, что при $p = q$ биномиальное распределение симметрично относительно математического ожидания, следовательно, эксцесс достигает наибольшее по модулю отрицательное значение. Если $c_x > 0$ имеется положительный эксцесс (вершина сильно вытянута); если $c_x < 0$ – отрицательный эксцесс (низковершинная кривая); если $c_x = 0$ – нормальное распределение.

Если $a_x > 0$ – асимметрия положительная;
если $a_x < 0$ – асимметрия отрицательная;
если $a_x = 0$ – распределение симметричное.

Пример 1.2. Техническая система состоит из пяти независимо друг от друга функционирующих узлов. *Определить* математическое ожидание, дисперсию, среднее квадратическое отклонение числа отказов узлов, если вероятность отказа любого из них $p = 0{,}2$.

Решение

1. Математическое ожидание числа отказов:

$$M[X] = np = 5 \cdot 0{,}2 = 1.$$

2. Дисперсия:

$$D_x = npq = 5 \cdot 0{,}2 \cdot 0{,}8 = 0{,}8.$$

3. Среднее квадратическое отклонение:

$$\sigma_x = \sqrt{D_x} = \sqrt{0{,}8} = 0{,}8944.$$

4. Коэффициент асимметрии:

$$a_x = \frac{q-p}{\sqrt{npq}} = \frac{0{,}8-0{,}2}{\sqrt{5 \cdot 0{,}2 \cdot 0{,}8}} = 0{,}671.$$

5. Коэффициент эксцесса:

$$c_x = \frac{1-6pq}{npq} = \frac{1-6 \cdot 0{,}2 \cdot 0{,}8}{5 \cdot 0{,}2 \cdot 0{,}8} = 0{,}05.$$

Б. Распределение Пуассона. Данное распределение является предельным случаем биномиального распределения. Предположим, что в биномиальном распределении $p \to 0$ и $n \to \infty$, так, что $n \cdot p \to$ $\to M[X] = a > 0$. Тогда плотность вероятности биномиального распределения принимает вид:

$$P(x = m) = \frac{(a^k \cdot e^{-a})}{k!} = \frac{a^k}{k!} \cdot e^{-a}, \quad k = 0, 1, 2, ..., \tag{1.47}$$

что и является распределением Пуассона. Формула (1.47) выражает ряд распределения Пуассона. Заметим, что распределение Пуассона зависит только от одного параметра – математического ожидания $M[X] = a$. Основные числовые характеристики случайной величины, имеющей распределение Пуассона, равны величине $a > 0$, а именно дисперсия случайной величины X, имеющей распределение Пуассона, численно равна ее математическому ожиданию. Этим свойством пользуются для оценки близости эмпирического распределения к распределению Пуассона.

На рис. 1.7 показаны кривые распределения Пуассона, отвечающие различным значениям математического ожидания:

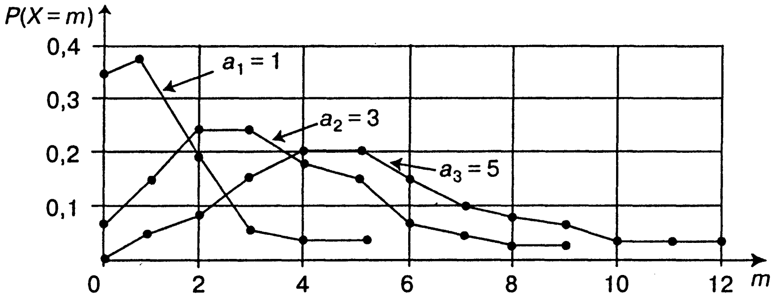


Рис. 1.7. Кривые распределения Пуассона

Из рис. 1.7 следует, что при увеличении математического ожидания a кривые распределения Пуассона становятся более симметричными. При $a \geq 10 \div 11$ несимметричность распределения практически не ощущается и закон Пуассона можно заменять нормальным законом распределения с определенными допущениями.

Пример 1.3. *Определить* вероятность того, что на АЗС находится один или хотя бы один автомобиль, если среднее число автомобилей, находящихся в данном интервале времени на АЗС, $a = 3$.

Решение

1. Вероятность нахождения одного автомобиля на АЗС следующая:

$$P(X=1)=\frac{a^m}{m!}\cdot e^{-a}=\frac{a^1}{1!}\cdot e^{-a}=a\cdot e^{-a}=3\cdot e^{-3}=0{,}149.$$

2. Вероятность того, что на АЗС находится хотя бы один автомобиль, равна вероятности того, что на АЗС находится не менее одного автомобиля, т. е.

$$P(X\geq 1)=1-P(X=0)=1-\frac{a^0}{0!}\cdot e^{-a}=1-e^{-3}=0{,}95.$$

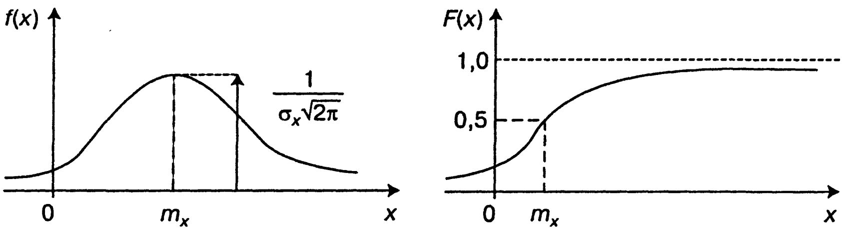
В. Нормальное распределение. Наиболее известным непрерывным распределением является нормальное. Плотность нормального распределения определяется по формуле

$$f(x)=\frac{1}{\sigma_x\sqrt{2\pi}}\cdot e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}. \qquad (1.48)$$

Непрерывная случайная величина X принимает значения от $-\infty$ до $+\infty$. Соответствующая функция распределения равна:

$$F(x)=\frac{1}{\sigma_x\sqrt{2\pi}}\cdot\int\limits_{-\infty}^{x} e^{-\frac{(x-m_x)}{2\sigma_x^2}}\,dx. \qquad (1.49)$$

Типичные графики плотности вероятности $f(x)$ и функции нормального распределения приведены на рис. 1.8.



а) плотность вероятности

б) функция нормального распределения

Рис. 1.8. Графики кривых нормального распределения

Кривой плотности вероятности $f(x)$ нормального распределения является плавная колоколообразная симметричная кривая, уравнение которой – формула (1.48).

Перечислим основные свойства нормального распределения.

1. Нормальное распределение полностью характеризуется математическим ожиданием и дисперсией.

2. Кривая плотности вероятности $f(x)$ нормального распределения симметрична относительно математического ожидания m_x. Максимум плотности распределения соответствует абсциссе, равной m_x.

3. При $|x| \to \infty$ ветви кривой распределения асимптотически приближаются к оси $0x$.

4. Математическое ожидание случайной величины X, распределенной в соответствии с нормальным законом, совпадает по величине с ее модой и медианой.

5. Коэффициенты асимметрии и эксцесса нормального распределения равны нулю.

Величина математического ожидания не влияет на форму кривой плотности распределения $f(x)$. С возрастанием σ_x максимальная ордината кривой $\frac{1}{\sigma_x\sqrt{2\pi}}$ постоянно убывает и нормальная кривая становится все более пологой. При уменьшении σ_x нормальная кривая становится все круче, т. е. растягивается вдоль оси ординат. При значении $\sigma_x = 1$ и $m_x = 0$ нормальную кривую называют ***нормированной***, а соответствующий закон распределения – ***стандартным нормальным законом распределения*** с плотностью

$$f(z)=\left(\frac{1}{\sqrt{2\pi}}\right)\cdot e^{-\frac{z^2}{2}}, \quad -\infty<z<\infty. \qquad (1.50)$$

Соответствующая функция распределения имеет вид:

$$\Phi(z)=\int_{-\infty}^{z}\left(\frac{1}{\sqrt{2\pi}}\right)\cdot e^{-\frac{z^2}{2}}\cdot dz. \qquad (1.51)$$

Путем подстановки $Z=\frac{(x-m_x)}{\sigma_x}$ нормальное распределение с произвольными параметрами m_x и σ_x приводится к стандартному виду. Вероятность попадания случайной величины в заданный интервал от α до β равна:

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha), \qquad (1.52)$$

где

$$F(\beta)=\frac{1}{\sigma_x\sqrt{2\pi}}\cdot\int_{-\infty}^{\beta}\exp\left[\frac{-(\beta-m_x)^2}{2\sigma_x^2}\right]dx. \qquad (1.53)$$

$$F(\alpha)=\frac{1}{\sigma_x\sqrt{2\pi}}\cdot\int_{-\infty}^{\alpha}\exp\left[\frac{-(\alpha-m_x)^2}{2\sigma_x^2}\right]dx. \qquad (1.54)$$

Интегралы (1.53) и (1.54) не выражаются через элементарные функции, поэтому для вычислений по формуле (1.52) обычно осуществляют замену $z_2 = \frac{(\beta - m_x)}{\sigma_x}$ и $z_1 = \frac{(\alpha - m_x)}{\sigma_x}$ и переходят к функции стандартного нормального закона распределения, которая имеет вид формулы (1.51). Тогда

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = \Phi(z_2) - \Phi(z_1).$$

Значения функции стандартного нормального закона распределения табулированы и приведены в приложении 6.

Отклонения случайной величины X от математического ожидания практически заключены в интервале $\pm 3\sigma_x$, при этом вероятность попадания X в данный интервал равна 0,9973.

Пример 1.4. Среднее время обслуживания персонального компьютера (ПК) $\bar{t} = 2$ ч. Среднее квадратическое отклонение времени обслуживания равно $\sigma_t = 0{,}403$ ч.

Определить вероятность окончания обслуживания ПК в течение интервала времени от 1,5 до 2,5 ч.

Решение

1. Вероятность попадания случайной величины t в интервал [1,5; 2,5] будет равна:

$$p(1{,}5 < t < 2{,}5) = F(2{,}5) - F(1{,}5).$$

2. Определим z:

$$z_2 = \frac{(t_2 - \bar{t})}{\sigma_t}; \qquad z_1 = \frac{(t_1 - \bar{t})}{\sigma_t};$$

$$z_2 = \frac{(2{,}5 - 2)}{0{,}403} = 1{,}24;$$

$$z_1 = \frac{(1{,}5 - 2)}{0{,}403} = -1{,}24.$$

3. По таблицам приложения 6 определим значение стандартной нормальной функции распределения:

$$\Phi(z_2) = \Phi(1{,}24) = 0{,}892;$$

$$\Phi(z_1) = \Phi(-1{,}24) = 0{,}107.$$

4. Вероятность окончания обслуживания ПК в течение интервала времени [1,5; 2,5] будет равна:

$$p(1{,}5 < t < 2{,}5) = \Phi(z_2) - \Phi(z_1) = 0{,}892 - 0{,}107 = 0{,}785.$$

Г. Гамма-распределение и распределение Эрланга. Неотрицательная случайная величина X имеет гамма-распределение, если ее плотность распределения вычисляется по формуле

$$f_k(x) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{\Gamma(k)} \text{ при } x > 0, \tag{1.55}$$

где $\lambda > 0$ и $k > 0$;

$\Gamma(k)$ – гамма-функция равна:

$$\Gamma(k) = \int_0^\infty e^{-t} \cdot t^{k-1} \cdot dt, \tag{1.56}$$

если k – целое неотрицательное число, то

$$\Gamma(k) = k! \tag{1.57}$$

Математическое ожидание случайной величины X, подчиненной гамме-распределению, равно:

$$m_x = \frac{k}{\lambda}. \tag{1.58}$$

При этом дисперсия величины X определяется по формуле

$$D_x = \frac{k}{\lambda^2}. \tag{1.59}$$

При целом $k > 1$ гамма-распределение превращается в распределение Эрланга k-го порядка, т. е.

$$f_k(x) = \frac{\lambda \cdot (\lambda x)^{k-1} \cdot e^{-\lambda x}}{(k-1)!} \quad (x > 0;\ k = 1, 2, \ldots). \tag{1.60}$$

Закону Эрланга k-го порядка подчинена сумма независимых случайных величин $x_1 + x_2 + \ldots + x_k$, каждая из которых распределена по показательному закону с параметром λ.

При $k = 1$ гамма-распределение превращается в показательное распределение с параметром λ.

Д. Показательное распределение. Непрерывная случайная величина X имеет показательное распределение, если ее плотность распределения выражается формулой

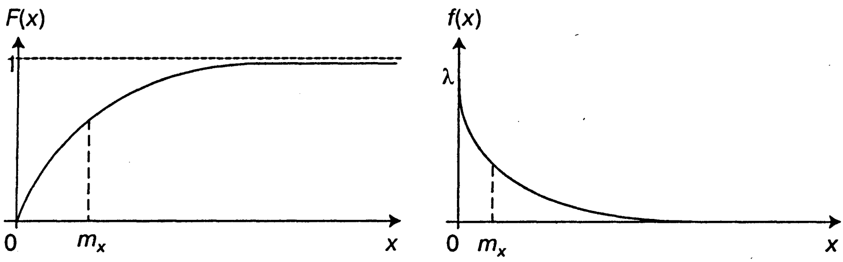
$$f(x) = \lambda \cdot e^{-\lambda x}, \quad x > 0. \qquad (1.61)$$

Положительная величина λ является параметром показательного распределения.

Функция распределения случайной величины X выглядит следующим образом:

$$F(x) = 1 - e^{-\lambda x}, \lambda > 0, 0 \le x < \infty. \qquad (1.62)$$

Графики функции и плотности показательного распределения приведены на рис. 1.9.



а) функция показательного распределения

б) плотность показательного распределения

Рис. 1.9. Графики показательного распределения

Математическое ожидание случайной величины X, имеющей показательное распределение, обратно его параметру, т. е.

$$m_x = \frac{1}{\lambda}. \qquad (1.63)$$

Дисперсия случайной величины X, имеющей показательное распределение, равна:

$$D_x = \frac{1}{\lambda^2}. \qquad (1.64)$$

Отсюда

$$\sigma_x = \frac{1}{\lambda}, \text{ т.е. } \sigma_x = m_x = \frac{1}{\lambda}. \qquad (1.65)$$

Коэффициент вариации случайной величины X, имеющей показательное распределение, равен единице:

$$V_x = \frac{\sigma_x}{m_x} = 1.$$

Существует важное соотношение между пуассоновским и экспоненциальным распределениями. Если случайная величина подчинена закону Пуассона и представляет собой число отказов в единицу времени, то случайная величина, которая определяет промежуток времени между двумя последовательными отказами, распределена по экспоненциальному закону. Экспоненциальное распределение можно в сущности вывести из распределения Пуассона.

Е. Равномерное распределение. Непрерывная случайная величина X имеет равномерное распределение на отрезке $[a, b]$, если на этом отрезке плотность распределения постоянна, а вне его – равна нулю:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{при } a < x < b; \\ 0 & x \ge b;\ x \le a. \end{cases} \tag{1.66}$$

Кривая равномерного распределения показана на рис. 1.10.

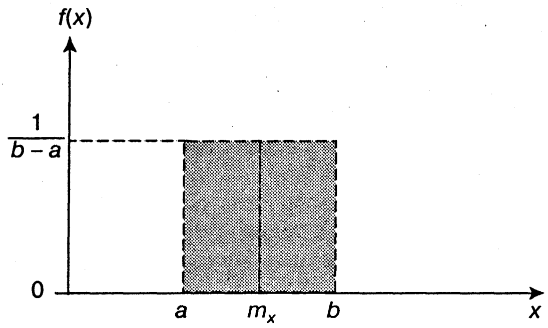


Рис. 1.10. Кривая равномерного распределения

Значения $f(x)$ в крайних точках a и b участка (a, b) не указываются, так как вероятность попадания в любую из этих точек для непрерывной случайной величины X равна нулю.

Кривая равномерного распределения (рис. 1.10) имеет вид прямоугольника, опирающегося на участок $[a, b]$.

Математическое ожидание случайной величины X, имеющей равномерное распределение на участке $[a, b]$, равно:

$$m_x = \frac{a+b}{2}. \quad (1.67)$$

Дисперсия случайной величины X, имеющей равномерное распределение на участке $[a, b]$, вычисляется по формуле

$$D_x = \frac{(b-a)^2}{12}. \quad (1.68)$$

Отсюда

$$\sigma_x = \sqrt{D_x} = \frac{b-a}{2\sqrt{3}}. \quad (1.69)$$

Вероятность попадания равномерно распределенной случайной величины X на участок $[\alpha, \beta]$ выразим формулой

$$P(\alpha < X < \beta) = \frac{\beta-\alpha}{b-a}. \quad (1.70)$$

Пример 1.5. Троллейбусы прибывают на остановку через 4 мин. Какова вероятность того, что время ожидания троллейбуса не превысит 3 мин?

Решение

Так как $(\beta - \alpha) = 3$ мин, а $(b - a) = 4$ мин, то

$$P(0 < X < 3) = \frac{3}{4} = 0{,}75.$$

1.5. Выбор теоретического закона распределения случайной величины

В любом статистическом распределении присутствуют элементы случайности, и, как следствие, экспериментальные точки гистограммы обычно колеблются от опыта к опыту около неизвестной кривой истинного распределения.

При наличии числовых характеристик случайной величины (математического ожидания, дисперсии, коэффициента вариации) законы ее распределения могут быть определены в первом приближении по табл. 1.5.

Таблица 1.5

Законы распределения случайной положительной величины в зависимости от коэффициента вариации

Пределы изменения коэффициента вариации V_x	Закон распределения случайной величины X
$V_x \leq 0{,}3$	Нормальный
$0{,}3 < V_x < 0{,}4$	Гамма-распределение
$0{,}4 \leq V_x < 1$	Вейбулла
$V_x = 1$	Экспоненциальный, Пуассона

Для более точного определения теоретического закона распределения проводят дополнительную статистическую обработку данных. При обработке статистических данных решают вопрос о том, как подобрать для исходного статистического ряда теоретическую кривую распределения, которая выражала бы лишь существенные черты статистического материала, но не случайности, обусловленные недостаточным объемом выборки экспериментальных данных. Под *построением теоретической кривой распределения* понимается такая обработка статистических данных, когда обеспечивается подбор наиболее подходящего теоретического закона распределения, задаваемого либо функцией распределения $F(x)$, либо плотностью распределения $f(x)$.

Для построения теоретической кривой распределения исходный статистический ряд распределения аппроксимируется одной из дифференциальных функций теоретического распределения $f(x)$. При этом выбирается такая функция $f(x)$, которая обеспечивала бы максимальное приближение теоретических данных к эмпирическим $f(x) \approx f^*(x)$. Для оценки правдоподобия этого приближенного равенства разработано несколько критериев согласия проверяемых гипотез относительно вида функции $f(x)$.

Наиболее употребительными критериями согласия являются *критерий χ^2 Пирсона* и *критерий Колмогорова*. Для примера подробно рассмотрим критерий χ^2 Пирсона.

Критерий χ^2 Пирсона. Согласно критерию χ^2 Пирсона в качестве меры расхождения между теоретическим законом распределения и статистическим распределением выбрана величина, определяемая следующим выражением:

$$\chi^2 = n \cdot \sum_{i=1}^{k} \frac{(p_i^* - p_i)^2}{p_i}, \tag{1.71}$$

где k – число интервалов статистического ряда;
p_i^* – статистическая вероятность попадания случайной величины в интервал;
p_i – теоретическая вероятность попадания случайной величины X в i-й интервал.

Учитывая соотношение $p_i^* = \dfrac{m_x}{n}$, выражение (1.71) после преобразований записываем в виде:

$$\chi^2 = \sum_{i=1}^{k} \frac{(m_i - n \cdot p_i)^2}{n \cdot p_i}, \tag{1.72}$$

где m_i – эмпирическое количество значений случайной величины, попадающих в i-й интервал.

Для того чтобы выяснить, является ли полученное расхождение χ^2 случайным за счет ограниченного объема выборки или свидетельствует о наличии существенной разницы между теоретическим и статистическим распределениями, необходимо вычислить вероятность такого расхождения $\Delta \geq \chi^2$, т. е. $P(\chi^2 \leq \Delta < \infty)$ вероятность того, что за счет чисто случайных причин мера расхождения теоретического и статистического распределений Δ будет не меньше, чем фактическое значение χ^2 для данной выборки. Величина вероятности расхождения определяется по специальным таблицам при известных значениях r и χ^2.

Число степеней свободы r вычисляется для данного статистического ряда распределения как

$$r = k - l, \tag{1.73}$$

где l – число исчисленных статистических характеристик (средняя, дисперсия и т. д.), использованных при вычислении теоретического распределения.

Если искомая вероятность окажется очень малой, практически меньше 0,1, то выбранное теоретическое распределение следует считать неудачным. При относительно большом значении искомой вероятности теоретическое распределение можно признать не противоречащим опытным данным.

Следует отметить, что критерий χ^2 Пирсона применим в тех случаях, когда объем выборки $n \geq 100$ и в каждом интервале число наблюдений не менее $m_i \geq 5$.

Пример 1.6. Пользуясь критерием χ^2 Пирсона, подобрать теоретический закон распределения для часовой выработки автомобилей КамАЗ-5511, статистическое распределение которой приведено в табл. 1.4 примера 1.1.

Решение

По форме гистограммы рис. 1.4 можно предположить, что часовая выработка автомобиля подчиняется нормальному закону.

Для оценки числовых характеристик нормального распределения вычислим:

математическое ожидание

$$m_x = \sum_{i=1}^{k} \overline{x}_i p_i^* = 4{,}75 \cdot 0{,}07 + 6{,}25 \cdot 0{,}14 + 7{,}75 \cdot 0{,}17 + 9{,}25 \cdot 0{,}17 +$$

$$+ 10{,}75 \cdot 0{,}15 + 12{,}25 \cdot 0{,}14 + 13{,}75 \cdot 0{,}11 + 15{,}25 \cdot 0{,}05 = 9{,}7 \text{ т};$$

дисперсию

$$D_x = \sum_{i=1}^{k} (m_x - \overline{x}_i) p_i^* = (9{,}7 - 4{,}75) \cdot 0{,}07 + (9{,}7 - 6{,}25) \cdot 0{,}14 +$$

$$+ (9{,}7 - 7{,}75) \cdot 0{,}17 + (9{,}7 - 9{,}25) \cdot 0{,}17 + (9{,}7 - 10{,}75) \cdot 0{,}15 +$$

$$+ (9{,}7 - 12{,}25) \cdot 0{,}14 + (9{,}7 - 13{,}75) \cdot 0{,}11 + (9{,}7 - 15{,}25) \cdot 0{,}05 \approx 8{,}48,$$

где $\overline{x}_i$ – значение середины i-го интервала.

Среднее квадратическое отклонение

$$\sigma_x = \sqrt{D_x} = \sqrt{8{,}483} = 2{,}91.$$

Коэффициент вариации

$$V_x = \frac{\sigma_x}{m_x} = \frac{2{,}91}{9{,}7} \approx 0{,}3.$$

Величина $V_x \approx 0{,}3$ свидетельствует о том, что теоретическое распределение близко к нормальному закону распределения. Проверим данную гипотезу, воспользовавшись критерием согласия χ^2.

Определим теоретическую вероятность попадания значений часовой выработки автомобиля в заданные интервалы, используя формулу (1.52)

$$p_i = F(x_{i+1}) - F(x_i) = \Phi\left(\frac{x_{i+1} - m_x}{\sigma_x}\right) - \Phi\left(\frac{x_i - m_x}{\sigma_x}\right),$$

где x_i, x_{i+1} – границы i-го интервала (табл. 1.4).

Затем составим сравнительную таблицу (табл. 1.6) чисел попаданий в интервалы m_i и соответствующих значений np_i (n = 100).

Таблица 1.6

Сравнительная таблица

Интервал, Δx_i, т	4–5,5	5,5–7,0	7,0–8,5	8,5–10	10–11,5	11,5–13,0	13,0–14,5	14,5–16
Количество наблюдений, m_i	7	14	17	17	15	14	11	5
Теоретическое количество наблюдений, np_i	5	11	17	21	20	14	8	4

Построим график теоретического распределения и совместим его с гистограммой статистического распределения (рис. 1.11).

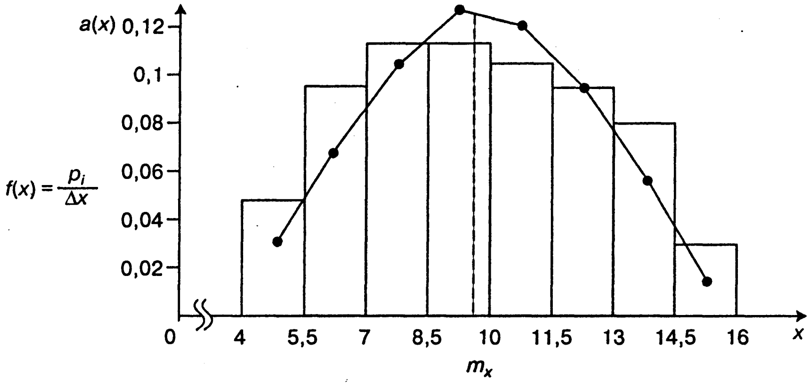


Рис. 1.11. Распределение часовой выработки автомобиля

Вычислим значение меры расхождения по формуле

$$\chi^2 = \sum_{i=1}^{8} \frac{(m_i - np_i)^2}{np_i} = 5{,}01.$$

Определим число степеней свободы

$$r = k - l = 8 - 2 = 6.$$

По приложению 8 для $r = 6$ находим следующее:

$$\text{при } \chi^2 = 3{,}83 \qquad p = 0{,}7;$$
$$\text{при } \chi^2 = 5{,}35 \qquad p = 0{,}5.$$

Следовательно, искомая вероятность p при $\chi^2 = 5{,}01$ приближенно равна $p \approx 0{,}545$. Эта вероятность малой не является; поэтому гипотезу о том, что часовая выработка автомобиля распределена по нормальному закону, можно считать правдоподобной.

Задачи

1.1. Из двадцати сотрудников малого предприятия пять опоздали к началу рабочего дня.

Определите частоту опозданий сотрудников.

1.2. Автомобилист совершает две попытки с целью преодоления дорожного препятствия. Вероятность преодоления препятствия при каждой попытке одинакова и равна 0,8.

Найдите вероятность того, что в результате двух попыток препятствие будет преодолено хотя бы один раз.

1.3. *Определите* математическое ожидание и моду числа остановок автобуса перед светофорами на маршруте, если случайная величина X – число остановок – задана следующей таблицей распределения:

x_i	0	1	2	3	4	5
$p(x_i)$	0,05	0,05	0,2	0,5	0,1	0,1

1.4. *Определите* среднее квадратическое отклонение числа отказов оборудования, если случайная величина X – число отказов оборудования – задана следующей таблицей распределения:

x_i	0	1	2	3	4	5	6
$p(x_i)$	0,3	0,1	0,05	0,1	0,2	0,2	0,05

1.5. В задаче **1.4** *определите* коэффициент вариации случайной величины X.

1.6. *Постройте* гистограмму часовой производительности одного рабочего в течение календарного периода. Объем выборки составил 200 наблюдений. Вариационный ряд производительности рабочего представлен в следующей таблице:

Интервал, Δx_i (ед.)	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11
Частота, p_i^*	0,03	0,10	0,15	0,19	0,24	0,12	0,11	0,06

1.7. По данным задачи 1.6 *постройте* статистическую функцию распределения часовой производительности рабочего.

1.8. Малое предприятие имеет 16 автомобилей, работающих независимо друг от друга.

Определите математическое ожидание, дисперсию, среднее квадратическое отклонение числа отказов автомобилей, если вероятность отказа любого из них равна $p = 0,3$.

1.9. Число проверок предприятия в течение года инспекцией является случайной величиной, имеющей распределение Пуассона.

Определите вероятность того, что на предприятии будет произведена в течение календарного года одна или хотя бы одна проверка, если среднее число проверок на данном временном интервале $a = 4$.

1.10. На предприятии работает 50 станков. Вероятность отказа каждого из них – 0,002. Число отказов станков – случайная величина, имеющая распределение Пуассона.

Определите вероятность безотказного функционирования всех элементов.

1.11. Поезда метрополитена следуют через 1,5 мин. *Какова вероятность* того, что время ожидания поезда не превысит 1 мин?

1.12. Средняя часовая выручка магазина $B = 100$ д. е. Среднее квадратическое отклонение часовой выручки $\sigma_B = 25$ д. е. Часовая выручка есть случайная величина, подчиненная нормальному закону распределения.

Определите вероятность получения в течение одного часа выручки в размере от 80 до 120 д. е.

1.13. Автобусы прибывают на остановку через 6 мин. *Какова вероятность* того, что время ожидания автобуса не превысит 5 мин?

1.14. Объем продаж товара в течение месяца есть случайная величина, подчиненная нормальному закону распределения с параметрами $\bar{x} = 500$ и $\sigma_x = 120$ д. е.

Определите вероятность продажи товара в течение одного месяца на сумму от 480 до 600 д. е.

1.15. На предприятии работает 50 специалистов, вероятность невыхода специалиста на работу по причине болезни равна 0,001. Число заболевших специалистов – случайная величина, имеющая распределение Пуассона.

Определите вероятность выхода на работу всех специалистов.

1.16. *Постройте* гистограмму часовой торговой выручки (X) магазина в течение календарного периода. Объем выборки составил

150 наблюдений. Вариационный ряд торговой выручки представлен в следующей таблице (д. е.):

Интервал, (ед.) Δx_i	3–4	4–5	5–6	6–7	7–8	8–9	9–10	10–11
Частота, p_i^*	0,03	0,10	0,15	0,19	0,24	0,12	0,11	0,06

1.17. По данным задачи 1.16 *постройте* статистическую функцию распределения часовой торговой выручки.

1.18. Пользуясь критерием χ^2 Пирсона, *подберите* теоретический закон распределения для часовой производительности рабочего, статистическое распределение которой приведено в задаче 1.6.

1.19. Пользуясь критерием χ^2 Пирсона, *подберите* теоретический закон распределения для часовой торговой выручки, статистическое распределение которой приведено в задаче 1.16.

1.20. Предприятие имеет 5 станков по производству камня, работающих независимо друг от друга. Вероятность отказа любого из них $p = 0{,}25$.

Определите параметры закона биномиального распределения случайной величины – число отказов станков.

1.21. *Определите* среднее квадратическое отклонение и дисперсию числа отказов автомобилей, если случайная величина X – число отказов автомобилей – задана следующей таблицей распределения:

x_i	0	1	2	3	4	5	6	7
$p(x_i)$	0,2	0,15	0,15	0,1	0,25	0,04	0,06	0,05

Глава 2
Моделирование экономических систем с использованием марковских случайных процессов

2.1. Основные понятия марковских процессов

Функция $X(t)$ называется *случайной*, если ее значение при любом аргументе t является случайной величиной.

Случайная функция $X(t)$, аргументом которой является время, называется *случайным процессом.*

Марковские процессы являются частным видом случайных процессов. Особое место марковских процессов среди других классов случайных процессов обусловлено следующими обстоятельствами: для марковских процессов хорошо разработан математический аппарат, позволяющий решать многие практические задачи; с помощью марковских процессов можно описать (точно или приближенно) поведение достаточно сложных систем.

> ***Определение.*** Случайный процесс, протекающий в какой-либо системе S, называется марковским (или процессом без последействия), если он обладает следующим свойством: для любого момента времени t_0 вероятность любого состояния системы в будущем (при $t > t_0$) зависит только от ее состояния в настоящем (при $t = t_0$) и не зависит от того, когда и каким образом система S пришла в это состояние.

Классификация марковских процессов. Классификация марковских случайных процессов производится в зависимости от непрерывности или дискретности множества значений функции $X(t)$ и параметра t.

Различают следующие основные виды марковских случайных процессов:

- с дискретными состояниями и дискретным временем (цепь Маркова);
- с непрерывными состояниями и дискретным временем (марковские последовательности);
- с дискретными состояниями и непрерывным временем (непрерывная цепь Маркова);
- с непрерывным состоянием и непрерывным временем.

В данной работе будут рассматриваться только марковские процессы с дискретными состояниями $S_1, S_2, ..., S_n$.

Граф состояний. Марковские процессы с дискретными состояниями удобно иллюстрировать с помощью так называемого графа состояний (рис. 2.1), где кружками обозначены состояния $S_1, S_2, ...$ системы S, а стрелками – возможные переходы из состояния в состояние. На графе отмечаются только непосредственные переходы, а не переходы через другие состояния. Возможные задержки в прежнем состоянии изображают «петлей», т. е. стрелкой, направленной из данного состояния в него же. Число состояний системы может быть как конечным, так и бесконечным (но счетным). Пример графа состояний системы S представлен на рис.2.1.

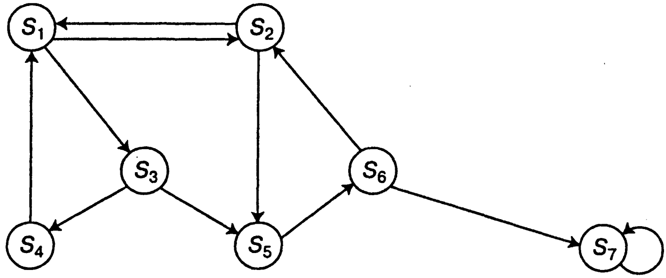


Рис. 2.1. Граф состояний системы S

2.2. Марковские цепи

Марковский случайный процесс с дискретными состояниями и дискретным временем называют *марковской цепью*. Для такого процесса моменты t_1, t_2, ..., когда система S может менять свое состояние, рассматривают как последовательные шаги процесса, а в качестве аргумента, от которого зависит процесс, выступает не время t, а номер шага 1, 2, ..., k, ... Случайный процесс в этом случае характеризуется последовательностью состояний $S(0)$, $S(1)$, $S(2)$, ..., $S(k)$, ..., где $S(0)$ – начальное состояние системы (перед первым шагом); $S(1)$ – состояние системы после первого шага; $S(k)$ – состояние системы после k-го шага...

Событие $\{S(k) = Si\}$, состоящее в том, что сразу после k-го шага система находится в состоянии $S_i(i = 1, 2, ...)$, является случайным событием. Последовательность состояний $S(0)$, $S(1)$, ..., $S(k)$, ... можно рассматривать как последовательность случайных событий. Такая случайная последовательность событий называется *марковской цепью*, если для каждого шага вероятность перехода из любого состояния S_i в любое S_j не зависит от того, когда и как система пришла в состояние S_i. Начальное состояние $S(0)$ может быть заданным заранее или случайным.

Вероятностями состояний цепи Маркова называются вероятности $P_i(k)$ того, что после k-го шага (и до $(k + 1)$-го) система S будет находиться в состоянии $S_i(i = 1, 2, ..., n)$. Очевидно, для любого k

$$\sum_{i=1}^{n} P_i(k) = 1. \tag{2.1}$$

Начальным распределением вероятностей марковской цепи называется распределение вероятностей состояний в начале процесса:

$$P_1(0),\ P_2(0),\ ...,\ P_i(0),\ ...,\ P_n(0). \tag{2.2}$$

В частном случае, если начальное состояние системы S в точности известно $S(0) = S_i$, то начальная вероятность $P_i(0) = 1$, а все остальные равны нулю.

Вероятностью перехода (переходной вероятностью) на k-м шаге из состояния S_i в состояние S_j называется условная вероятность того, что система S после k-го шага окажется в состоянии S_j при условии, что непосредственно перед этим (после $k - 1$ шага) она находилась в состоянии S_i.

Поскольку система может пребывать в одном из n состояний, то для каждого момента времени t необходимо задать n^2 вероятностей перехода P_{ij}, которые удобно представить в виде следующей матрицы:

$$\|P_{ij}\| = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \dots & \dots & \dots & \dots \\ P_{i1} & P_{i2} & \dots & P_{in} \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix}, \tag{2.3}$$

где P_{ij} – вероятность перехода за один шаг из состояния S_i в состояние S_j;
P_{ii} – вероятность задержки системы в состоянии S_i.

Матрица (2.3) называется *переходной* или *матрицей переходных вероятностей*.

Если переходные вероятности не зависят от номера шага (от времени), а зависят только от того, из какого состояния в какое осуществляется переход, то соответствующая цепь Маркова называется *однородной*.

Переходные вероятности однородной марковской цепи P_{ij} *образуют квадратную матрицу размера $n \times n$*. Отметим некоторые ее особенности:

1. Каждая строка характеризует выбранное состояние системы, а ее элементы представляют собой вероятности всех возможных переходов за один шаг из выбранного (из i-го) состояния, в том числе и переход в самое себя.

2. Элементы столбцов показывают вероятности всех возможных переходов системы за один шаг в заданное (j-е) состояние (иначе говоря, строка характеризует вероятность перехода системы из состояния, столбец – в состояние).

3. Сумма вероятностей каждой строки равна единице, так как переходы образуют полную группу несовместных событий:

$$\sum_{j=1}^{n} P_{ij} = 1, \quad i = \overline{1,n}. \qquad (2.4)$$

4. По главной диагонали матрицы переходных вероятностей стоят вероятности P_{ii} того, что система не выйдет из состояния S_i, а останется в нем.

Если для однородной марковской цепи заданы начальное распределение вероятностей (2.2) и матрица переходных вероятностей $\|P_{ij}\|$ (2.3), то вероятности состояний системы $P_i(k)$ ($i = \overline{1,n}$; $j = \overline{1,n}$) определяются по рекуррентной формуле:

$$P_i(k) = \sum_{j=1}^{n} P_j(k-1) \cdot P_{ji}, \quad (i = \overline{1,n};\ j = \overline{1,n}). \qquad (2.5)$$

Пример 2.1. Рассмотрим процесс функционирования системы автомобиля. Пусть автомобиль (система) в течение одной смены (суток) может находиться в одном из двух состояний: исправном (S_1) и неисправном (S_2). Граф состояний системы представлен на рис. 2.2.

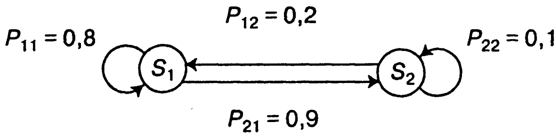


Рис. 2.2. Граф состояний автомобиля

В результате проведения массовых наблюдений за работой автомобиля составлена следующая матрица вероятностей перехода:

$$\|P_{ij}\| = \begin{bmatrix} 0,8 & 0,2 \\ 0,9 & 0,1 \end{bmatrix}, \qquad (2.6)$$

где $P_{11} = 0,8$ – вероятность того, что автомобиль останется в исправном состоянии;

$P_{12} = 0,2$ – вероятность перехода автомобиля из состояния «исправен» в состояние «неисправен»;

$P_{21} = 0,9$ – вероятность перехода автомобиля из состояния «неисправен» в состояние «исправен»;

$P_{22} = 0,1$ – вероятность того, что автомобиль останется в состоянии «неисправен».

Вектор начальных вероятностей состояний автомобиля задан $P(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, т.е. $P_1(0) = 0$ и $P_2(0) = 1$.

Требуется *определить* вероятности состояний автомобиля через трое суток.

Используя матрицу переходных вероятностей, определим вероятности состояний $P_i(k)$ после первого шага (после первых суток):

$$P_1(1) = P_1(0)P_{11} + P_2(0)P_{21} = 0 \cdot 0{,}8 + 1 \cdot 0{,}9 = 0{,}9;$$
$$P_2(1) = P_1(0)P_{12} + P_2(0)P_{22} = 0 \cdot 0{,}2 + 1 \cdot 0{,}1 = 0{,}1.$$

Вероятности состояний после второго шага (после вторых суток) таковы:

$$P_1(2) = P_1(1)P_{11} + P_2(1)P_{21} = 0{,}9 \cdot 0{,}8 + 0{,}1 \cdot 0{,}9 = 0{,}81;$$
$$P_2(2) = P_1(1)P_{12} + P_2(1)P_{22} = 0{,}9 \cdot 0{,}2 + 0{,}1 \cdot 0{,}1 = 0{,}19.$$

Вероятности состояний после третьего шага (после третьих суток) равны

$$P_1(3) = P_1(2)P_{11} + P_2(2)P_{21} = 0{,}81 \cdot 0{,}8 + 0{,}19 \cdot 0{,}9 = 0{,}819;$$
$$P_2(3) = P_1(2)P_{12} + P_2(2)P_{22} = 0{,}81 \cdot 0{,}2 + 0{,}19 \cdot 0{,}1 = 0{,}181.$$

Таким образом, после третьих суток автомобиль будет находиться в исправном состоянии с вероятностью 0,819 и в состоянии «неисправен» с вероятностью 0,181.

Пример 2.2. В процессе эксплуатации ЭВМ может рассматриваться как физическая система S, которая в результате проверки может оказаться в одном из следующих состояний:

S_1 – ЭВМ полностью исправна;

S_2 – ЭВМ имеет неисправности в оперативной памяти, при которых она может решать задачи;

S_3 – ЭВМ имеет существенные неисправности и может решать ограниченный класс задач;

S_4 – ЭВМ полностью вышла из строя.

В начальный момент времени ЭВМ полностью исправна (состояние S_1). Проверка ЭВМ производится в фиксированные моменты времени t_1, t_2, t_3. Процесс, протекающий в системе S, может рассматриваться как однородная марковская цепь с тремя шагами (первая, вторая, третья проверки ЭВМ). Матрица переходных вероятностей имеет вид:

$$\|P_{ij}\| = \begin{Vmatrix} 0{,}3 & 0{,}4 & 0{,}1 & 0{,}2 \\ 0 & 0{,}2 & 0{,}5 & 0{,}3 \\ 0 & 0 & 0{,}4 & 0{,}6 \\ 0 & 0 & 0 & 1{,}0 \end{Vmatrix}.$$

Определите вероятности состояний ЭВМ после трех проверок.

Решение

Граф состояний имеет вид, показанный на рис. 2.3. Против каждой стрелки проставлена соответствующая вероятность перехода. Начальные вероятности состояний $P_1(0) = 1$; $P_2(0) = P_3(0) = P_4(0) = 0$.

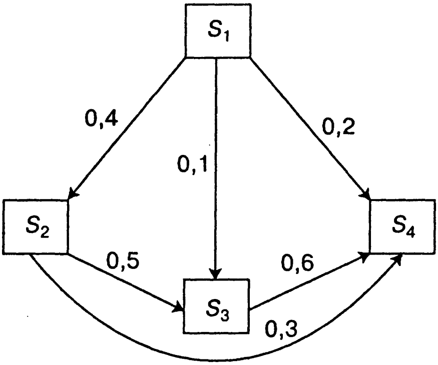


Рис. 2.3. Граф состояний ЭВМ

По формуле (2.5), учитывая в сумме вероятностей только те состояния, из которых возможен непосредственный переход в данное состояние, находим:

$P_1(1) = P_1(0)P_{11} = 1 \cdot 0{,}3 = 0{,}3;$
$P_2(1) = P_1(0)P_{12} = 1 \cdot 0{,}4 = 0{,}4;$
$P_3(1) = P_1(0)P_{13} = 1 \cdot 0{,}1 = 0{,}1;$
$P_4(1) = P_1(0)P_{14} = 1 \cdot 0{,}2 = 0{,}2;$

$P_1(2) = P_1(1)P_{11} = 0{,}3 \cdot 0{,}3 = 0{,}09;$
$P_2(2) = P_1(1)P_{12} + P_2(1)P_{22} = 0{,}3 \cdot 0{,}4 + 0{,}4 \cdot 0{,}2 = 0{,}20;$
$P_3(2) = P_1(1)P_{13} + P_2(1)P_{23} + P_3(1)P_{33} = 0{,}27;$
$P_4(2) = P_1(1)P_{14} + P_2(1)P_{24} + P_3(1)P_{34} + P_4(1)P_{44} = 0{,}44;$

$P_1(3) = P_1(2)P_{11} = 0{,}09 \cdot 0{,}3 = 0{,}027;$
$P_2(3) = P_1(2)P_{12} + P_2(2)P_{22} = 0{,}09 \cdot 0{,}4 + 0{,}20 \cdot 0{,}2 = 0{,}076;$
$P_3(3) = P_1(2)P_{13} + P_2(2)P_{23} + P_3(2)P_{33} = 0{,}217;$
$P_4(3) = P_1(2)P_{14} + P_2(2)P_{24} + P_3(2)P_{34} + P_4(2)P_{44} = 0{,}680.$

Итак, вероятности состояний ЭВМ после трех проверок следующие: $P_1(3) = 0{,}027$; $P_2(3) = 0{,}076$; $P_3(3) = 0{,}217$; $P_{4(3)} = 0{,}680$.

2.3. Непрерывные цепи Маркова

Марковский случайный процесс с дискретными состояниями и непрерывным временем называется *непрерывной цепью Маркова* при условии, что переход системы из состояния в состояние происходит не в фиксированные, а в случайные моменты времени.

В экономике часто встречаются ситуации, которые указать заранее невозможно. Например, любая деталь или агрегат автомобиля могут выйти из строя в любой, непредсказуемый заранее момент времени. Для описания таких систем в отдельных случаях можно использовать математический аппарат непрерывной цепи Маркова.

Пусть система характеризуется n состояниями $S_0, S_1, S_2, ..., S_n$, а переход из состояния в состояние может осуществляться в любой момент времени. Обозначим через $P_i(t)$ вероятность того, что в момент времени t система S будет находиться в состоянии S_i ($i = 0,1, ..., n$). Требуется определить для любого t вероятности состояний $P_0(t), P_1(t), ..., P_n(t)$. Очевидно, что

$$\sum_{i=0}^{n} P_i(t) = 1.$$

Для процесса с непрерывным временем вместо переходных вероятностей P_{ij} рассматриваются плотности вероятностей перехода λ_{ij}, представляющие собой предел отношения вероятности перехода системы за время Δt из состояния S_i в состояние S_j к длине промежутка Δt:

$$\lambda_{ij}(t) = \lim_{\Delta t \to 0} \frac{P_{ij}(t; \Delta t)}{\Delta t}, \tag{2.7}$$

где $P_{ij}(t; \Delta t)$ – вероятность того, что система, пребывавшая в момент t в состоянии S_i, за время Δt перейдет из него в состояние S_j (при этом всегда $i \neq j$).

Если λ_{ij} = const, то процесс называется *однородным*, если плотность вероятности зависит от времени $\lambda_{ij} = \lambda_{ij}(t)$, то – *неоднородным*.

При рассмотрении непрерывных марковских процессов принято представлять переходы системы S из состояния в состояние как происходящие под влиянием некоторых потоков событий[1]. *Потоком событий* называется последовательность однородных событий,

[1] *Вентцель Е. С., Овчаров Л. А.* Прикладные задачи теории вероятностей. – М.: Радио и связь, 1983.

следующих одно за другим через какие-то, вообще говоря, случайные интервалы времени. Плотность вероятности перехода интерпретируется как интенсивность λ_{ij} соответствующих потоков событий. Если все эти потоки пуассоновские, то процесс, протекающий в системе S, будет марковским.

При изучении марковских случайных процессов с дискретными состояниями и непрерывным временем в графе состояний над стрелками, ведущими из состояния S_i в S_j, проставляют соответствующие интенсивности λ_{ij}. Такой граф состояний называют *размеченным.*

Пусть система S имеет конечное число состояний $S_0, S_1, ..., S_n$. Случайный процесс, протекающий в этой системе, описывается вероятностями состояний $P_0(t), P_1(t), ... P_n(t)$, где $P_i(t)$ – вероятность того, что система S в момент t находится в состоянии S_i. Для любого t

$$\sum_{i=0}^{n} P_i(t) = 1.$$

Вероятности состояний $P_i(t)$ находят путем решения системы *дифференциальных уравнений* (*уравнений Колмогорова*), имеющих вид

$$\frac{dP_i(t)}{dt} = \sum_{j=1}^{n} \lambda_{ji} P_j(t) - P_i(t) \sum_{j=1}^{n} \lambda_{ij}, \tag{2.8}$$

где $i = 0, 1, ..., n$.

Величина $\lambda_{ij} P_i(t)$ называется *потоком вероятности перехода из состояния* S_i в S_j, причем интенсивность потоков λ_{ij} может зависеть от времени или быть постоянной.

Уравнения (2.8) составляют по размеченному графу состояний системы, пользуясь следующим *мнемоническим правилом*:

производная вероятности каждого состояния равна сумме всех потоков вероятности, идущих из других состояний в данное состояние, минус сумма всех потоков вероятности, идущих из данного состояния в другие.

Чтобы решить систему дифференциальных уравнений (2.8), нужно задать начальное распределение вероятностей $P_0(0), P_1(0), ..., P_i(0), ..., P_n(0)$. Для решения применяют численные методы.

Финальные вероятности состояний

Если процесс, протекающий в системе, длится достаточно долго, то имеет смысл говорить о предельном поведении вероятностей $P_i(t)$ при $t \to \infty$. В некоторых случаях существуют финальные (предельные) вероятности состояний:

$$P_i = \lim_{t \to \infty} P_i(t),$$

где $i = 0, 1, ..., n,$

не зависящие от того, в каком состоянии система S находилась в начальный момент. Говорят, что в системе S устанавливается предельный стационарный режим, в ходе которого она переходит из состояния в состояние, но вероятности состояний P_i уже не меняются. Система, для которой существуют финальные вероятности, называется *эргодической*, а соответствующий случайный процесс — *эргодическим*.

Финальные вероятности состояний (если они существуют) могут быть получены путем решения системы линейных алгебраических уравнений, которые получаются из дифференциальных уравнений Колмогорова, если приравнять производные к нулю, а вероятностные функции состояний $P_1(t)$, ..., $P_n(t)$ в правых частях уравнений (2.8) заменить соответственно на неизвестные финальные вероятности P_1, ..., P_n.

Таким образом, для системы S с n состояниями получается система n линейных однородных алгебраических уравнений с n неизвестными P_0, P_1, ..., P_n, которые можно найти с точностью до произвольного множителя. Для нахождения точного значения P_0, P_1, ..., P_n к уравнениям добавляют нормировочное условие $P_0 + P_1 + + ... + P_n = 1$, пользуясь которым можно выразить любую из вероятностей P_i через другие и отбросить одно из уравнений.

Пример 2.3. Имеется размеченный граф состояний системы S (рис. 2.4). Необходимо составить систему дифференциальных уравнений Колмогорова и записать начальные условия для решения этой системы, если известно, что в начальный момент система находилась в состоянии S_1.

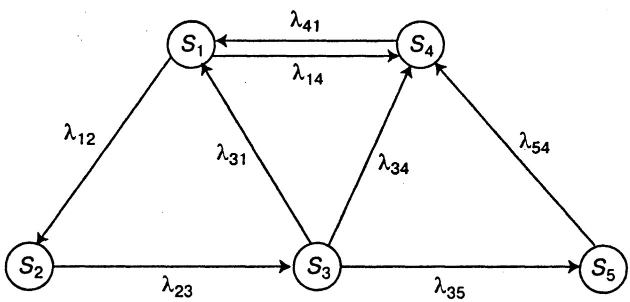


Рис. 2.4. Граф состояний системы

Решение

Согласно приведенному мнемоническому правилу система дифференциальных уравнений Колмогорова имеет вид

$$\begin{cases} \dfrac{dP_1}{dt} = \lambda_{31}P_3 + \lambda_{41}P_4 - \lambda_{12}P_1 - \lambda_{14}P_1 \\ \dfrac{dP_2}{dt} = \lambda_{12}P_1 - \lambda_{23}P_2 \\ \dfrac{dP_3}{dt} = \lambda_{23}P_2 - (\lambda_{31}P_3 + \lambda_{34}P_3 + \lambda_{35}P_3); \\ \dfrac{dP_4}{dt} = \lambda_{14}P_1 + \lambda_{34}P_3 + \lambda_{54}P_5 - \lambda_{41}P_4 \\ \dfrac{dP_5}{dt} = \lambda_{35}P_3 + \lambda_{54}P_5. \end{cases} \qquad (2.9)$$

Начальные условия при $t = 0$:

$$P_1 = 1;\ P_2 = P_3 = P_4 = P_5 = 0.$$

Рассмотрим, что произойдет с системой S, описываемой дифференциальными уравнениями Колмогорова, при $t \to \infty$. Известно, что в случае сообщающихся состояний функции $P_1(t)$, $P_2(t)$, ..., $P_n(t)$ стремятся к предельным (финальным) вероятностям состояний системы S. Финальные вероятности не зависят от времени. Поэтому в системе дифференциальных уравнений Колмогорова все левые части уравнений (производные) принимают равными нулю. При этом система дифференциальных уравнений превратится в систему линейных алгебраических уравнений.

Для нашего примера система (2.9) будет иметь вид

$$\begin{cases} 0 = \lambda_{31}P_3 + \lambda_{41}P_4 - (\lambda_{12}P_1 - \lambda_{14}P_1) \\ 0 = \lambda_{12}P_1 - \lambda_{23}P_2 \\ 0 = \lambda_{23}P_2 - (\lambda_{31}P_3 + \lambda_{34}P_3 + \lambda_{35}P_3) \\ 0 = \lambda_{14}P_1 + \lambda_{34}P_3 + \lambda_{54}P_5 - \lambda_{41}P_4 \\ 0 = \lambda_{35}P_3 - \lambda_{54}P_5. \end{cases} \qquad (2.10)$$

Решая ее с учетом условия $P_1 + P_2 + P_3 + P_4 + P_5 = 1$, получим все предельные вероятности. Эти вероятности представляют собой не что иное, как среднее относительное время пребывания системы в данном состоянии.

Необходимые и достаточные условия существования финальных вероятностей

Для существования финальных вероятностей одного условия λ_{ij} = const недостаточно, требуется выполнение еще некоторых условий, проверить которые можно по графу состояний, выделив в нем так называемые существенные и несущественные состояния.

Состояние S_i называется *существенным*, если нет другого состояния S_j, т. е. такого, что, перейдя однажды каким-то способом из S_i в S_j, система уже не может вернуться в S_i.

Все состояния, не обладающие таким свойством, называются *несущественными.*

Рассмотрим **пример**, представленный на рис. 2.5.

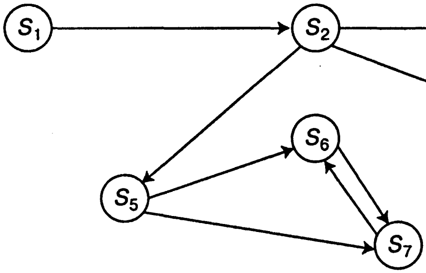


Рис. 2.5. Граф состояний системы S

Состояния S_1, S_2 и S_5 – несущественные, так как из S_1 можно уйти, например, в состояние S_2 и не вернуться, а из состояния S_2 – в состояние S_3 или S_4 и не вернуться, аналогично из состояния S_5 – в состояние S_6 и S_7. Состояния S_3, S_4, S_6 и S_7 – существенные состояния.

Теорема. При конечном числе состояний для существования финальных вероятностей *необходимо и достаточно*, чтобы из каждого существенного состояния можно было (за какое-то число шагов) перейти в каждое другое существенное состояние.

Граф из примера рис. 2.5 этому условию не удовлетворяет, так как из существенного состояния S_4 нельзя перейти в существенное состояние S_7. Если система S имеет конечное число состояний S_1, S_2, ..., S_n, то для существования финальных вероятностей достаточно, чтобы из любого состояния системы можно было (за какое-то число шагов) перейти в любое другое состояние.

Если число состояний S_1, S_2, ..., S_n бесконечно, то это условие перестает быть достаточным, и существование финальных вероятностей зависит не только от графа состояний, но и от интенсивности λ_{ij}.

При исследовании непрерывных марковских цепей, как было уже отмечено, часто бывает удобно представить переход системы из состояния в состояние как воздействие каких-то потоков событий (поток заявок на обслуживание, поток автомобилей, поток документов и т. п.). Различают следующие *основные свойства,* которыми могут обладать случайные потоки событий:

- стационарность;
- ординарность;
- отсутствие последействия.

Свойство *стационарности* проявляется в том, что вероятность попадания того или иного числа событий на участок времени τ зависит только от длины участка и не зависит от расположения на оси $0t$. Другими словами, стационарность означает неизменность вероятностного режима потока событий во времени. Поток, обладающий свойством стационарности, называют *стационарным.* Для стационарного потока среднее число событий, воздействующих на систему в течение единицы времени, остается постоянным. Реальные потоки событий в экономике предприятия являются в действительности стационарными лишь на ограниченных участках времени.

Свойство *ординарности* потока присутствует, если вероятность попадания на элементарный участок времени двух и более событий пренебрежимо мала по сравнению с длиной этого участка. Свойство ординарности означает, что за малый промежуток времени практически невозможно появление более одного события. Поток, обладающий свойством *ординарности*, называют *ординарным.* Реальные потоки событий в различных экономических системах либо являются ординарными, либо могут быть достаточно просто приведены к ординарным.

Отсутствие последействия – это свойство потока, которое состоит в том, что для любых непересекающихся участков времени количество событий, попадающих на один из них, не зависит от того, сколько событий попало на другие участки времени. Поток, обладающий свойством отсутствия последействия, называют *потоком без последействия.* Поток событий, одновременно обладающий свойствами стационарности, ординарности и отсутствия последействия, называется простейшим потоком событий.

Под интенсивностью потока понимают

$$\lambda(t)=\lim_{\tau\to 0}\frac{m(t,t+\tau)}{\tau}, \tag{2.11}$$

где $m(t,\ t+\tau)$ – среднее число событий в $(t,\ t+\tau)$.

Для простейшего потока интенсивность λ = const. Если поток событий не имеет последействия, ординарен, но не стационарен, то его называют *нестационарным пуассоновским потоком*, а его интенсивность зависит от времени, т. е. $\lambda = \lambda\ (t)$.

В пуассоновском потоке событий (стационарном и нестационарном) число событий потока, попадающих на любой участок, распределено по *закону Пуассона*:

$$P_m = \frac{a_m}{m!}e^{-a} \qquad m = 0, 1, \ldots, \tag{2.12}$$

где P_m – вероятность попадания на участок m событий (приложение 7);
a – среднее число событий, приходящееся на участок.

Для простейшего потока $a = \lambda\tau$, а для нестационарного пуассоновского потока

$$a = \int_{t_0}^{t_0+\tau} \lambda(t)dt, \tag{2.13}$$

где τ – длина участка времени;
t_0 – начало участка τ.

Отметим еще одно важное свойство простейшего потока событий. Промежуток времени t между соседними событиями распределен по показательному закону, а его среднее значение $\overline{T}$ и среднее квадратическое отклонение σ равны, т. е.

$$\overline{T} = \sigma = \frac{1}{\lambda}, \tag{2.14}$$

где λ – интенсивность потока.

Для нестационарного пуассоновского потока закон распределения промежутка t уже не является показательным, так как зависит от положения на оси $0t$ и вида зависимости $\lambda(t)$. Однако для некоторых задач при сравнительно небольших изменениях $\lambda(t)$ его можно приближенно считать показательным с интенсивностью λ, равной среднему значению $\lambda(t)$.

Таким образом, для исследуемой системы S с дискретными состояниями и непрерывным временем переходы из состояния в состояние происходят под действием пуассоновских потоков событий с определенной интенсивностью λ_{ij}.

Рассмотрим еще одну типичную схему непрерывных марковских цепей, так называемую схему гибели и размножения, часто встречающуюся в разнообразных практических задачах.

Марковский процесс с дискретными состояниями S_0, S_1, S_2, ..., S_n называется процессом *гибели и размножения*, если все состояния можно вытянуть в одну цепочку, в которой каждое из средних состояний (S_1, S_2, ..., S_{n-1}) может переходить только в соседние состояния, которые, в свою очередь, переходят обратно, а крайние состояния (S_0 и S_n) переходят только в соседние состояния (рис. 2.6).

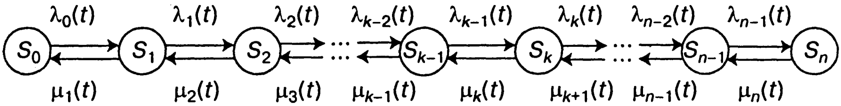


Рис. 2.6. Граф состояний для процесса гибели и размножения

Название взято из биологических задач, где состояние популяции S_k означает наличие в ней k единиц особей.

Переход вправо связан с размножением единиц, а влево – с их гибелью.

$\lambda_0(t)$, $\lambda_1(t)$, $\lambda_2(t)$, ..., $\lambda_{n-1}(t)$ – интенсивности размножения,

$\mu_1(t)$, $\mu_2(t)$, ..., $\mu_n(t)$ – интенсивности гибели.

У λ и μ индекс того состояния, из которого стрелка выходит.

С состоянием S_k связана неслучайная величина X_k: если система S в момент времени t находится в состоянии S_k, то дискретная случайная величина $X(t)$, связанная с функционированием системы, принимает значение k. Таким образом, получаем случайный процесс $X(t)$, который в случайные, заранее неизвестные моменты времени скачком изменяет свое состояние.

Марковским процессом гибели и размножения с непрерывным временем называется такой случайный процесс, который может принимать только целые неотрицательные значения. Изменения этого процесса могут происходить в любой момент времени, т. е. в любой момент времени он может либо увеличиться на единицу, либо уменьшиться на единицу, либо остаться неизменным.

В практике встречаются процессы чистого размножения и чистой гибели. *Процессом чистого размножения* называется такой процесс гибели и размножения, у которого интенсивности всех потоков гибели равны нулю; аналогично *процессом чистой гибели* называется такой процесс гибели и размножения, у которого интенсивности всех потоков размножения равны нулю.

Пример 2.4. Рассмотрим эксплуатацию моделей автомобилей одной марки в крупной транспортной фирме (на предприятии). Интенсивность поступления автомобилей на предприятие равна $\lambda(t)$. Каждый поступивший на предприятие автомобиль списывается через случайное время T_C. Срок службы автомобиля T_C распределен по показательному закону с параметром μ. Процесс эксплуатации автомобилей является случайным процессом. $A(t)$ – число автомобилей данной марки, находящихся в эксплуатации в момент t.

Найдем одномерный закон распределения случайного процесса $P_i(t) = P\{(A(t) = i\}$, если:

1) нет ограничений на число эксплуатируемых машин;

2) на предприятии может эксплуатироваться не более n автомобилей.

Решение

1. Случайный процесс эксплуатации автомобилей есть процесс гибели и размножения, размеченный граф которого представлен на рис. 2.7.

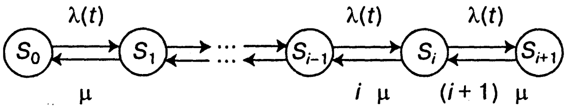


Рис. 2.7. Граф состояний

Система уравнений Колмогорова, соответствующая этому графу, имеет вид

$$\frac{dP_0(t)}{dt} = \mu P_1(t) - \lambda(t)P_0(t)$$

....

$$\frac{dP_i(t)}{dt} = \lambda(t)P_{i-1}(t) + (i+1)\mu P_{i+1}(t) - (\lambda(t) + i\mu)P_i(t), \qquad (2.15)$$

....

где $i = 1, 2,...$

Если в начальный момент времени $t = 0$ на предприятии не было ни одного автомобиля, то решать эту систему уравнений нужно при начальных условиях $P_0(0) = 1$, $P_i(0) = 0$ ($i = 1, 2, ...$). Если при $t = 0$ на предприятии было k автомобилей ($k = 1, 2, ...$), то начальные условия будут иметь вид

$$P_k(0) = 1,\ P_i(0) = 0\ (i = 1, 2, ..., i \neq k).$$

2. Если на предприятии может эксплуатироваться не более n автомобилей моделей одной марки, то имеет место процесс гибели и размножения с ограниченным числом состояний n, размеченный граф которого представлен на рис. 2.8.

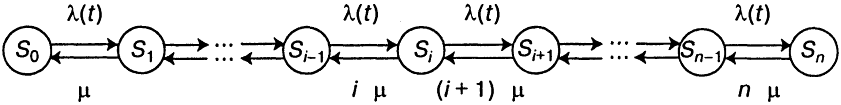


Рис. 2.8. Граф состояний

Система уравнений Колмогорова для размеченного графа (рис. 2.8) имеет вид:

$$\frac{dP_0(t)}{dt} = \mu P_1(t) - \lambda(t) P_0(t)$$

....

$$\frac{dP_i(t)}{dt} = \lambda(t) P_{i-1}(t) + (i+1)\mu P_{i+1}(t) - (\lambda(t) + i\mu) P_i(t) \quad i = \overline{1, n-1} \qquad (2.16)$$

....

$$\frac{dP_n(t)}{dt} = \lambda(t) P_{n-1}(t) - n\mu P_n(t).$$

Эту систему надо решать при начальных условиях, рассмотренных выше. Решения систем уравнений (2.15) и (2.16) являются одномерными законами распределения $P_i(t)$. Отыскание решений систем (2.15) и (2.16) в общем виде при произвольном виде функции $\lambda(t)$ представляет значительные трудности и не имеет практических приложений.

При постоянных интенсивностях потоков гибели и размножения и конечном числе состояний будет существовать стационарный режим. Система S с конечным числом состояний $(n + 1)$, в которой протекает процесс гибели и размножения с постоянными интенсивностями потоков гибели и размножения, является простейшей эргодической системой. Размеченный граф состояний для такой системы представлен на рис. 2.9.

Предельные (финальные) вероятности состояний для простейшего эргодического процесса гибели и размножения, находящегося в стационарном режиме, определяются по следующим формулам:

$$P_k = \frac{\lambda_0 \lambda_1 \cdot \ldots \cdot \lambda_{k-1}}{\mu_1 \mu_2 \cdot \ldots \cdot \mu_k} \cdot P_0, \quad k = 1, 2, \ldots, n; \qquad (2.17)$$

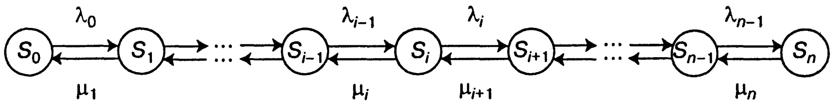


Рис. 2.9. Граф состояний

$$P_0 = \left\{1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0\lambda_1}{\mu_1\mu_2} + \ldots + \frac{\lambda_0\lambda_1 \cdot \ldots \cdot \lambda_{n-1}}{\mu_1\mu_2 \cdot \ldots \cdot \mu_n}\right\}^{-1}. \qquad (2.18)$$

Правило. Вероятность k-го состояния в схеме гибели и размножения равна дроби, в числителе которой стоит произведение всех интенсивностей размножения, стоящих левее S_k, а в знаменателе – произведение всех интенсивностей гибели, стоящих левее S_k, умноженной на вероятность крайнего левого состояния системы P_0.

В примере 2.4 для стационарного режима если интенсивность поступления автомобилей постоянная ($\lambda(t) = \lambda = \text{const}$), то финальные вероятности состояний при условии, что нет ограничений на число автомобилей на предприятии, равны:

$$P_0 = e^{-\frac{\lambda}{\mu}}, \qquad (2.19)$$

$$P_k = \left[\frac{\left(\frac{\lambda}{\mu}\right)^k}{k!}\right] e^{-\frac{\lambda}{\mu}}. \qquad (2.20)$$

При этом математическое ожидание числа эксплуатируемых автомобилей равно его дисперсии:

$$M[A(t)] = D[A(t)] = \frac{\lambda}{\mu}. \qquad (2.21)$$

Если существует ограничение по числу автомобилей на предприятии (не более n), то финальные вероятности равны

$$P_0 = \frac{e^{-\alpha}}{\sum\limits_{k=0}^{n} \alpha^k \cdot e^{-\alpha} / k!}, \qquad (2.22)$$

где $\alpha = \dfrac{\lambda}{\mu}$.

$$P_k = P_0 \cdot \frac{\alpha^k}{k!}, \qquad (2.23)$$

где $k = 0,1,2, \ldots, n$.

Математическое ожидание числа эксплуатируемых автомобилей в стационарном режиме

$$M[A(t)] = \sum_{k=0}^{n} k \cdot P_k. \qquad (2.24)$$

Пример 2.5. В состав ЭВМ входят четыре накопителя на магнитных дисках (НМД). Бригада в составе четырех человек обслуживающего персонала проводит профилактический ремонт каждого диска. Суммарный поток моментов окончания ремонтов для всей бригады – пуассоновский с интенсивностью $\lambda(t)$. После окончания ремонта диск проверяется; с вероятностью P он оказывается работоспособным (время проверки мало, и им можно пренебречь по сравнению со временем профилактики). Если диск оказался неработоспособным, то вновь проводится его профилактика (время на которую не зависит от того, проводилась ли она ранее) и т.д. В начальный момент все НМД нуждаются в профилактическом ремонте[1].

Требуется:

1) построить граф состояний для системы S (четыре НМД);

2) написать дифференциальные уравнения для вероятностей состояний;

3) найти математическое ожидание числа дисков M_τ, успешно прошедших профилактику к моменту τ.

Решение

1. Граф состояний показан на рис. 2.10, в котором

S_0 – все четыре НМД нуждаются в профилактическом ремонте;

S_1 – один НМД успешно прошел профилактику, а три НМД нуждаются в профилактическом ремонте;

S_2 – два НМД успешно прошли профилактику, а два нуждаются в профилактическом ремонте;

S_3 – три НМД успешно прошли профилактику, один нуждается в профилактическом ремонте;

S_4 – все четыре НМД успешно прошли профилактику.

[1] *Вентцель Е. С., Овчаров Л. А.* Прикладные задачи теории вероятностей. – М.: Радио и связь, 1983.

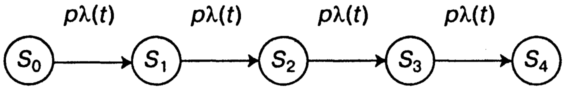


Рис. 2.10. Граф состояний системы

Каждый профилактический ремонт успешно заканчивается с вероятностью p, что равносильно p-преобразованию потока окончаний ремонтов, после которого он остается пуассоновским, но с интенсивностью $p\lambda(t)$. В этом примере мы имеем дело с процессом чистого размножения с ограниченным числом состояний.

2. Уравнения Колмогорова имеют следующий вид:

$$\begin{aligned}
&\frac{dP_0(t)}{dt} = -p\lambda(t)P_0(t);\\
&\frac{dP_1(t)}{dt} = p\lambda(t)(P_0(t) - P_1(t));\\
&\frac{dP_2(t)}{dt} = p\lambda(t)(P_1(t) - P_2(t));\\
&\frac{dP_3(t)}{dt} = p\lambda(t)(P_2(t) - P_3(t));\\
&\frac{dP_4(t)}{dt} = p\lambda(t)P_3(t).
\end{aligned} \tag{2.25}$$

Начальные условия $P_0(0) = 1$; $P_1(0) = ... = P_4(0) = 0$. При постоянной интенсивности $\lambda(t) = \lambda$ и вероятности состояний определяются по следующим формулам:

$$\begin{aligned}
&P_0(t) = e^{-\lambda pt}; \qquad && P_1(t) = \lambda pte^{-\lambda pt};\\
&P_2(t) = \frac{(\lambda pt)^2}{2!}e^{-\lambda pt}; && P_2(t) = \frac{(\lambda pt)^3}{3!}e^{-\lambda pt};\\
&P_4(t) = 1 - \sum_{i=0}^{3} P_i(t).
\end{aligned} \tag{2.26}$$

3. Математическое ожидание числа дисков, успешно прошедших профилактику к моменту τ, равно:

$$M_\tau = \sum_{i=0}^{n-1} iP_i(t) + nP_n(t), \tag{2.27}$$

где $n = 4$.

Пример 2.6. Рассмотрим производство автомобилей на заводе. Поток производимых автомобилей – нестационарный пуассоновский с интенсивностью $\lambda(t)$. Найдем одномерный закон распределения случайного процесса $X(t)$ – число выпущенных автомобилей к моменту времени t, если в момент $t = 0$ начат выпуск автомобилей.

Решение

Очевидно, что здесь процесс чистого размножения без ограничения на число состояний, при этом $\lambda_i(t) = \lambda(t)$, так как интенсивность выпуска автомобилей не зависит от того, сколько их уже выпущено. Граф состояний такого процесса показан на рис. 2.11.

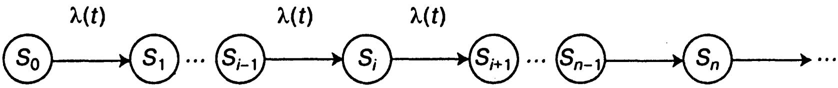


Рис. 2.11. Граф состояний

Одномерный закон распределения случайного процесса $X(t)$ для графа, изображенного на рис. 2.11, определяется следующей системой уравнений Колмогорова:

$$\frac{dP_0(t)}{dt} = -\lambda(t)P_0(t)$$

....

$$\frac{dP_i(t)}{dt} = -\lambda(t)P_i(t) + \lambda(t)P_{i-1}(t), \quad i = 1, 2, \ldots .$$

....

Так как число выпущенных автомобилей $X(t)$ на любой фиксированный момент t распределено по закону Пуассона с параметром

$$a(t) = \int_0^t \lambda(t)\,dt,$$

то

$$P\{X(t) = i\} = P_i(t) = \frac{[a(t)]^i}{i!} \cdot e^{-a(t)}, \quad i = 0,\ 1,\ 2,\ \ldots .$$

$$M[X(t)] = D[X(t)] = a(t).$$

Рассмотренный в этом примере процесс $X(t)$ называется *неоднородным процессом Пуассона*. Если интенсивность $\lambda(t) = \lambda = \text{const}$, то получим *однородный процесс Пуассона.* Для такого процесса при $P_0(0) = 1$, $P_i(0) = 0$ $(i > 0)$

$$P\{X(t)=i\} = P_i(t) = \frac{(\lambda t)^i}{i!} e^{-\lambda t}, \quad i = 0,\ 1,\ 2,\ \ldots .$$

Характеристиками процесса Пуассона будут

$$M[X(t)] = D[X(t)] = \lambda t.$$

2.4. Моделирование работы подвижного состава с использованием марковских случайных процессов

Представим автомобиль как некоторую систему S с дискретными состояниями S_0, S_1, ..., S_n, которая переходит из состояния в состояние под влиянием случайных событий (отказов).

На стадии прогнозирования (планирования) работы автомобиля целесообразно рассматривать следующие состояния, в которых подвижной состав может находиться в процессе эксплуатации и которые характеризуются целодневными простоями:

S_0 – исправен, работает;

S_1 – находится на капитальном ремонте (КР);

S_2 – проходит ТО-2;

S_3 – находится в текущем ремонте (ТР);

S_4 – исправен, не работает по организационным причинам (без водителя, шин, запасных частей);

S_5 – не работает, снятие агрегата для отправки на капитальный ремонт;

S_6 – не работает, списание агрегата, замена на новый;

S_7 – исправен, не работает (выходные и праздничные дни);

S_8 – списывается.

Надо отметить, что в настоящее время вышеперечисленные состояния автомобиля планируются при разработке годовой программы работы автотранспортного предприятия (АТП), при этом состояния S_3, S_5, S_6 объединяются в одно состояние «находится в ТР».

Для анализа процесса эксплуатации автомобиля как случайного процесса с дискретными состояниями удобно воспользоваться геометрической схемой, так называемым графом состояний (рис. 2.12). Граф состояний изображает возможные состояния автомобиля и его возможные переходы из состояния в состояние.

На рис. 2.12 через λ_{ij} и μ_{ji} обозначены плотности вероятностей перехода автомобиля из состояния S_i в состояние S_j. Напри-

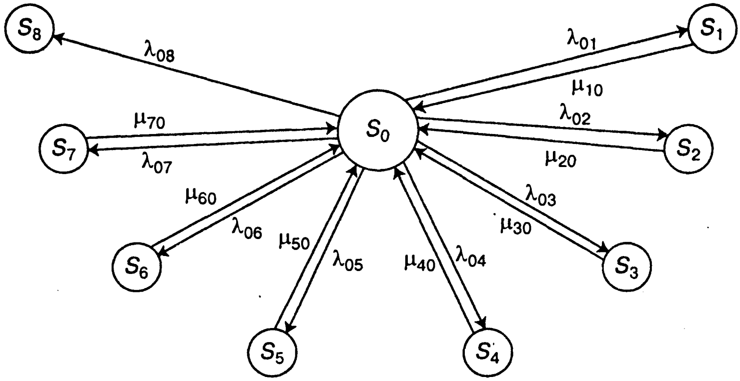


Рис. 2.12. Граф состояний автомобиля

мер, λ_{03} – плотность вероятности перехода автомобиля из состояния «исправен, работает» в состояние «находится в текущем ремонте».

Можно считать, что события, переводящие автомобиль из состояния в состояние, представляют собой потоки событий (например, потоки отказов). Если все потоки событий, переводящие систему (автомобиль) из состояния в состояние, пуассоновские (стационарные или нестационарные), то процесс, протекающий в системе, будет марковским, а плотности вероятности перехода λ_{ij} в непрерывной цепи Маркова представляют собой интенсивности потока событий, переводящего систему из состояния S_i в состояние S_j. Например, λ_{03} – интенсивность потока отказов автомобиля, который переводит автомобиль из состояния «исправен, работает» в состояние «находится в ТР».

Рассматриваемые состояния автомобиля S_j характеризуются средним числом дней пребывания автомобиля в каждом состоянии $Д_j$. Показатели $Д_j$ находят отражение в статистической отчетности автотранспортного предприятия. Отношение

$$P_j = \frac{Д_j}{Д_к}, \qquad (2.28)$$

где $Д_к$ – число календарных дней в году.

$Д_к$ можно трактовать как вероятность нахождения автомобиля в j-м состоянии.

Вероятности состояний автомобиля $P_0, P_1, P_2, ..., P_j, ..., P_n$ как функции пробега в случае марковского процесса с дискретными состояниями и непрерывным временем удовлетворяют определенного вида дифференциальным уравнениям (уравнениям Колмогорова), записываемым в виде

$$\begin{cases} \dfrac{dP_0(L)}{dL} = -\ell_c \sum\limits_{i=1}^{n} \lambda_{0i}(L) \cdot P_0(L) + \sum\limits_{i=1}^{n-1} \mu_{i0}(L) P_i(L) \\ \dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dP_i(L)}{dL} = \ell_c \lambda_{0i}(L) \cdot P_0(L) - \mu_{i0}(L) P_i(L) \\ \dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dP_n(L)}{dL} = \ell_c \lambda_{0n}(L) P_0(L), \end{cases} \tag{2.29}$$

где $P_i(L)$ – вероятность нахождения автомобиля в i-м состоянии, $i = \overline{0, n}$;

$\lambda_{0i}(L)$ – интенсивность перехода автомобиля из нулевого в i-е состояние, $i = \overline{1, n}$;

$\mu_{i0}(L)$ – интенсивность перехода из i-го в нулевое состояние, $i = \overline{1, n-1}$;

ℓ_c – коэффициент, отражающий связь между наработками в днях и километрах пробега (среднесуточный пробег, тыс. км).

Число уравнений в системе (2.29) зависит от числа состояний автомобиля. Вероятность нахождения автомобиля в состоянии «исправен-работает» $P_0(L)$ представляет собой коэффициент выпуска $\alpha_в(L)$, а сумма вероятностей $P_0(L) + P_4(L) + P_7(L) = k_{тг}(L)$ – коэффициент технической готовности автомобиля.

Поскольку большинство интенсивностей перехода зависят от пробега, то решение системы (2.29) производится с помощью методов численного интегрирования, например *метода Рунге-Кутта.*

Необходимо учесть, что для расчета производственной программы АТП требуется зачастую определять показатели работы группы автомобилей определенной модели j-го возраста (коэффициент выпуска и годовой пробег автомобиля j-й возрастной группы).

Для описания процесса функционирования группы автомобилей может быть использован *метод динамики средних.* Этот метод вытекает из теории марковских случайных процессов. Удобство его заключается в том, что, зная возможные состояния одного (условного) автомобиля, можно моделировать процесс функционирования группы из любого числа автомобилей.

Схема, изображающая процесс работы условного автомобиля определенной модели, аналогична схеме рис. 2.12, лишь с той разницей, что через λ_{ij} и μ_{ji} обозначены средние интенсивности потоков событий, переводящих автомобиль из состояния S_i в состояние S_j, и наоборот. При этом каждое состояние характеризуется средней численностью автомобилей $N_j(t)$, находящихся в нем в момент времени t. Очевидно, что для любого t сумма численностей всех состояний равна общей численности автомобилей исследуемой группы:

$$N = \sum_{j=0}^{8} N_j(t).$$

Величина $N_j(t)$ для любого t представляет собой случайную величину, а при меняющемся t – случайную функцию времени.

Зная граф состояний (рис. 2.12) и соответствующие интенсивности перехода λ_{ij} и μ_{ji}, определим средние численности автомобилей $N_0(L)$, $N_1(L)$, $N_2(L)$, ... , $N_8(L)$ как функции пробега L.

Согласно графу состояний (рис. 2.12) система дифференциальных уравнений для средних численностей состояний запишется следующим образом:

$$\frac{dN0(L)}{dL} = -\ell c \cdot N_0(L)[\lambda_{01}(L) + \lambda_{02}(L) + \lambda_{03}(L) + \lambda_{04}(L) + \lambda_{05}(L) +$$
$$+ \lambda_{06}(L) + \lambda_{07}(L) + \lambda_{08}(L)] + \mu_{10}(L) \cdot N_1(L) + \mu_{20}(L) \cdot N_2(L) +$$
$$+ \mu_{30}(L) \cdot N_3(L) + \mu_{40}(L) \cdot N_4(L) + \mu_{50}(L)\, N_5(L) + \mu_{60}(L) \cdot N_6(L) +$$
$$+ \mu_{70}(L) \cdot N_7(L);$$

$$\frac{dN_1(L)}{dL} = \ell c \cdot N_0(L) \cdot \lambda_{01}(L) - \mu_{10}(L) \cdot N_1(L);$$

$$\frac{dN_2(L)}{dL} = \ell_c \cdot N_0(L) \cdot \lambda_{02}(L) - \mu_{20}(L) \cdot N_2(L);$$

$$\frac{dN_3(L)}{dL} = \ell_c \cdot N_0(L) \cdot \lambda_{03}(L) - \mu_{30}(L) \cdot N_3(L); \qquad (2.30)$$

$$\frac{dN_4(L)}{dL} = \ell_c \cdot N_0(L) \cdot \lambda_{04}(L) - \mu_{40}(L) \cdot N_4(L);$$

$$\frac{dN_5(L)}{dL} = \ell_c \cdot N_0(L) \cdot \lambda_{05}(L) - \mu_{50}(L) \cdot N_5(L);$$

$$\frac{dN_6(L)}{dL} = \ell_c \cdot N_0(L) \cdot \lambda_{06}(L) - \mu_{60}(L) \cdot N_6(L);$$

$$\frac{dN_7(L)}{dL} = \ell_c \cdot N_0(L) \cdot \lambda_{07}(L) - \mu_{70}(L) \cdot N_7(L);$$

$$\frac{dN_8(L)}{dL} = \ell_c \cdot N_0(L) \cdot \lambda_{08}(L).$$

Отношение $N_0(L)/N$ равно коэффициенту выпуска автомобилей определенной модели на пробеге L с начала их эксплуатации, а отношение $[N_0(L) + N_4(L) + N_7(L)]/N$ – коэффициенту технической готовности автомобилей.

Докажем, что формулы для определения коэффициентов технической готовности ($\kappa_{\text{тг}}$), выпуска подвижного состава ($\alpha_{\text{в}}$) являются частным случаем, соответствующим стационарному решению системы уравнений (2.30), описывающей функционирование автопарка.

Для расчета средней численности автомобилей, находящихся в исправном состоянии, можно предварительно объединить состояния S_1, S_5, S_6 в одно состояние: «исправен» – S_1'. Тогда граф состояний условного автомобиля примет вид, представленный на рис. 2.13.

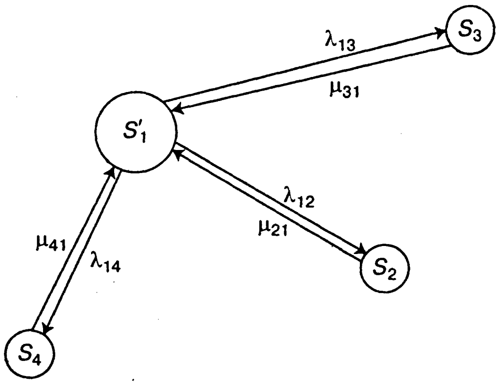


Рис. 2.13. Граф состояний условного автомобиля

Система дифференциальных уравнений для средних численностей подвижного состава запишется следующим образом:

$$\begin{cases} \dfrac{dN_1'(\ell)}{d\ell} = -\ell_c(\lambda_{12}(\ell)+\lambda_{13}(\ell)+\lambda_{14}(\ell))\cdot N_1'(\ell)+\mu_{21}(\ell)\cdot N_2(\ell)+ \\ +\mu_{31}(\ell)\cdot N_3(\ell)+\mu_{41}(\ell)\cdot N_4(\ell) \\ \dfrac{dN_2(\ell)}{d\ell} = -\mu_{21}(\ell)\cdot N_2(\ell)+\ell_c\cdot N_1'(\ell)\cdot\lambda_{12}(\ell) \\ \dfrac{dN_3(\ell)}{d\ell} = -\mu_{31}(\ell)\cdot N_3(\ell)+\ell_c\cdot N_1'(\ell)\cdot\lambda_{13}(\ell) \\ \dfrac{dN_4(\ell)}{d\ell} = -\mu_{41}(\ell)\cdot N_4(\ell)+\ell_c\cdot N_1'(\ell)\cdot\lambda_{14}(\ell). \end{cases} \tag{2.31}$$

Положим левые части уравнений равными нулю, получим систему алгебраических уравнений для средних численностей состояний автопарка, работающего в стационарном режиме:

$$\begin{cases} 0=-\ell_c(\lambda_{12}+\lambda_{13}+\lambda_{14})N_1' +\mu_{21}N_2+\mu_{31}N_3+\mu_{41}N_4 \\ 0=-\mu_{21}N_2+\ell_c\lambda_{12}N_1' \\ 0=-\mu_{31}N_3+\ell_c\lambda_{13}N_1' \\ 0=-\mu_{41}N_4+\ell_c\lambda_{14}N_1'. \end{cases} \qquad (2.32)$$

Решим систему алгебраических уравнений с учетом так называемого нормировочного условия:

$$N_0 = N_1 + N_2 + N_3 + N_4, \qquad (2.33)$$

где N_0 – среднесписочная численность автопарка, шт.

Для **примера** из системы (2.32) определим неизвестные средние численности состояний, используя N_1'. Так, из второго и третьего уравнений имеем

$$N_2=\ell_c\frac{\lambda_{12}}{\mu_{21}}N_1'; \qquad (2.34)$$

$$N_3=\ell_c\frac{\lambda_{13}}{\mu_{31}}N_1'. \qquad (2.35)$$

Согласно нормировочному условию

$$N_4 = N_0 - N_1' - N_2 - N_3. \qquad (2.36)$$

Тогда

$$N_4=N_0-\left(N_1'+N_1'\ell_c\frac{\lambda_{12}}{\mu_{21}}+N_1'\ell_c\frac{\lambda_{13}}{\mu_{31}}\right)=$$

$$=N_0-\left(1+\ell_c\frac{\lambda_{12}}{\mu_{21}}+\ell_c\frac{\lambda_{13}}{\mu_{31}}\right)N_1'.$$

Подставляя в первое уравнение (2.32), получим:

$$0 = -\ell_c(\lambda_{12} + \lambda_{13} + \lambda_{14})N_1' + \ell_c\lambda_{12}N_1' + \ell_c\lambda_{13}N_1' +$$

$$+\mu_{41}\left[N_0-\left(1+\ell_c\frac{\lambda_{12}}{\mu_{21}}+\ell_c\frac{\lambda_{13}}{\mu_{31}}\right)N_1'\right]=$$

$$=-\ell_c\lambda_{14}N_1'+\mu_{41}N_0-\mu_{41}\left(1+\ell_c\frac{\lambda_{12}}{\mu_{21}}+\ell_c\frac{\lambda_{13}}{\mu_{31}}\right)N_1'.$$

Разделим полученное уравнение на μ_{41}:

$$-\ell_c \frac{\lambda_{14}}{\mu_{41}} N_1' + N_0 - \left(1 + \ell_c \frac{\lambda_{12}}{\mu_{21}} + \ell_c \frac{\lambda_{13}}{\mu_{31}}\right) N_1' = 0.$$

Последнее уравнение можно записать следующим образом:

$$N_0 = \left(1 + \ell_c \frac{\lambda_{12}}{\mu_{21}} + \ell_c \frac{\lambda_{13}}{\mu_{31}} + \frac{\lambda_{14}}{\mu_{41}} \ell_c\right) N_1'.$$

Тогда коэффициент технической готовности равен:

$$\text{К}_{\text{тг}} = \frac{N_1'}{N_0} = \frac{1}{1 + \ell_c \left(\frac{\lambda_{12}}{\mu_{21}} + \frac{\lambda_{13}}{\mu_{31}} + \frac{\lambda_{14}}{\mu_{41}}\right)}, \tag{2.37}$$

где λ_{12}, λ_{13}, λ_{14} – интенсивности перехода автомобиля в состояния «техническое обслуживание», «текущий ремонт», «капитальный ремонт» соответственно, отк/тыс. км;

μ_{21}, μ_{31}, μ_{41} – интенсивности восстановления, равные обратным средним величинам продолжительности соответствующих ремонтных воздействий технического обслуживания (ТО-2), текущего ремонта (ТР), капитального ремонта (КР), отк/день.

Отношение

$$\frac{\lambda_{ij}}{\mu_{ji}} = \frac{\dfrac{\text{отк}}{\text{тыс. км}}}{\dfrac{\text{отк}}{\text{день}}} = \frac{\text{день}}{\text{тыс. км}}.$$

Таким образом, $\dfrac{\lambda_{ij}}{\mu_{ji}} = d_j$ – удельная величина, характеризующая количество дней в j-м состоянии (ТО-2, ТР, КР) на 1 тыс. км пробега. Тогда формулу (2.37) можно записать в виде:

$$\text{к}_{\text{тг}} = \frac{1}{1 + \ell_c (d_2 + d_3 + d_4)} = \frac{1}{1 + \ell_c d_{\text{ТО-Р}}}. \tag{2.38}$$

Очевидно, формула (2.38) есть частный случай, соответствующий стационарному режиму работы автомобиля, который является решением системы алгебраических уравнений.

Рассмотрим все потоки событий, переводящие условный автомобиль из состояния в состояние. Характер потока отказов автомобиля, переводящего условный автомобиль из состояния «исправен, работает» в состояние «находится в текущем ремонте», не изменяется. При определении его величины учитывается возрастная структура автомобилей данной модели.

Наработка до первого капитального ремонта автомобиля подчиняется нормальному закону распределения с коэффициентом вариации 0,1–0,33. Вместе с тем следует отметить значительное абсолютное рассеивание пробегов до первого капитального ремонта автомобиля в исследуемых группах подвижного состава. Размах между минимальным и максимальным пробегами может составить пробег, примерно равный среднему пробегу до первого капитального ремонта этих автомобилей.

Таким образом, поток событий, который переводит автомобиль в состояние «капитальный ремонт», протекает на значительном интервале пробега. В этом потоке интенсивность $\lambda_{01}(L)$ (среднее число событий в единицу пробега) зависит от пробега, т. е. поток является нестационарным.

Очевидно, на малом интервале пробега автомобиля (1–2 тыс. км) интенсивность $\lambda_{01}(L)$ меняется сравнительно медленно. В этом случае закон распределения наработки до капитального ремонта можно приближенно считать показательным, а интенсивность λ_{01} принимать равной среднему значению $\lambda_{01}(L)$ на этом интервале. Аналогичные утверждения справедливы относительно потоков отказов, переводящих условный автомобиль в состояния «капитальный ремонт агрегата» и «списание агрегата».

Общий поток отказов, связанный с попаданием автомобилей исследуемой группы в ТО-2, получается путем наложения (суперпозиции) потоков «ТО-2» этих автомобилей. Как показывают расчеты, распределение интервала пробега между событиями в этом потоке подчиняется показательному закону. При этом поток «ТО-2» всех исследуемых автомобилей является пуассоновским.

Образ потока отказов, связанного со списанием автомобиля, является условным. Действительно, если автомобиль отказывает в тот момент, когда происходит первое событие данного потока, то совершенно все равно, продолжается после этого поток отказов или прекращается: судьба автомобиля от этого уже не зависит. В случае когда элемент (автомобиль) не подлежит восстановлению, поток отказов является пуассоновским.

Поток отказов автомобиля, связанный со списанием, является нестационарным, так как пробег до списания подвижного состава подчиняется закону распределения, отличному от показательного. Очевидно, на малом интервале пробега автомобиля (1–2 тыс. км) интенсивность отказов меняется сравнительно медленно, в таком случае закон распределения событий можно приблизительно считать показательным и для описания процесса эксплуатации автомобиля использовать марковскую схему.

Характер остальных потоков событий, связанных с процессом работы группы автомобилей, не изменяется.

Таким образом, все средние потоки, переводящие условный автомобиль из состояния в состояние, либо пуассоновские, либо сводятся к ним путем рассмотрения процесса эксплуатации на малых интервалах пробега (1–2 тыс. км) и корректировки исходного потока отказов деталей для исключения последействия. Это позволяет использовать метод динамики средних для описания процесса эксплуатации группы автомобилей.

В табл. 2.1 приведены формулы для расчета интенсивностей перехода λ_{ij} и μ_{ji}.

Таблица 2.1

Интенсивности перехода λ_{ij} и μ_{ij} для расчета комплексных показателей надежности автомобилей МАЗ

Интенсивность	Формула, принятая в расчете	Примечание
Исправен – капитальный ремонт	$\lambda_{01} = \sum\limits_{k=1}^{K} \varphi_k(L)$	Плотность распределения наработки до k-го капитального ремонта автомобиля – $\varphi_k(L)$
Исправен – проходит техническое обслуживание (ТО-2)	$\lambda_{02}(L) = \sum\limits_{i=1}^{\infty} f_{\text{то}i}(L)$ $\lambda_{02}(L) = (\bar{L}_{\text{то}})^{-1}$	$f_{\text{то}i}$ – плотность распределения наработки до i-го ТО-2; $\bar{L}_{\text{то}}$ – средняя периодичность ТО-2
Исправен – находится в текущем ремонте	$\lambda_{03}(L) = \sum\limits_{f=1}^{F} \omega_f(L)$	$\omega_f(L)$ – параметр потока отказов f детали по интервалам пробега L; F – число ДЛН автомобиля, шт.
Исправен – простаивает по организационным причинам (без водителя и т. п.)	$\lambda_{04}(L) = \rho(L)$ $\lambda_{04}(L) = (l_c \bar{T}_{\text{пр}})^{-1}$	$\bar{T}_{\text{пр}}$ – среднее время между простоями; l_c – среднесуточный пробег, тыс. км;

Интенсивность	Формула, принятая в расчете	Примечание
Исправен – капитальный ремонт агрегата	$\lambda_{05}(L)=\sum\limits_{n=1}^{N}\omega_{п}^{кр\ аг}(L)$	$\omega_{п}^{кр\ аг}$ – параметр потока отказов автомобиля, связанных с капитальным ремонтом его агрегатов
Исправен – списание агрегата	$\lambda_{06}(L)=\sum\limits_{n=1}^{N}\omega_{п}^{сп\ аг}$	$\omega_{п}^{сп\ аг}(L)$ – параметр потока отказов автомобиля, связанных со списанием агрегатов; N – число агрегатов
Исправен – не работает (праздничные и выходные дни)	$\lambda_{07}(L)=\rho(L)$ $\lambda_{07}(L)=(l_c\cdot\bar{T}_{вых})^{-1}$	$\bar{T}_{вых}$ – среднее время между простоями
Исправен – списание автомобиля	$\lambda_{08}(L)=$ $=f_c(L)/[1-F_c(L)]$ $\lambda_{08}(L)=(L-L_0)/\sigma^2$ $L>270$ тыс. км	$F_c(L), f_c(L)$ – функция и плотность распределения наработки до списания автомобиля; принято распределение Рэлея с параметрами $\sigma^2=40^2$ тыс. км; $L_0=270$ тыс. км
Капитальный ремонт – исправен	$\mu_{10}=(\bar{T}_{кр})^{-1}$	$\bar{T}_{кр}$ – средняя продолжительность капитального ремонта
ТО-2 – исправен	$\mu_{20}=(\bar{T}_{то})^{-1}$	$\bar{T}_{то}$ – средняя продолжительность ТО-2
Находится в текущем ремонте – исправен	$\mu_{30}(L)=\eta(L)$ $\mu_{30}(L)=(\bar{T}_{т})^{-1}$	$\bar{T}_{т}$ – средняя продолжительность текущего ремонта
Простаивает по организационным причинам – исправен	$\mu_{40}=(\bar{T}_{п})^{-1}$	$\bar{T}_{п}$ – средняя продолжительность простоя
Капитальный ремонт агрегата – исправен	$\mu_{50}(L)=(\bar{T}_{кр}^{а})^{-1}$	$\bar{T}_{кр}^{а}$ – средняя продолжительность простоя при снятии агрегата

Интенсивность	Формула, принятая в расчете	Примечание
Списание агрегата – исправен	$\mu_{60}(L) = (\overline{T}_{р}^{с})^{-1}$	$\overline{T}_{р}^{с}$ – среднее время замены агрегата
Исправен, не работает (праздничные и выходные дни) – исправен, работает	$\mu_{70}(L) = (\overline{T}_{пр})^{-1}$	$\overline{T}_{пр}$ – средняя продолжительность простоя

Значения параметров модели (2.30) $\lambda_{03}(L)$, $\lambda_{05}(L)$, $\lambda_{06}(L)$ могут быть определены двумя способами. Согласно первому способу полученные значения параметров потока отказов автомобиля, связанных с его текущим ремонтом, капитальным ремонтом и списанием его агрегатов, аппроксимируются экспоненциальными зависимостями следующего вида:

$$\lambda_{0i}(L) = \exp(a_0 + a_1 \,.\, x + ... + a_n \,.\, x^n),$$

где x – пробег автомобиля с начала эксплуатации, тыс. км;
i – номер состояния, в котором находится автомобиль, i = 3,5,6.

Ошибка аппроксимации при небольших n бывает высокой и может достигать 10–20%. Это один из главных недостатков первого способа, существенно снижающий точность последующих расчетов годового пробега. Указанный недостаток можно исключить.

Согласно второму способу параметры λ_{03}, λ_{05}, λ_{06} задаются дискретно для каждого интервала пробега и являются постоянными величинами на каждом заданном интервале пробега, составляющем 10–20 тыс. км, но значения этих параметров меняются в течение пробега с начала эксплуатации автомобиля скачкообразно от одного интервала к другому.

Метод динамики средних может быть использован и для определения коэффициента выпуска автопарка, состоящего из автомобилей разных моделей.

Указанная задача может быть решена двумя способами. Первый способ состоит в рассмотрении изолированного процесса эксплуатации совокупности автомобилей одной модели.

Второй способ предполагает рассмотрение процесса функционирования моделей автомобилей многомарочного парка в целом. В этом случае без принципиальных изменений может быть использован изложенный выше способ, разница будет только в том, что число дифференциальных уравнений увеличится в n раз, где n – число моделей подвижного состава, обслуживаемых на одних и тех же постах ТО и ТР. Использование метода динамики средних для определения коэффициентов технической готовности и выпуска автомобилей моделей разномарочного парка[1] позволяет учесть ограниченное количество постов для проведения ТО и ТР.

При определении коэффициентов технической готовности и выпуска автомобилей разномарочного парка необходимо разбить все модели подвижного состава, эксплуатирующегося в АТП, на группы, включающие автомобили тех моделей, которые обслуживаются на одних и тех же постах ТО-2 и ТР. Для каждой группы моделей подвижного состава строится единая система дифференциальных уравнений, описывающая функционирование соответствующей группы автомобилей[1].

Задачи

2.1. В моменты времени t_1, t_2, t_3 проводится осмотр ЭВМ. Возможны следующие состояния ЭВМ:

S_0 – полностью исправна;

S_1 – незначительные неисправности, которые позволяют эксплуатировать ЭВМ;

S_2 – существенные неисправности, дающие возможность решать ограниченное число задач;

S_3 – ЭВМ полностью вышла из строя.

Матрица переходных вероятностей имеет вид

$$\left\|P_{ij}\right\| = \begin{Vmatrix} 0{,}5 & 0{,}3 & 0{,}2 & 0 \\ 0 & 0{,}4 & 0{,}4 & 0{,}2 \\ 0 & 0 & 0{,}3 & 0{,}7 \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

Постройте граф состояний. *Найдите* вероятности состояний ЭВМ после одного, двух, трех осмотров, если вначале (при $t = 0$) ЭВМ была полностью исправна.

2.2. Магазин продает две марки автомобилей А и В. Опыт эксплуатации этих марок автомобилей свидетельствует о том, что для

[1] Примеры расчета коэффициента выпуска и коэффициента технической готовности для различных марок подвижного состава см.: *Бережной В. И., Бережная Е. В.* Методы и модели управления материальными потоками микрологистической системы автопредприятия. – Ставрополь: Интеллект-сервис, 1996.

них имеют место различные матрицы переходных вероятностей, соответствующие состояниям: работает хорошо (состояние 1) и требует ремонта (состояние 2):

автомобиль марки A $P_A = \begin{Vmatrix} 0{,}9 & 0{,}1 \\ 0{,}6 & 0{,}4 \end{Vmatrix}$;

автомобиль марки B $P_B = \begin{Vmatrix} 0{,}8 & 0{,}2 \\ 0{,}7 & 0{,}3 \end{Vmatrix}$.

Элементы матрицы перехода определены на годовой период эксплуатации автомобиля.

Требуется:

1) найти вероятности состояний для каждой марки автомобиля после двухлетней эксплуатации, если в начальном состоянии автомобиль «работает хорошо»;

2) определить марку автомобиля, являющуюся более предпочтительной для приобретения в личное пользование.

2.3. Система S-автомобиль может находиться в одном из пяти возможных состояний:

исправен, работает;

неисправен, ожидает осмотра;

осматривается;

ремонтируется;

списывается.

Постройте граф состояний системы.

2.4. Организация по прокату автомобилей в городе выдает автомобили напрокат в трех пунктах города: А, В, С. Клиенты могут возвращать автомобили в любой из трех пунктов. Анализ процесса возвращения автомобилей из проката в течение года показал, что клиенты возвращают автомобили в пункты А, В, С в соответствии со следующими вероятностями:

Пункты выдачи	Пункты приема автомобилей		
	А	В	С
А	0,8	0,2	0
В	0,2	0	0,8
С	0,2	0,2	0,6

Требуется:

1) в предположении, что число клиентов в городе не изменяется, найти процентное распределение клиентов, возвращающих автомобили по станциям проката к концу года, если в начале года оно было равномерным;

2) найти вероятности состояний в установившемся режиме;

3) определить пункт проката, у которого более целесообразно строить станцию по ремонту автомобилей.

2.5. Рассматривается процесс накопления терминов в динамическом словаре (тезаурусе) при функционировании автоматизированного банка данных (АБД). Сущность процесса в том, что термины заносятся в словарь по мере их появления в той информации, которая вводится в АБД. Например, в АБД автоматизированной системы управления производством (АСУП) могут в качестве терминов заноситься наименования организаций, с которыми данное предприятие поддерживает производственные отношения. Динамический словарь наименований таких организаций будет накапливаться в АБД АСУП по мере появления этих наименований в единицах информации, вводимых в АБД.

В каждой единице информации, поступающей в АБД, в среднем встречается x терминов словаря, а интенсивность поступления единиц информации в АБД равна $\tilde{\lambda}(t)$. Следовательно, интенсивность потока терминов словаря в информации, поступающей в АБД, будет $\lambda(t) = x\tilde{\lambda}(t)$. Предполагается, что поток терминов словаря является пуассоновским. Число терминов словаря n является конечным и неслучайным, хотя, возможно, и неизвестным нам заранее. Все термины словаря могут находиться в единице информации с одинаковой вероятностью. В словарь заносятся, естественно, лишь те термины, которые до сих пор еще не встречались в единицах информации.

Требуется найти математическое ожидание и дисперсию числа терминов, накопленных в динамическом словаре[1].

2.6. Водитель такси обнаружил, что если он находится в городе А, то в среднем в 8 случаях из 10 он везет следующего пассажира в город Б, в остальных случаях будет поездка по городу А. Если же он находится в городе Б, то в среднем в 4 случаях из 10 он везет следующего пассажира в город А, в остальных же случаях будет поездка по городу Б.

Требуется:

1) перечислить возможные состояния процесса и построить граф состояний;
2) записать матрицу переходных вероятностей;
3) найти вероятности состояний после двух шагов процесса, если:
   а) в начальном состоянии водитель находится в городе А;
   б) в начальном состоянии водитель находится в городе Б;
4) найти вероятности состояний в установившемся режиме.

[1] *Вентцель Е. С., Овчаров Л. А.* Прикладные задачи теории вероятностей. – М.: Радио и связь, 1983.

2.7. Система S представляет собой техническое устройство, состоящее из m узлов и время от времени (в моменты $t_1, t_2, ..., t_k$) подвергающееся профилактическому осмотру и ремонту. После каждого шага (момента осмотра и ремонта) система может оказаться в одном из следующих состояний:

S_0 – все узлы исправны (ни один не заменялся новым);

S_1 – один узел заменен новым, остальные исправны;

S_2 – два узла заменены новыми, остальные исправны;

S_i – i узлов ($i < m$) заменены новыми, остальные исправны;

S_m – все m узлов заменены новыми.

Суммарный поток моментов окончания осмотров для всех узлов – пуассоновский с интенсивностью $\lambda = 4$. Вероятность того, что в момент профилактики узел придется заменить новым, равна $P = 0{,}4$.

Рассматривая процесс профилактического осмотра и ремонта (замены) как марковский процесс размножения, *вычислите* вероятности состояний системы (S) в стационарном режиме (для $m = 3$), если в начальный момент времени все узлы исправны[1].

2.8. Техническое устройство имеет два возможных состояния:

S_1 – исправно, работает;

S_2 – неисправно, ремонтируется.

Матрица переходных вероятностей имеет вид:

$$\|P_{ij}\| = \begin{Vmatrix} 0{,}7 & 0{,}3 \\ 0{,}8 & 0{,}2 \end{Vmatrix}.$$

Постройте граф состояний. *Найдите* вероятности состояний после третьего шага и в установившемся режиме, если в начальном состоянии техническое устройство исправно.

2.9. Система S состоит из двух узлов – I и II, каждый из которых может в ходе работы системы отказать (выйти из строя).

Перечислите возможные состояния системы и *постройте* граф состояний для двух случаев:

1) ремонт узлов в процессе работы системы не производится (чистый процесс гибели системы);
2) отказавший узел немедленно начинает восстанавливаться.

2.10. В городе издаются три журнала: C_1, C_2, C_3, и читатели выписывают только один из них. Пусть в среднем читатели стремятся поменять журнал, т. е. подписаться на другой не более одного раза в год, и вероятности таких изменений постоянны. Результаты маркетинговых исследований спроса читателей на журналы дали следующее процентное соотношение:

---

[1] *Вентцель Е. С., Овчаров Л. А.* Прикладные задачи теории вероятностей. – М.: Радио и связь, 1983.

80% читателей C_1 подписываются на C_2;
15% читателей C_2 подписываются на C_3;
8% читателей C_3 подписываются на C_1.

Требуется:

1) записать матрицу переходных вероятностей для средних годовых изменений;
2) предположить, что общее число подписчиков в городе постоянно, и определить, какая доля из их числа будет подписываться на указанные журналы через два года, если по состоянию на 1 января текущего года каждый журнал имел одинаковое число подписчиков;
3) найти вероятности состояний в установившемся режиме и определить журнал, который будет пользоваться наибольшим спросом у читателей.

2.11. Техническое устройство состоит из двух узлов и может находиться в одном из следующих состояний:

оба узла исправны, работают;
неисправен только первый узел;
неисправен только второй узел;
неисправны оба узла.

Вероятность выхода из строя (отказов) после месячной эксплуатации для первого узла – $P_1 = 0{,}4$; для второго узла – $P_2 = 0{,}3$, а вероятность совместного выхода их из строя – $P_{1,2} = 0{,}1$. В исходном состоянии оба узла исправны, работают.

Запишите матрицу переходных вероятностей и *найдите* вероятности состояний после двухмесячной эксплуатации.

2.12. Размеченный граф состояний системы S имеет вид, показанный на рис. 2.14.

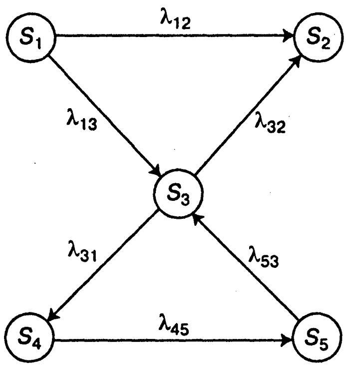


Рис. 2.14. Граф состояний

Запишите систему дифференциальных уравнений Колмогорова и начальные условия для решения системы, если известно, что в начальный момент система находится в состоянии S_1.

2.13. Экономическая система S имеет возможные состояния: S_1, S_2, S_3, S_4. Размеченный граф состояний системы с указанием численных значений интенсивностей перехода показан на рис. 2.15.

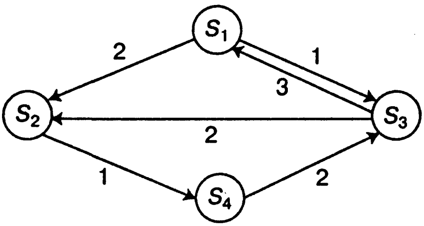


Рис. 2.15. Граф состояний системы

Вычислите вероятности состояний в стационарном режиме.

2.14. Размеченный граф состояний экономической системы с указанием численных значений интенсивностей перехода системы показан на рис. 2.16.

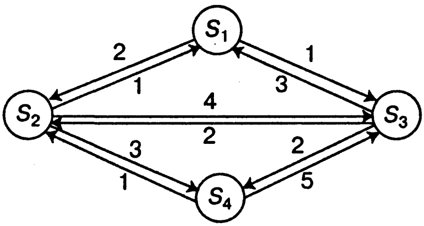


Рис. 2.16. Граф состояний системы

Напишите алгебраические уравнения для вероятностей состояний в установившемся режиме. *Определите* финальные вероятности состояний системы.

2.15. Граф состояний системы имеет вид, приведенный на рис. 2.17.

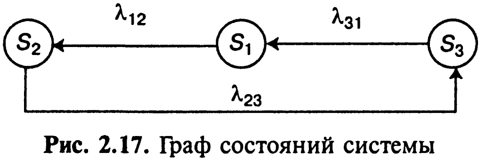


Рис. 2.17. Граф состояний системы

Напишите алгебраические уравнения для вероятностей состояний в стационарном режиме и *найдите* выражение для этих вероятностей.

2.16. *Найдите* вероятности состояний в установившемся режиме для процесса гибели и размножения, граф которого представлен на рис. 2.18.

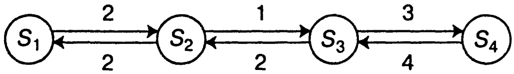


Рис. 2.18. Граф состояний системы

2.17. На автотранспортном предприятии (АТП) эксплуатируются модели автомобилей одной марки. Интенсивность поступления на АТП новых автомобилей $\lambda = 5$ авт/год. Средний срок службы автомобиля до списания $\overline{T}_{\text{сп}} = 7$ лет. Величина $T_{\text{сп}}$ распределена – по показательному закону с параметром $\mu = \dfrac{1}{\overline{T}_{\text{сп}}}$.

Найдите финальные вероятности и математическое ожидание числа эксплуатируемых автомобилей в стационарном режиме, если число автомобилей в АТП не ограничено.

2.18. В задаче 2.17 число эксплуатируемых автомобилей ограничено, $n = 60$ единиц.

Найдите финальные вероятности и математическое ожидание числа эксплуатируемых автомобилей в стационарном режиме на АТП.

2.19. *Найдите* вероятности состояний в стационарном режиме для процесса гибели и размножения, граф которого показан на рис. 2.19.

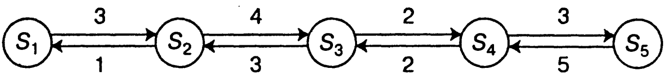


Рис. 2.19. Граф состояний системы

2.20. Система учета на предприятии использует компьютерную сеть, в состав которой входит $n = 6$ персональных компьютеров (ПК). Ежегодно обслуживающий персонал проводит профилактический осмотр каждого ПК. Суммарный поток моментов окончания профилактических осмотров для всего участвующего персонала – пуассоновский с интенсивностью $\lambda = 0{,}5$ ч$^{-1}$ (число событий в единицу времени). После окончания осмотра с вероятностью $P = 0{,}86$ устанавливается, что ПК – работоспособный. Если ПК оказался неработоспособным, то вновь проводится профилактика. В начальный момент все ПК компьютерной сети нуждаются в профилактическом осмотре.

Постройте граф состояний для системы S (6 ПК), *напишите* дифференциальные уравнения для вероятностей состояний. *Найдите* вероятности состояний $P_i(3)$ и математическое ожидание числа персональных компьютеров (M_3), успешно прошедших профилактику после трех часов с начала обслуживания ($t = 3$).

2.21. Используйте условие задачи **2.20**, за исключением того, что система учета предприятия применяет не шесть, а десять персональных компьютеров.

2.22. Размеченный граф состояний в установившемся режиме для процесса гибели и размножения приведен на рис. 2.20.

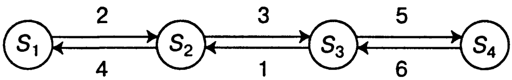


Рис. 2.20. Граф состояний системы

2.23. Граф состояний системы имеет вид, приведенный на рис. 2.21.

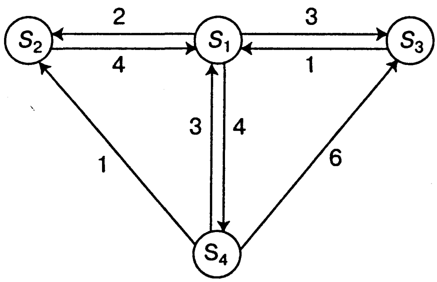


Рис. 2.21. Граф состояний системы

Найдите вероятности состояний системы в стационарном режиме.

2.24. Рассматривается производство персональных компьютеров на заводе. Поток производимых компьютеров – простейший пуассоновский с интенсивностью $\lambda(t) = \lambda = 1200$ год$^{-1}$ (число компьютеров в год).

Определите вероятность выпуска 5000 компьютеров. За четыре года работы завода *вычислите* характеристики процесса производства ПК $m[X(t)]$ и $D[X(t)]$, при $t = 4$ года.

Постройте граф состояний процесса производства ПК.

2.25. Аудиторская фирма разрабатывает проекты отдельных документов для 6 предприятий. Поток разрабатываемых документов – простейший пуассоновский с интенсивностью $\lambda = 2$ месяца$^{-1}$. Определите закон распределения случайного процесса $X(t)$ – число разрабатываемых документов на момент времени $t = 2$ месяца, если в момент $t = 0$ начата разработка документов.

Вычислите математическое ожидание $M[X(t)]$ случайного процесса $X(t)$, предварительно построив размеченный граф состояний.

2.26. Размеченный граф состояний в установившемся режиме для процесса гибели и размножения приведен на рис. 2.22.

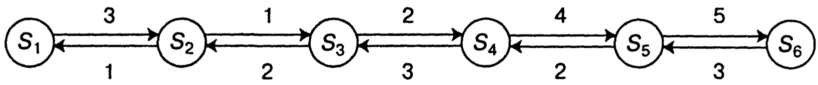


Рис. 2.22. Граф состояний системы

Найдите вероятности состояний.

2.27. Граф состояний системы имеет вид, приведенный на рис. 2.23.

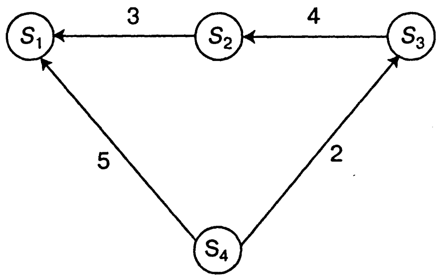


Рис. 2.23. Граф состояний системы

Найдите вероятности состояний в стационарном режиме.

2.28. Размеченный граф состояний представлен на рис. 2.24.

Найдите вероятности состояний S_i и характеристику $M[X(t)]$ на момент времени $t = 3$.

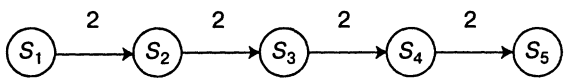


Рис. 2.24. Граф состояний системы

2.29. Размеченный граф состояний представлен на рис. 2.25. *Найдите* вероятности состояний S_i и характеристику $M[X(t)]$ на момент времени $t = 4$.

$$S_1 \xrightarrow{3} S_2 \xrightarrow{3} S_3 \xrightarrow{3} S_4$$

Рис. 2.25. Граф состояний системы

2.30. Размеченный граф состояний представлен на рис. 2.26.

$$S_0 \xrightarrow{5} S_1 \xrightarrow{5} S_2 \xrightarrow{5} S_3 \xrightarrow{5} S_4 \xrightarrow{5} S_5 \xrightarrow{5} \ldots$$

Рис. 2.26. Граф состояний системы

Найдите вероятности состояний S_i и характеристики $M[X(t)]$ и $D[X(t)]$ для $t = 2$.

Глава 3
Моделирование систем массового обслуживания

3.1. Компоненты и классификация моделей массового обслуживания

Рассмотренный в гл. 2 марковский случайный процесс с дискретными состояниями и непрерывным временем имеет место в системах массового обслуживания.

Системы массового обслуживания – это такие системы, в которые в случайные моменты времени поступают заявки на обслуживание, при этом поступившие заявки обслуживаются с помощью имеющихся в распоряжении системы каналов обслуживания.

С позиции моделирования процесса массового обслуживания ситуации, когда образуются очереди заявок (требований) на обслуживание, возникают следующим образом. Поступив в обслуживающую систему, требование присоединяется к очереди других (ранее поступивших) требований. Канал обслуживания выбирает требование из находящихся в очереди, с тем чтобы приступить к его обслуживанию. После завершения процедуры обслуживания очередного требования канал обслуживания приступает к обслуживанию следующего требования, если такое имеется в блоке ожидания. Цикл функционирования системы массового обслуживания подобного рода повторяется многократно в течение всего периода работы обслуживающей системы. При этом предполагается, что переход системы на обслуживание очередного требования после завершения обслуживания предыдущего требования происходит мгновенно, в случайные моменты времени.

Примерами систем массового обслуживания могут служить:

- посты технического обслуживания автомобилей;
- посты ремонта автомобилей;
- персональные компьютеры, обслуживающие поступающие заявки или требования на решение тех или иных задач;
- станции технического обслуживания автомобилей;
- аудиторские фирмы;
- отделы налоговых инспекций, занимающиеся приемкой и проверкой текущей отчетности предприятий;
- телефонные станции и т. д.

Основными компонентами системы массового обслуживания любого вида являются:

- входной поток поступающих требований или заявок на обслуживание;
- дисциплина очереди;
- механизм обслуживания.

Раскроем содержание каждого из указанных выше компонентов.

Для описания *входного потока требований* нужно задать вероятностный закон, определяющий последовательность моментов поступления требований на обслуживание и указать количество таких требований в каждом очередном поступлении. При этом, как правило, оперируют понятием «вероятностное распределение моментов поступления требований». Здесь могут поступать как единичные, так и групповые требования (требования поступают группами в систему). В последнем случае обычно речь идет о системе обслуживания с параллельно-групповым обслуживанием.

Дисциплина очереди – это важный компонент системы массового обслуживания, он определяет принцип, в соответствии с которым поступающие на вход обслуживающей системы требования подключаются из очереди к процедуре обслуживания. Чаще всего используются дисциплины очереди, определяемые следующими правилами:

- первым пришел – первым обслуживаешься;
- пришел последним – обслуживаешься первым;
- случайный отбор заявок;
- отбор заявок по критерию приоритетности;

• ограничение времени ожидания момента наступления обслуживания (имеет место очередь с ограниченным временем ожидания обслуживания, что ассоциируется с понятием «допустимая длина очереди»).

Механизм обслуживания определяется характеристиками самой процедуры обслуживания и структурой обслуживающей системы. К характеристикам процедуры обслуживания относятся: продолжительность процедуры обслуживания и количество требований, удовлетворяемых в результате выполнения каждой такой процедуры. Для аналитического описания характеристик процедуры обслуживания оперируют понятием «вероятностное распределение времени обслуживания требований».

Следует отметить, что время обслуживания заявки зависит от характера самой заявки или требований клиента и от состояния и возможностей обслуживающей системы. В ряде случаев приходится также учитывать вероятность выхода обслуживающего прибора по истечении некоторого ограниченного интервала времени.

Структура обслуживающей системы определяется количеством и взаимным расположением каналов обслуживания (механизмов, приборов и т. п.). Прежде всего следует подчеркнуть, что система обслуживания может иметь не один канал обслуживания, а несколько; система такого рода способна обслуживать одновременно несколько требований. В этом случае все каналы обслуживания предлагают одни и те же услуги, и, следовательно, можно утверждать, что имеет место параллельное обслуживание.

Система обслуживания может состоять из нескольких разнотипных каналов обслуживания, через которые должно пройти каждое обслуживаемое требование, т. е. в обслуживающей системе процедуры обслуживания требований реализуются последовательно. Механизм обслуживания определяет характеристики выходящего (обслуженного) потока требований.

Рассмотрев основные компоненты систем обслуживания, можно констатировать, что функциональные возможности любой системы массового обслуживания определяются следующими основными факторами:

• вероятностным распределением моментов поступлений заявок на обслуживание (единичных или групповых);

• вероятностным распределением времени продолжительности обслуживания;

• конфигурацией обслуживающей системы (параллельное, последовательное или параллельно-последовательное обслуживание);

• количеством и производительностью обслуживающих каналов;

• дисциплиной очереди;

• мощностью источника требований.

В качестве основных критериев эффективности функционирования систем массового обслуживания в зависимости от характера решаемой задачи могут выступать:

- вероятность немедленного обслуживания поступившей заявки;
- вероятность отказа в обслуживании поступившей заявки;
- относительная и абсолютная пропускная способность системы;
- средний процент заявок, получивших отказ в обслуживании;
- среднее время ожидания в очереди;
- средняя длина очереди;
- средний доход от функционирования системы в единицу времени и т. п.

Предметом теории массового обслуживания является установление зависимости между факторами, определяющими функциональные возможности системы массового обслуживания, и эффективностью ее функционирования. В большинстве случаев все параметры, описывающие системы массового обслуживания, являются случайными величинами или функциями, поэтому эти системы относятся к стохастическим системам.

Случайный характер потока заявок (требований), а также, в общем случае, и длительности обслуживания приводит к тому, что в системе массового обслуживания происходит случайный процесс. По характеру случайного процесса, происходящего в системе массового обслуживания (СМО), различают системы марковские и немарковские. В *марковских системах* входящий поток требований и выходящий поток обслуженных требований (заявок) являются пуассоновскими. Пуассоновские потоки позволяют легко описать и построить математическую модель системы массового обслуживания. Данные модели имеют достаточно простые решения, поэтому большинство известных приложений теории массового обслуживания используют марковскую схему. В случае *немарковских процессов* задачи исследования систем массового обслуживания значительно усложняются и требуют применения статистического моделирования, численных методов с использованием ЭВМ.

Независимо от характера процесса, протекающего в системе массового обслуживания, различают два основных вида СМО:

- системы с отказами, в которых заявка, поступившая в систему в момент, когда все каналы заняты, получает отказ и сразу же покидает очередь;
- системы с ожиданием (очередью), в которых заявка, поступившая в момент, когда все каналы обслуживания заняты, становится в очередь и ждет, пока не освободится один из каналов.

Системы массового обслуживания с ожиданием делятся на системы с ограниченным ожиданием и системы с неограниченным ожиданием.

В системах с *ограниченным ожиданием* может ограничиваться:

- длина очереди;
- время пребывания в очереди.

В системах с *неограниченным ожиданием* заявка, стоящая в очереди, ждет обслуживания неограниченно долго, т.е. пока не подойдет очередь.

Все системы массового обслуживания различают по числу каналов обслуживания:

- одноканальные системы;
- многоканальные системы.

Приведенная классификация СМО является условной. На практике чаще всего системы массового обслуживания выступают в качестве смешанных систем. Например, заявки ожидают начала обслуживания до определенного момента, после чего система начинает работать как система с отказами.

3.2. Определение характеристик систем массового обслуживания

Одноканальная модель с пуассоновским входным потоком с экспоненциальным распределением длительности обслуживания

Простейшая одноканальная модель. Такой моделью с вероятностными входным потоком и процедурой обслуживания является модель, характеризуемая показательным распределением как длительностей интервалов между поступлениями требований, так и длительностей обслуживания. При этом плотность распределения длительностей интервалов между поступлениями требований имеет вид

$$f_1(t) = \lambda e^{-\lambda t}, \tag{3.1}$$

где λ – интенсивность поступления заявок в систему.

Плотность распределения длительностей обслуживания:

$$f_2(t) = \mu e^{-\mu t}, \tag{3.2}$$

где μ – интенсивность обслуживания.

Потоки заявок и обслуживаний простейшие.

Пусть система работает с *отказами*. Необходимо определить абсолютную и относительную пропускную способность системы.

Представим данную систему массового обслуживания в виде графа (рис. 3.1), у которого имеются два состояния:

S_0 – канал свободен (ожидание);

S_1 – канал занят (идет обслуживание заявки).

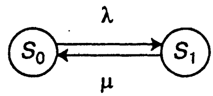


Рис. 3.1. Граф состояний одноканальной СМО с отказами

Обозначим вероятности состояний:

$P_0(t)$ – вероятность состояния «канал свободен»;

$P_1(t)$ – вероятность состояния «канал занят».

По размеченному графу состояний (рис. 3.1) составим систему дифференциальных уравнений Колмогорова для вероятностей состояний:

$$\begin{cases} \dfrac{dP_0(t)}{dt} = -\lambda P_0(t) + \mu P_1(t); \\ \dfrac{dP_1(t)}{dt} = -\mu P_1(t) + \lambda P_0(t). \end{cases} \tag{3.3}$$

Система линейных дифференциальных уравнений (3.3) имеет решение с учетом нормировочного условия $P_0(t) + P_1(t) = 1$. Решение данной системы называется неустановившимся, поскольку оно непосредственно зависит от t и выглядит следующим образом:

$$P_0(t) = \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda+\mu}; \tag{3.4}$$

$$P_1(t) = 1 - P_0(t). \tag{3.5}$$

Нетрудно убедиться, что для одноканальной СМО с отказами вероятность $P_0(t)$ есть не что иное, как относительная пропускная способность системы q.

Действительно, P_0 – вероятность того, что в момент t канал свободен и заявка, пришедшая к моменту t, будет обслужена, а следовательно, для данного момента времени t среднее отношение числа обслуженных заявок к числу поступивших также равно $P_0(t)$, т. е.

$$q = P_0(t). \tag{3.6}$$

По истечении большого интервала времени (при $t \to \infty$) достигается стационарный (установившийся) режим:

$$q = P_0 = \frac{\mu}{\mu+\lambda}. \tag{3.7}$$

Зная относительную пропускную способность, легко найти абсолютную. Абсолютная пропускная способность (A) – среднее число заявок, которое может обслужить система массового обслуживания в единицу времени:

$$A = \lambda q = \frac{\lambda\mu}{\lambda + \mu}. \qquad (3.8)$$

Вероятность отказа в обслуживании заявки будет равна вероятности состояния «канал занят»:

$$P_{\text{отк}} = P_1 = 1 - P_0 = 1 - \frac{\mu}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}. \qquad (3.9)$$

Данная величина $P_{\text{отк}}$ может быть интерпретирована как средняя доля необслуженных заявок среди поданных.

Пример 3.1. Пусть одноканальная СМО с отказами представляет собой один пост ежедневного обслуживания (ЕО) для мойки автомобилей. Заявка – автомобиль, прибывший в момент, когда пост занят, – получает отказ в обслуживании. Интенсивность потока автомобилей λ = 1,0 (автомобиль в час). Средняя продолжительность обслуживания – 1,8 часа. Поток автомобилей и поток обслуживаний являются простейшими.

Требуется определить в установившемся режиме предельные значения:

относительной пропускной способности q;

абсолютной пропускной способности A;

вероятности отказа $P_{\text{отк}}$.

Сравните фактическую пропускную способность СМО с номинальной, которая была бы, если бы каждый автомобиль обслуживался точно 1,8 часа и автомобили следовали один за другим без перерыва.

Решение

1. Определим интенсивность потока обслуживания:

$$\mu = \frac{1}{\bar{t}_{\text{об}}} = \frac{1}{1{,}8} = 0{,}555.$$

2. Вычислим относительную пропускную способность:

$$q = \frac{\mu}{\lambda + \mu} = \frac{0{,}555}{1 + 0{,}555} = 0{,}356.$$

Величина q означает, что в установившемся режиме система будет обслуживать примерно 35% прибывающих на пост ЕО автомобилей.

3. Абсолютную пропускную способность определим по формуле:

$$A = \lambda \cdot q = 1 \cdot 0{,}356 = 0{,}356.$$

Это означает, что система (пост ЕО) способна осуществить в среднем 0,356 обслуживания автомобилей в час.

3. Вероятность отказа:

$$P_{\text{отк}} = 1 - q = 1 - 0{,}356 = 0{,}644.$$

Это означает, что около 65% прибывших автомобилей на пост ЕО получат отказ в обслуживании.

4. Определим номинальную пропускную способность системы:

$$A_{\text{ном}} = \frac{1}{\bar{t}_{\text{обсл}}} = \frac{1}{1{,}8} = 0{,}555 \text{ (автомобилей в час).}$$

Оказывается, что $A_{\text{ном}}$ в 1,5 раза $\left(\frac{0{,}555}{0{,}356} \approx 1{,}5\right)$ больше, чем фактическая пропускная способность, вычисленная с учетом случайного характера потока заявок и времени обслуживания.

Одноканальная СМО с ожиданием. Система массового обслуживания имеет один канал. Входящий поток заявок на обслуживание – простейший поток с интенсивностью λ. Интенсивность потока обслуживания равна μ (т. е. в среднем непрерывно занятый канал будет выдавать μ обслуженных заявок). Длительность обслуживания – случайная величина, подчиненная показательному закону распределения. Поток обслуживаний является простейшим пуассоновским потоком событий. Заявка, поступившая в момент, когда канал занят, становится в очередь и ожидает обслуживания.

Предположим, что независимо от того, сколько требований поступает на вход обслуживающей системы, данная система (очередь + обслуживаемые клиенты) не может вместить более N-требований (заявок), т. е. клиенты, не попавшие в ожидание, вынуждены обслуживаться в другом месте. Наконец, источник, порождающий заявки на обслуживание, имеет неограниченную (бесконечно большую) емкость.

Граф состояний СМО в этом случае имеет вид, показанный на рис. 3.2.

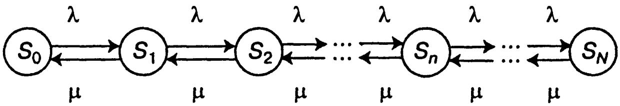


Рис. 3.2. Граф состояний одноканальной СМО с ожиданием (схема гибели и размножения)

Состояния СМО имеют следующую интерпретацию:

S_0 – канал свободен;

S_1 – канал занят (очереди нет);

S_2 – канал занят (одна заявка стоит в очереди);

....

S_n – канал занят (n – 1 заявок стоит в очереди);

....

S_N – канал занят (N – 1 заявок стоит в очереди).

Стационарный процесс в данной системе будет описываться следующей системой алгебраических уравнений:

$$\begin{cases} -\rho P_0 + P_1 = 0, \quad n = 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ -(1-\rho)P_n + P_{n+1} + \rho P_{n-1} = 0, \quad 0 < n < N; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ -P_N + \rho \cdot P_{N-1} = 0, \quad n = N, \end{cases} \tag{3.10}$$

где $\rho = \dfrac{\lambda}{\mu}$;

n – номер состояния.

Решение приведенной выше системы уравнений (3.10) для нашей модели СМО имеет вид

$$P_n = \begin{cases} \left(\dfrac{1-\rho}{1-\rho^{N+1}}\right) \cdot \rho^n, \quad \rho \neq 1, \quad n = 0, 1, 2, \ldots, N \\ \dfrac{1}{(N+1)}, \quad \rho = 1; \end{cases} \tag{3.11}$$

$$P_0 = \frac{1-\rho}{1-\rho^{N+1}}, \tag{3.12}$$

Тогда

$$P_n = \begin{cases} P_0 \cdot \rho^n & \rho \neq 1, \quad n = 1, 2, ..., N, \\ \dfrac{1}{(N+1)}, & \rho = 1. \end{cases}$$

Следует отметить, что выполнение условия стационарности $\rho = \dfrac{\lambda}{\mu} < 1$ для данной СМО необязательно, поскольку число допускаемых в обслуживающую систему заявок контролируется путем введения ограничения на длину очереди (которая не может превышать $N - 1$), а не соотношением между интенсивностями входного потока, т. е. не отношением $\lambda/\mu = \rho$.

Определим *характеристики одноканальной СМО* с ожиданием и ограниченной длиной очереди, равной $(N - 1)$:

вероятность отказа в обслуживании заявки:

$$P_{\text{отк}} = P_N = \begin{cases} \left(\dfrac{1-\rho}{1-\rho^{N+1}}\right)\rho^N, & \rho \neq 1, \\ \dfrac{1}{(N+1)}, & \rho = 1; \end{cases} \tag{3.13}$$

относительная пропускная способность системы:

$$q = 1 - P_{\text{отк}} = \begin{cases} 1 - \left(\dfrac{1-\rho}{1-\rho^{N+1}}\right)\rho^N, & \rho \neq 1, \\ 1 - \dfrac{1}{(N+1)}, & \rho = 1; \end{cases} \tag{3.14}$$

абсолютная пропускная способность:

$$A = q \cdot \lambda; \tag{3.15}$$

среднее число находящихся в системе заявок:

$$L_S = \sum_{n=0}^{N} n \cdot P_n = \begin{cases} \dfrac{\rho \cdot \left[1 - (N+1) \cdot \rho^N + N \cdot \rho^{N+1}\right]}{(1-\rho) \cdot (1-\rho^{N+1})}, & \rho \neq 1 \\ N/2, \quad \rho = 1; \end{cases} \tag{3.16}$$

среднее время пребывания заявки в системе:

$$W_S = \frac{L_S}{\lambda(1-P_N)}; \tag{3.17}$$

средняя продолжительность пребывания клиента (заявки) в очереди:

$$W_q = W_S - 1/\mu; \tag{3.18}$$

среднее число заявок (клиентов) в очереди (длина очереди):

$$L_q = \lambda(1 - P_N)W_q. \tag{3.19}$$

Рассмотрим пример одноканальной СМО с ожиданием.

Пример 3.2. Специализированный пост диагностики представляет собой одноканальную СМО. Число стоянок для автомобилей, ожидающих проведения диагностики, ограничено и равно 3 $[(N - 1) = 3]$. Если все стоянки заняты, т. е. в очереди уже находится три автомобиля, то очередной автомобиль, прибывший на диагностику, в очередь на обслуживание не становится. Поток автомобилей, прибывающих на диагностику, распределен по закону Пуассона и имеет интенсивность $\lambda = 0{,}85$ (автомобиля в час). Время диагностики автомобиля распределено по показательному закону и в среднем равно 1,05 час.

Требуется определить вероятностные характеристики поста диагностики, работающего в стационарном режиме.

Решение

1. Параметр потока обслуживаний автомобилей:

$$\mu = \frac{1}{\bar{t}} = \frac{1}{1{,}05} = 0{,}952.$$

2. Приведенная интенсивность потока автомобилей определяется как отношение интенсивностей λ и μ, т. е.

$$\rho = \frac{\lambda}{\mu} = \frac{0{,}85}{0{,}952} = 0{,}893.$$

3. Вычислим финальные вероятности системы:

$$P_0 = \frac{1-\rho}{1-\rho^{N+1}} = \frac{1-0{,}893}{1-0{,}893^5} \approx 0{,}248;$$

$$P_1 = \rho P_0 = 0,893 \cdot 0,248 \approx 0,221;$$
$$P_2 = \rho^2 P_0 = 0,893^2 \cdot 0,248 \approx 0,198;$$
$$P_3 = \rho^3 P_0 = 0,893^3 \cdot 0,248 \approx 0,177;$$
$$P_4 = \rho^4 P_0 = 0,893^4 \cdot 0,248 \approx 0,158.$$

4. Вероятность отказа в обслуживании автомобиля:

$$P_{\text{отк}} = P_4 = \rho^4 P_0 \approx 0,158.$$

5. Относительная пропускная способность поста диагностики:

$$q = 1 - P_{\text{отк}} = 1 - 0,158 = 0,842.$$

6. Абсолютная пропускная способность поста диагностики

$$A = \lambda \cdot q = 0,85 \cdot 0,842 = 0,716 \text{ (автомобиля в час).}$$

7. Среднее число автомобилей, находящихся на обслуживании и в очереди (т.е. в системе массового обслуживания):

$$L_S = \frac{\rho\left[1-(N+1)\rho^N + N\rho^{N+1}\right]}{(1-\rho)(1-\rho^{N+1})} =$$

$$= \frac{0,893 \cdot \left[1-(4+1) \cdot 0,893^4 + 4 \cdot 0,893^5\right]}{(1-0,893) \cdot (1-0,893^5)} = 1,77.$$

8. Среднее время пребывания автомобиля в системе:

$$W_S = \frac{L_S}{\lambda(1-P_N)} = \frac{1,77}{0,85(1-0,158)} \approx 2,473 \text{ часа.}$$

9. Средняя продолжительность пребывания заявки в очереди на обслуживание:

$$W_q = W_S - 1/\mu = 2,473 - 1/0,952 = 1,423 \text{ часа.}$$

10. Среднее число заявок в очереди (длина очереди):

$$L_q = \lambda(1 - P_N) W_q = 0,85 \cdot (1 - 0,158) \cdot 1,423 = 1,02.$$

Работу рассмотренного поста диагностики можно считать удовлетворительной, так как пост диагностики не обслуживает автомобили в среднем в 15,8% случаев ($P_{\text{отк}} = 0,158$).

Одноканальная СМО с ожиданием без ограничения на вместимость блока ожидания (т. е. $N \to \infty$). Остальные условия функционирования СМО остаются без изменений.

Стационарный режим функционирования данной СМО существует при $t \to \infty$ для любого $n = 0, 1, 2, \ldots$ и когда $\lambda < \mu$. Система алгебраических уравнений, описывающих работу СМО при $t \to \infty$ для любого $n = 0, 1, 2, \ldots$, имеет вид

$$\begin{cases} -\lambda P_0 + \mu P_1 = 0, & n = 0 \\ \lambda P_{n-1} + \mu P_{n+1} - (\lambda + \mu) P_n = 0, & n > 0. \end{cases} \tag{3.20}$$

Решение данной системы уравнений имеет вид

$$P_n = (1 - \rho)\, \rho^n, \quad n = 0, 1, 2, \ldots, \tag{3.21}$$

где $\rho = \lambda/\mu < 1$.

Характеристики одноканальной СМО с ожиданием, без ограничения на длину очереди, следующие:

среднее число находящихся в системе клиентов (заявок) на обслуживание:

$$L_S = \sum_{n=0}^{\infty} n P_n = \frac{\rho}{1-\rho}; \tag{3.22}$$

средняя продолжительность пребывания клиента в системе:

$$W_S = \frac{L_S}{\lambda} = \frac{1}{[\mu(1-\rho)]}; \tag{3.23}$$

среднее число клиентов в очереди на обслуживании:

$$L_q = L_S - \frac{\lambda}{\mu} = \frac{\rho^2}{(1-\rho)}; \tag{3.24}$$

средняя продолжительность пребывания клиента в очереди:

$$W_q = \frac{L_q}{\lambda} = \frac{\rho}{[\mu(1-\rho)]}. \tag{3.25}$$

Пример 3.3. Вспомним о ситуации, рассмотренной в примере 3.2, где речь идет о функционировании поста диагностики. Пусть рассматриваемый пост диагностики располагает неограниченным количеством площадок для стоянки прибывающих на обслуживание автомобилей, т. е. длина очереди не ограничена.

Требуется определить финальные значения следующих вероятностных характеристик:

• вероятности состояний системы (поста диагностики);

• среднее число автомобилей, находящихся в системе (на обслуживании и в очереди);

• среднюю продолжительность пребывания автомобиля в системе (на обслуживании и в очереди);

• среднее число автомобилей в очереди на обслуживании;

• среднюю продолжительность пребывания автомобиля в очереди.

Решение

1. Параметр потока обслуживания μ и приведенная интенсивность потока автомобилей ρ определены в примере 3.2:

$$\mu = 0{,}952; \ \ \rho = 0{,}893.$$

2. Вычислим предельные вероятности системы по формулам

$$P_0 = 1 - \rho = 1 - 0{,}893 = 0{,}107;$$
$$P_1 = (1 - \rho)\,\rho = (1 - 0{,}893) \cdot 0{,}893 = 0{,}096;$$
$$P_2 = (1 - \rho)\,\rho^2 = (1 - 0{,}893) \cdot 0{,}893^2 = 0{,}085;$$
$$P_3 = (1 - \rho)\,\rho^3 = (1 - 0{,}893) \cdot 0{,}893^3 = 0{,}076;$$
$$P_4 = (1 - \rho)\,\rho^4 = (1 - 0{,}893) \cdot 0{,}893^4 = 0{,}068;$$
$$P_5 = (1 - \rho)\,\rho^5 = (1 - 0{,}893) \cdot 0{,}893^5 = 0{,}061 \text{ и т. д.}$$

Следует отметить, что P_0 определяет долю времени, в течение которого пост диагностики вынужденно бездействует (простаивает). В нашем примере она составляет 10,7%, так как $P_0 = 0{,}107$.

3. Среднее число автомобилей, находящихся в системе (на обслуживании и в очереди):

$$L_S = \frac{\rho}{1-\rho} = \frac{0{,}893}{1-0{,}893} = 8{,}346 \text{ ед.}$$

4. Средняя продолжительность пребывания клиента в системе:

$$W_S = \frac{L_S}{\lambda} = \frac{1}{\left[\mu(1-\rho)\right]} = \frac{1}{\left[0{,}952\cdot(1-0{,}893)\right]} = 9{,}817 \text{ час.}$$

5. Среднее число автомобилей в очереди на обслуживание:

$$L_q = L_S - \frac{\lambda}{\mu} = \frac{\rho^2}{(1-\rho)} = \frac{0{,}893^2}{(1-0{,}893)} = 7{,}453.$$

6. Средняя продолжительность пребывания автомобиля в очереди:

$$W_q = \frac{\rho}{\mu(1-\rho)} = \frac{0,893}{0,952(1-0,893)} = 8,766 \text{ час.}$$

7. Относительная пропускная способность системы:

$$q = 1,$$

т. е. каждая заявка, пришедшая в систему, будет обслужена.

8. Абсолютная пропускная способность:

$$A = \lambda q = 0,85 \cdot 1 = 0,85.$$

Следует отметить, что предприятие, осуществляющее диагностику автомобилей, прежде всего интересует количество клиентов, которое посетит пост диагностики при снятии ограничения на длину очереди.

Допустим, в первоначальном варианте количество мест для стоянки прибывающих автомобилей было равно трем (см. пример 3.2). Частота m возникновения ситуаций, когда прибывающий на пост диагностики автомобиль не имеет возможности присоединиться к очереди:

$$m = \lambda\ P_N.$$

В нашем примере при $N = 3 + 1 = 4$ и $\rho = 0,893$,

$$m = \lambda\ P_0\ \rho^4 = 0,85 \cdot 0,248 \cdot 0,8934 = 0,134 \text{ автомобиля в час.}$$

При 12-часовом режиме работы поста диагностики это эквивалентно тому, что пост диагностики в среднем за смену (день) будет терять $12 \cdot 0,134 = 1,6$ автомобиля.

Снятие ограничения на длину очереди позволяет увеличить количество обслуженных клиентов в нашем примере в среднем на 1,6 автомобиля за смену (12 ч. работы) поста диагностики. Ясно, что решение относительно расширения площади для стоянки автомобилей, прибывающих на пост диагностики, должно основываться на оценке экономического ущерба, который обусловлен потерей клиентов при наличии всего трех мест для стоянки этих автомобилей.

Многоканальная модель с пуассоновским входным потоком и экспоненциальным распределением длительности обслуживания

Модели с *n* обслуживающими каналами. В подавляющем большинстве случаев на практике системы массового обслуживания являются многоканальными, и, следовательно, *модели с n обслуживающими каналами* (где $n > 1$) представляют несомненный интерес.

Процесс массового обслуживания, описываемый данной моделью, характеризуется интенсивностью входного потока λ, при этом параллельно может обслуживаться не более n клиентов (заявок). Средняя продолжительность обслуживания одной заявки равняется $1/\mu$. Входной и выходной потоки являются пуассоновскими. Режим функционирования того или иного обслуживающего канала не влияет на режим функционирования других обслуживающих каналов системы, причем длительность процедуры обслуживания каждым из каналов является случайной величиной, подчиненной экспоненциальному закону распределения. Конечная цель использования n параллельно включенных обслуживающих каналов заключается в повышении (по сравнению с одноканальной системой) скорости обслуживания требований за счет обслуживания одновременно n клиентов.

Граф состояний многоканальной системы массового обслуживания с отказами имеет вид, показанный на рис. 3.3.

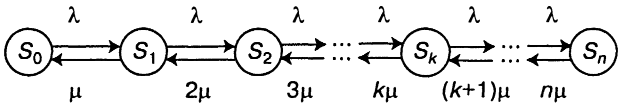


Рис. 3.3. Граф состояний многоканальной СМО с отказами

Состояния данной СМО имеют следующую интерпретацию:

S_0 – все каналы свободны;

S_1 – занят один канал, остальные свободны;

....

S_k – заняты ровно k каналов, остальные свободны;

....

S_n – заняты все n каналов, заявка получает отказ в обслуживании.

Уравнения Колмогорова для вероятностей состояний системы P_0, ..., P_k, ..., P_n будут иметь следующий вид:

$$\begin{cases} \dfrac{dP_0}{dt} = -\lambda \cdot P_0 + \mu \cdot P_1; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dP_k}{dt} = \lambda \cdot P_{k-1} - (\lambda + k \cdot \mu) \cdot P_k + \mu \cdot (k+1) \cdot P_{k+1} \quad 1 \le k \le n-1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \dfrac{dP_n}{dt} = \lambda \cdot P_{n-1} - \mu \cdot n \cdot P_n. \end{cases} \qquad (3.26)$$

Начальные условия решения системы таковы:

$$P_0(0) = 1,\ P_1(0) = P_2(0) = \dots = P_k(0) = \dots = P_n(0) = 0.$$

Стационарное решение системы имеет вид:

$$\begin{cases} P_k = \dfrac{\dfrac{\rho^k}{k!}}{\sum\limits_{k=0}^{n} \dfrac{\rho^k}{k!}} = \dfrac{\rho^k}{k!} \cdot P_0, \quad k = 0, 1, 2, \dots, n \\ P_0 = \dfrac{1}{\left[\sum\limits_{k=0}^{n} \dfrac{\rho^k}{k!} \right]}, \quad k = 0, 1, 2, \dots, n, \end{cases} \tag{3.27}$$

где $\rho = \dfrac{\lambda}{\mu}$.

Формулы для вычисления вероятностей P_k называются *формулами Эрланга.*

Определим вероятностные характеристики функционирования *многоканальной СМО с отказами в стационарном режиме.*

Вероятность отказа определяет формула

$$P_{\text{отк}} = P_n = \frac{\rho^n}{n!} \cdot P_0. \tag{3.28}$$

Заявка получает отказ, если приходит в момент, когда все n каналов заняты. Величина $P_{\text{отк}}$ характеризует полноту обслуживания входящего потока.

Вероятность того, что заявка будет принята к обслуживанию (она же — относительная пропускная способность системы q) дополняет $P_{\text{отк}}$ до единицы:

$$q = 1 - P_{\text{отк}} = 1 - \frac{\rho_n}{n!} \cdot P_0. \tag{3.29}$$

Абсолютная пропускная способность показывается формулой

$$A = \lambda \cdot q = \lambda \cdot (1 - P_{\text{отк}}). \tag{3.30}$$

Среднее число каналов, занятых обслуживанием ($\bar{k}$), следующее:

$$\bar{k} = \sum_{k=1}^{n} k \cdot P_k = \rho \cdot (1 - P_{\text{отк}}). \tag{3.31}$$

Величина $\bar{k}$ характеризует степень загрузки СМО.

Пример 3.4. Пусть n-канальная СМО представляет собой вычислительный центр (ВЦ) с тремя ($n = 3$) взаимозаменяемыми ПЭВМ для решения поступающих задач. Поток задач, поступающих на ВЦ, имеет интенсивность $\lambda = 1$ задаче в час. Средняя продолжительность обслуживания $\bar{t}_{обсл} = 1,8$ час. Поток заявок на решение задач и поток обслуживания этих заявок являются простейшими.

Требуется вычислить финальные значения:

вероятности состояний ВЦ;

вероятности отказа в обслуживании заявки;

относительной пропускной способности ВЦ;

абсолютной пропускной способности ВЦ;

среднего числа занятых ПЭВМ на ВЦ.

Определите, сколько дополнительно надо приобрести ПЭВМ, чтобы увеличить пропускную способность ВЦ в 2 раза.

Решение

1. Определим параметр μ потока обслуживаний:

$$\mu = \frac{1}{\bar{t}_{обсл}} = \frac{1}{1,8} = 0,555.$$

2. Приведенная интенсивность потока заявок равна:

$$\rho = \lambda/\mu = 1/0,555 = 1,8.$$

3. Предельные вероятности состояний найдем по формулам Эрланга (3.27):

$$P_1 = \frac{\rho}{1!} P_0 = 1,8 \cdot P_0;$$

$$P_2 = \frac{\rho^2}{2!} P_0 = 1,62 \cdot P_0;$$

$$P_3 = \frac{\rho^3}{3!} P_0 = 0,97 \cdot P_0;$$

$$P_0 = \frac{1}{\sum\limits_{k=0}^{3} \frac{\rho^k}{k!}} = \frac{1}{1 + 1,8 + 1,62 + 0,97} \approx 0,186;$$

$P_1 \approx 1,8 \cdot 0,186 \approx 0,334;$
$P_2 \approx 1,62 \cdot 0,186 \approx 0,301;$
$P_3 \approx 0,97 \cdot 0,186 \approx 0,180.$

4. Вероятность отказа в обслуживании заявки

$$P_{\text{отк}} = P_3 = 0{,}180.$$

5. Относительная пропускная способность ВЦ

$$q = 1 - P_{\text{отк}} = 1 - 0{,}180 = 0{,}820.$$

6. Абсолютная пропускная способность ВЦ

$$A = \lambda \cdot q = 1 \cdot 0{,}820 = 0{,}820.$$

7. Среднее число занятых каналов – ПЭВМ

$$\bar{k} = \rho \cdot (1 - P_{\text{отк}}) = 1{,}8 \cdot (1 - 0{,}180) = 1{,}476.$$

Таким образом, при установившемся режиме работы СМО в среднем будет занято 1,5 компьютера из трех – остальные полтора будут простаивать. Работу рассмотренного ВЦ вряд ли можно считать удовлетворительной, так как центр не обслуживает заявки в среднем в 18% случаев ($P_3 = 0{,}180$). Очевидно, что пропускную способность ВЦ при данных λ и μ можно увеличить только за счет увеличения числа ПЭВМ.

Определим, сколько нужно использовать ПЭВМ, чтобы сократить число необслуженных заявок, поступающих на ВЦ, в 10 раз, т.е. чтобы вероятность отказа в решении задач не превосходила 0,0180. Для этого используем формулу (3.28):

$$P_{\text{отк}} = \frac{\rho^n}{n!} \cdot P_0.$$

Составим следующую таблицу:

n	1	2	3	4	5	6
P_0	0,357	0,226	0,186	0,172	0,167	0,166
$P_{\text{отк}}$	0,643	0,367	0,18	0,075	0,026	0,0078

Анализируя данные таблицы, следует отметить, что расширение числа каналов ВЦ при данных значениях λ и μ до 6 единиц ПЭВМ позволит обеспечить удовлетворение заявок на решение задач на 99,22%, так как при $n = 6$ вероятность отказа в обслуживании ($P_{\text{отк}}$) составляет 0,0078.

Многоканальная система массового обслуживания с ожиданием. Процесс массового обслуживания с ожиданием характеризуется следующим: входной и выходной потоки являются пуассоновскими с интенсивностями λ и μ соответственно; параллельно могут обслуживаться не более C клиентов. Система имеет C каналов обслуживания. Средняя продолжительность обслуживания одного клиента равна $1/\mu$.

В установившемся режиме функционирование многоканальной СМО с ожиданием и неограниченной очередью может быть описано с помощью системы алгебраических уравнений:

$$\begin{aligned}
&0 = \lambda P_{n-1} - (\lambda + n \cdot \mu) P_n + (n + 1) \mu P_{n+1} \\
&\text{при } 1 \le n < C; \\
&0 = \lambda P_{n-1} - (\lambda + C \cdot \mu) P_n + C \mu P_{n+1} \\
&\text{при } n \ge C.
\end{aligned} \tag{3.32}$$

Решение системы уравнений (3.32) имеет вид:

$$\begin{cases}
P_n = \dfrac{\rho^n}{n!} P_0, & \text{при } 0 \le n < C, \quad (3.33) \\
P_n = \dfrac{\rho^n}{C! C^{n-c}} P_0, & \text{при } n \ge C, \quad (3.34)
\end{cases}$$

где

$$P_0 = \left\{ \sum_{n=0}^{C-1} \frac{\rho^n}{n!} + \frac{\rho^C}{C! \left[1 - \left(\dfrac{\rho}{C} \right) \right]} \right\}^{-1}. \tag{3.35}$$

Решение будет действительным, если выполняется следующее условие: $\left[\dfrac{\lambda}{\mu C} \right] < 1.$

Вероятностные характеристики функционирования в стационарном режиме *многоканальной СМО с ожиданием и неограниченной очередью* определяются по следующим формулам:

вероятность того, что в системе находится n клиентов на обслуживании, определяется по формулам (3.33) и (3.34);

среднее число клиентов в очереди на обслуживание

$$L_q = \left[\frac{C \cdot \rho}{(C - \rho)^2} \right] P_C; \tag{3.36}$$

среднее число находящихся в системе клиентов (заявок на обслуживание и в очереди)

$$L_S = L_q + \rho; \tag{3.37}$$

средняя продолжительность пребывания клиента (заявки на обслуживание) в очереди

$$W_q = \frac{L_q}{\lambda}; \tag{3.38}$$

средняя продолжительность пребывания клиента в системе

$$W_S = W_q + \frac{1}{\mu}. \tag{3.39}$$

Рассмотрим примеры многоканальной системы массового обслуживания с ожиданием.

Пример 3.5. Механическая мастерская завода с тремя постами (каналами) выполняет ремонт малой механизации. Поток неисправных механизмов, прибывающих в мастерскую, – пуассоновский и имеет интенсивность $\lambda = 2{,}5$ механизма в сутки, среднее время ремонта одного механизма распределено по показательному закону и равно $\bar{t} = 0{,}5$ сут. Предположим, что другой мастерской на заводе нет, и, значит, очередь механизмов перед мастерской может расти практически неограниченно.

Требуется вычислить следующие предельные значения вероятностных характеристик системы:

вероятности состояний системы;

среднее число заявок в очереди на обслуживание;

среднее число находящихся в системе заявок;

среднюю продолжительность пребывания заявки в очереди;

среднюю продолжительность пребывания заявки в системе.

Решение

1. Определим параметр потока обслуживаний

$$\mu = \frac{1}{\bar{t}} = 1/0{,}5 = 2.$$

2. Приведенная интенсивность потока заявок

$$\rho = \lambda/\mu = 2{,}5/2{,}0 = 1{,}25,$$

при этом $\lambda/\mu \cdot c = 2{,}5/2 \cdot 3 = 0{,}41$.

Поскольку λ/μ · c <, то очередь не растет безгранично и в системе наступает предельный стационарный режим работы.

3. Вычислим вероятности состояний системы:

$$P_0 = \left\{ \sum_{n=0}^{C-1} \frac{\rho^n}{n!} + \frac{\rho^C}{C!\left[1 - \left(\frac{\rho}{C}\right)\right]} \right\}^{-1} = \frac{1}{\left[1 + \frac{\rho^1}{1!} + \frac{\rho^2}{2!} + \frac{\rho^3}{3!\left(1 - \frac{\rho}{3}\right)}\right]} =$$

$$= \frac{1}{\left[1 + \rho + \frac{\rho^2}{2} + \frac{\rho^3}{6 \cdot (1 - \frac{\rho}{3})}\right]} = \frac{1}{\left[1 + 1{,}25 + \frac{1{,}25^2}{2} + \frac{1{,}25^3}{6 \cdot \left(1 - \frac{1{,}25}{3}\right)}\right]} = 0{,}279;$$

$$P_1 = \frac{\rho^1}{1!} P_0 = 1{,}25 \cdot 0{,}279 = 0{,}349;$$

$$P_2 = \frac{\rho^2}{2!} P_0 = \frac{1{,}25^2}{2!} \cdot 0{,}279 = 0{,}218;$$

$$P_3 = \frac{\rho^3}{3!} P_0 = \frac{1{,}25^3}{3!} \cdot 0{,}279 = 0{,}091;$$

$$P_4 = \frac{\rho^4}{4!} P_0 = \frac{1{,}25^4}{4!} \cdot 0{,}279 = 0{,}028.$$

4. Вероятность отсутствия очереди у мастерской

$$P_{\text{от.о}} = P_0 + P_1 + P_2 + P_3 \approx$$
$$\approx 0{,}279 + 0{,}349 + 0{,}218 + 0{,}091 = 0{,}937.$$

5. Среднее число заявок в очереди на обслуживание

$$L_q = \left[\frac{C\rho}{(C - \rho)^2}\right] P_C = \left[\frac{3 \cdot 1{,}25}{(3 - 1{,}25)^2}\right] \cdot 0{,}091 = 0{,}111.$$

6. Среднее число находящихся в системе заявок

$$L_S = L_q + \rho = 0{,}111 + 1{,}25 = 1{,}361.$$

7. Средняя продолжительность пребывания механизма в очереди на обслуживание

$$W_q = \frac{L_q}{\lambda} = \frac{0{,}111}{2{,}5} = 0{,}044 \text{ суток.}$$

8. Средняя продолжительность пребывания механизма в мастерской (в системе)

$$W_S = W_q + \frac{1}{\mu} = 0{,}044 + \frac{1}{2} = 0{,}544 \text{ суток.}$$

Модель обслуживания машинного парка

Модель обслуживания машинного парка представляет собой *модель замкнутой системы массового обслуживания.*

До сих пор мы рассматривали только такие системы массового обслуживания, для которых интенсивность λ входящего потока заявок не зависит от состояния системы. В этом случае источник заявок является внешним по отношению к СМО и генерирует неограниченный поток требований. Рассмотрим системы массового обслуживания, для которых λ зависит от состояния системы, причем источник требований является внутренним и генерирует ограниченный поток заявок.

Например, обслуживается машинный парк, состоящий из N машин, бригадой R механиков ($N > R$), причем каждая машина может обслуживаться только одним механиком. Здесь машины являются источниками требований (заявок на обслуживание), а механики – обслуживающими каналами. Неисправная машина после обслуживания используется по своему прямому назначению и становится потенциальным источником возникновения требований на обслуживание. Очевидно, что интенсивность λ зависит от того, сколько машин в данный момент находится в эксплуатации ($N - k$) и сколько машин обслуживается или стоит в очереди, ожидая обслуживания (k).

В рассматриваемой модели емкость источника требований следует считать ограниченной. Входящий поток требований исходит из ограниченного числа эксплуатируемых машин ($N - k$), которые в случайные моменты времени выходят из строя и требуют обслуживания. При этом каждая машина из ($N - k$) находится в эксплуатации. Генерирует пуассоновский поток требований с интенсивностью λ независимо от других объектов; общий (суммарный) входящий поток имеет интенсивность $(N - k) \cdot \lambda$. Требование, поступившее в систему в момент, когда свободен хотя бы один канал, немедленно идет на обслуживание. Если требование застает все каналы занятыми обслуживанием других требований, то оно не покидает систему, а становится в очередь и ждет, пока один из каналов не станет свободным.

Таким образом, в замкнутой системе массового обслуживания входящий поток требований формируется из выходящего.

Состояние S_k системы характеризуется общим числом требований, находящихся на обслуживании и в очереди, равным k. Для рассматриваемой замкнутой системы, очевидно, $k = 0, 1, 2, ..., N$. При этом, если система находится в состоянии S_k, то число объектов, находящихся в эксплуатации, равно $(N - k)$.

Если λ – интенсивность потока требований в расчете на одну машину, то

$$\lambda_k = \begin{cases} (N-k)\cdot\lambda & 0 \le k \le N, \\ 0 & k > N, \end{cases}$$

$$\mu_k = \begin{cases} k\cdot\mu, & 0 \le k < R, \\ R\cdot\mu & R \le k \le N, \\ 0 & k > N. \end{cases}$$

Система алгебраических уравнений, описывающих работу замкнутой СМО в стационарном режиме, выглядит следующим образом:

$$\begin{cases} 0 = -\rho N P_0 + P_1; \\ 0 = (N-k+1)\rho P_{k-1} - \left[(N-k)\rho + k\right]P_k + (k+1)P_{k+1} & 0 < k < R, \\ 0 = (N-k+1)\rho P_{k-1} - \left[(N-k)\cdot\rho + R\right]P_k + RP_{k+1} & R \le k < N, \\ 0 = \rho P_{N-1} - RP_N. \end{cases} \tag{3.40}$$

Решая данную систему, находим вероятность k-го состояния:

$$P_k = \begin{cases} \dfrac{N!\cdot\rho^k}{k!\cdot(N-k)!}\cdot P_0 & 1 \le k < R, \\ \dfrac{N!\cdot\rho^k}{R!\cdot R^{k-R}\cdot(N-k)!}\cdot P_0 & R \le k \le N. \end{cases} \tag{3.41}$$

Величина P_0 определяется из условия нормирования $\sum_{k=0}^{N} P_k = 1$ полученных результатов по формулам (3.41) для P_k, $k = 1, 2, ..., N$.

Определим следующие *вероятностные характеристики системы*: среднее число требований в очереди на обслуживание

$$L_q = \sum_{k=R}^{N} (k-R)P_k; \tag{3.42}$$

среднее число требований, находящихся в системе (на обслуживании и в очереди)

$$L_S = \sum_{R=1}^{N} kP_k; \tag{3.43}$$

среднее число механиков (каналов), простаивающих из-за отсутствия работы

$$\overline{R}_n = \sum_{k=0}^{R-1} (R-k)P_k; \tag{3.44}$$

коэффициент простоя обслуживаемого объекта (машины) в очереди

$$\alpha_1 = \frac{L_q}{N}; \tag{3.45}$$

коэффициент использования объектов (машин)

$$\alpha_2 = 1 - \left(\frac{L_S}{N}\right); \tag{3.46}$$

коэффициент простоя обслуживающих каналов (механиков)

$$\alpha_3 = \frac{\overline{R}_n}{R}; \tag{3.47}$$

среднее время ожидания обслуживания (время ожидания обслуживания в очереди)

$$W_q = \frac{1}{\lambda} \cdot \left(\frac{1-\alpha_2}{\alpha_2}\right) - \frac{1}{\mu}. \tag{3.48}$$

Пример 3.6. Пусть для обслуживания десяти персональных компьютеров (ПК) выделено два инженера одинаковой производительности. Поток отказов (неисправностей) одного компьютера – пуассоновский с интенсивностью $\lambda = 0{,}2$. Время обслуживания ПК подчиняется показательному закону. Среднее время обслуживания одного ПК одним инженером составляет: $\bar{t} = 1{,}25$ час.

Возможны следующие варианты организации обслуживания ПК:

- оба инженера обслуживают все десять компьютеров, так что при отказе ПК его обслуживает один из свободных инженеров, в этом случае $R = 2$, $N = 10$;
- каждый из двух инженеров обслуживает по пять закрепленных за ним ПК. В этом случае $R = 1$, $N = 5$.

Необходимо выбрать наилучший вариант организации обслуживания ПК.

Решение

1. Вычислим параметр обслуживания

$$\mu = \frac{1}{\bar{t}} = \frac{1}{1{,}25} = 0{,}8.$$

2. Приведенная интенсивность

$$\rho = \frac{\lambda}{\mu} = \frac{0{,}2}{0{,}8} = 0{,}25.$$

3. Вычислим вероятностные характеристики СМО для двух вариантов организации обслуживания ПК.

Вариант 1

- Определим вероятности состояний системы:

$$P_k = \begin{cases} \dfrac{N!\cdot\rho^k}{k!\cdot(N-k)!}\cdot P_0, & 1 \le k < R \\[2ex] \dfrac{N!\cdot\rho^k}{R!\cdot R^{k-R}\cdot(N-k)!}\cdot P_0 & R \le k \le N; \end{cases}$$

$$P_1 = \frac{10!\cdot 0{,}25^1}{1!\cdot(10-1)!}\cdot P_0 = 2{,}5\cdot P_0;$$

$$P_2 = \frac{10! \cdot 0{,}25^2 \cdot P_0}{2! \cdot 2^{2-2} \cdot (10-2)!} = 2{,}812 \cdot P_0;$$

$$P_3 = \frac{10! \cdot 0{,}25^3}{2! \cdot 2^{3-2} \cdot (10-3)!} \cdot P_0 = 2{,}812 \cdot P_0;$$

$$P_4 = \frac{10! \cdot 0{,}25^4 \cdot P_0}{2! \cdot 2^{4-2} \cdot (10-4)!} = 2{,}461 \cdot P_0;$$

$$P_5 = \frac{10! \cdot 0{,}25^5 \cdot P_0}{2! \cdot 2^{5-2} \cdot (10-5)!} = 1{,}846 \cdot P_0;$$

$$P_6 = \frac{10! \cdot 0{,}25^6 \cdot P_0}{2! \cdot 2^{6-2} \cdot (10-6)!} = 1{,}154 \cdot P_0;$$

$$P_7 = \frac{10! \cdot 0{,}25^7 \cdot P_0}{2! \cdot 2^{7-2} \cdot (10-7)!} = 0{,}577 \cdot P_0;$$

$$P_8 = \frac{10! \cdot 0{,}25^8 \cdot P_0}{2! \cdot 2^{8-2} \cdot (10-8)!} = 0{,}216 \cdot P_0;$$

$$P_9 = \frac{10! \cdot 0{,}25^9 \cdot P_0}{2! \cdot 2^{9-2} \cdot (10-9)!} = 0{,}054 \cdot P_0;$$

$$P_{10} = \frac{10! \cdot 0{,}25^{10} \cdot P_0}{2! \cdot 2^{10-2} \cdot (10-10)!} = 0{,}007 \cdot P_0.$$

• Учитывая, что $\sum\limits_{k=0}^{N} P_k = 1$, и используя результаты расчета P_k, вычислим P_0:

$$\sum_{k=0}^{N} P_k = P_0 + 2{,}5 \cdot P_0 + 2{,}812 \cdot P_0 + 2{,}81 \cdot P_0 + \ldots + 0{,}007 \cdot P_0 = 1.$$

Откуда $P_0 = 0{,}065$,
тогда

$P_1 \approx 0{,}162$; $P_2 \approx 0{,}183$; $P_3 \approx 0{,}182$; $P_4 \approx 0{,}160$;
$P_5 \approx 0{,}11$; $P_6 \approx 0{,}075$; $P_7 \approx 0{,}037$; $P_8 \approx 0{,}014$;
$P_9 \approx 0{,}003$; $P_{10} \approx 0{,}000$.

Определим среднее число компьютеров в очереди на обслуживание:

$$L_q = \sum_{k=R}^{N} (k-R)\cdot P_k =$$

$$= 0 + (3-2)\cdot 0{,}182 + (4-2)\cdot 0{,}160 + (5-2)\cdot 0{,}11 + (6-2)\cdot$$
$$\cdot 0{,}075 + (7-2)\cdot 0{,}037 + (8-2)\cdot 0{,}014 + (9-2)\cdot 0{,}003 =$$
$$= 0{,}182 + 0{,}32 + 0{,}33 + 0{,}3 + 0{,}185 + 0{,}084 + 0{,}021 = 1{,}42.$$

Определим среднее число ПК, находящихся в системе (на обслуживании и в очереди):

$$L_S = \sum_{k=1}^{N} k\cdot P_k =$$

$$= 1\cdot P_1 + 2\cdot P_2 + 3\cdot P_3 + 4\cdot P_4 + 5\cdot P_5 + 6\cdot P_6 + 7\cdot P_7 +$$
$$+ 8\cdot P_8 + 9\cdot P_9 + 10\cdot P_{10} =$$
$$= 0{,}162 + 2\cdot 0{,}183 + 3\cdot 0{,}182 + 4\cdot 0{,}16 + 5\,.\,0{,}11 +$$
$$+ 6\cdot 0{,}075 + 7\cdot 0{,}037 + 8\cdot 0{,}014 + 9\cdot 0{,}003 + 10\cdot 0 = 3{,}11.$$

Определим среднее число инженеров, простаивающих из-за отсутствия работы:

$$\overline{R}_n = \sum_{k=0}^{R-1} (R-k)\cdot P_k = (2-0)\cdot P_0 + (2-1)\cdot P_1 = 2\cdot 0{,}065 + 1\cdot 0{,}162 = 0{,}292.$$

Коэффициент простоя персонального компьютера в очереди следующий:

$$\alpha_1 = \frac{L_q}{N} = \frac{1{,}42}{10} = 0{,}142.$$

Коэффициент использования компьютеров определяется по формуле

$$\alpha_2 = 1 - \left(\frac{L_S}{N}\right) = 1 - \left(\frac{3{,}11}{10}\right) = 0{,}689.$$

Коэффициент простоя обслуживающих инженеров рассчитывается так:

$$\alpha_3 = \frac{\overline{R}_n}{R} = \frac{0{,}292}{2} \approx 0{,}146.$$

Среднее время ожидания ПК обслуживания

$$W_q = \frac{1}{\lambda} \cdot \left[\left(\frac{1-\alpha_2}{\alpha_2} \right) \right] - \frac{1}{\mu} = \frac{1}{0,2} \cdot \frac{1-0,689}{0,689} - \frac{1}{0,8} = 1,01 \text{ час.}$$

Вариант 2

Определим вероятности состояний системы:

$$P_k = \begin{cases} \dfrac{N! \cdot \rho^k}{k! \cdot (N-k)!} \cdot P_0 & 1 \le k < R; \\ \dfrac{N! \cdot \rho^k}{R! \cdot R^{k-R} \cdot (N-k)!} \cdot P_0 & R \le k \le N; \end{cases}$$

$$P_1 = \frac{5! \cdot 0,25}{(5-1)!} \cdot P_0 = 1,25 \cdot P_0;$$

$$P_2 = \frac{5! \cdot 0,25^2}{1! \cdot 1^{2-1} \cdot (5-2)!} \cdot P_0 = 1,25 \cdot P_0;$$

$$P_3 = \frac{5! \cdot 0,25^3}{(5-3)!} \cdot P_0 = 0,938 \cdot P_0;$$

$$P_4 = \frac{5! \cdot 0,25^4}{(5-4)!} \cdot P_0 = 0,469 \cdot P_0;$$

$$P_5 = 5! \cdot 0,25^5 \cdot P_0 = 0,117 \cdot P_0;$$

$$\sum_{k=0}^{5} P_k = P_0 + 1,25 \cdot P_0 + 1,25 \cdot P_0 + 0,938 \cdot P_0 + 0,469 \cdot P_0 + 0,117 \cdot P_0 = 1.$$

Откуда $P_0 = 0,199$,
тогда

$$P_1 \approx 0,249;\ P_2 \approx 0,249;\ P_3 \approx 0,187;\ P_4 \approx 0,093;\ P_5 \approx 0,023.$$

Среднее число компьютеров в очереди на обслуживание таково:

$$L_q = \sum_{k=R}^{N} (k-R) \cdot P_k =$$

$$= (2-1) \cdot 0,249 + (3-1) \cdot 0,187 + (4-1) \cdot 0,093 + (5-1) \cdot 0,023 = 0,994.$$

Среднее число компьютеров, находящихся на обслуживании и в очереди, рассчитывается так:

$$L_S = \sum_{k=1}^{N} k \cdot P_k = P_1 + 2 \cdot P_2 + 3 \cdot P_3 + 4 \cdot P_4 + 5 \cdot P_5 =$$

$$= 0{,}249 + 2 \cdot 0{,}249 + 3 \cdot 0{,}187 + 4 \cdot 0{,}093 + 5 \cdot 0{,}023 = 1{,}8.$$

Среднее число инженеров, простаивающих из-за отсутствия работы:

$$\overline{R}_n = \sum_{k=0}^{R-1} (R-k) \cdot P_k = (1-0) \cdot P_0 = 0{,}199.$$

Коэффициент простоя персонального компьютера в очереди:

$$\alpha_1 = \frac{L_q}{N} = \frac{0{,}994}{5} = 0{,}199.$$

Коэффициент использования компьютеров:

$$\alpha_2 = 1 - \left(\frac{L_S}{N}\right) = 1 - \left(\frac{1{,}8}{5}\right) = 0{,}64.$$

Коэффициент простоя обслуживающих инженеров:

$$\alpha_3 = \frac{\overline{R}_n}{R} = \frac{0{,}199}{1} = 0{,}199.$$

Среднее время ожидания ПК обслуживания:

$$W_q = \left(\frac{1}{\lambda}\right) \cdot \left(\frac{1-\alpha_2}{\alpha_2}\right) - \frac{1}{\mu} = \frac{1}{0{,}2} \cdot \left(\frac{1-0{,}64}{0{,}64}\right) - \frac{1}{0{,}8} = 1{,}56 \text{ час.}$$

Сведем полученные результаты по двум вариантам в следующую таблицу:

Итоговые вероятностные характеристики	Варианты	
	1	2
α_1	0,142	0,199
α_2	0,689	0,64
α_3	0,146	0,199
W_q, час.	1,01	1,56

Таким образом, в варианте 1 каждый компьютер стоит в очереди в ожидании начала его обслуживания приблизительно 0,142 части рабочего времени, что меньше этого показателя при варианте 2 организации работ. Далее в варианте 1 вероятность того, что ПК в любой момент времени будет работать выше, чем в варианте 2, и равна $\alpha^1{}_2 = 0{,}689 > \alpha^2{}_2 = 0{,}64$. Очевидно, вариант 1 организации работ по обслуживанию ПК эффективнее, чем вариант 2.

Задачи

3.1. Одноканальная СМО с отказами представляет собой одну телефонную линию. Заявка (вызов), пришедшая в момент, когда линия занята, получает отказ. Все потоки событий простейшие. Интенсивность потока $\lambda = 0{,}95$ вызова в минуту. Средняя продолжительность разговора $\bar{t} = 1$ мин.

Определите вероятностные характеристики СМО в установившемся режиме работы.

3.2. В одноканальную СМО с отказами поступает простейший поток заявок с интенсивностью $\lambda = 0{,}5$ заявки в минуту. Время обслуживания заявки имеет показательное распределение с $\bar{t} = 1{,}5$ мин.

Определите вероятностные характеристики СМО в установившемся режиме работы.

3.3. В вычислительном центре работает 5 персональных компьютеров (ПК). Простейший поток задач, поступающих на ВЦ, имеет интенсивность $\lambda = 10$ задач в час. Среднее время решения задачи равно 12 мин. Заявка получает отказ, если все ПК заняты.

Найдите вероятностные характеристики системы обслуживания (ВЦ).

3.4. В аудиторскую фирму поступает простейший поток заявок на обслуживание с интенсивностью $\lambda = 1{,}5$ заявки в день. Время обслуживания распределено по показательному закону и равно в среднем трем дням. Аудиторская фирма располагает пятью независимыми бухгалтерами, выполняющими аудиторские проверки (обслуживание заявок). Очередь заявок не ограничена. Дисциплина очереди не регламентирована.

Определите вероятностные характеристики аудиторской фирмы как системы массового обслуживания, работающей в стационарном режиме.

3.5. На пункт техосмотра поступает простейший поток заявок (автомобилей) интенсивности $\lambda = 4$ машины в час. Время осмотра распределено по показательному закону и равно в среднем 17 мин., в очереди может находиться не более 5 автомобилей.

Определите вероятностные характеристики пункта техосмотра в установившемся режиме.

3.6. Используйте условия задачи 3.5 ($\lambda = 4$; $\overline{t} = 17$ мин.). Однако ограничения на очередь сняты.

Вычислите вероятностные характеристики пункта техосмотра в установившемся режиме.

Определите, эффективно ли снятие ограничения на длину очереди.

3.7. На промышленном предприятии решается вопрос о том, сколько потребуется механиков для работы в ремонтном цехе. Пусть предприятие имеет 10 машин, требующих ремонта с учетом числа ремонтирующихся. Отказы машин происходят с частотой $\lambda = 10$ отк/час. Для устранения неисправности механику требуется в среднем $\overline{t} = 3$ мин. Распределение моментов возникновения отказов является пуассоновским, а продолжительность выполнения ремонтных работ распределена экспоненциально. Возможно организовать 4 или 6 рабочих мест в цехе для механиков предприятия.

Необходимо выбрать наиболее эффективный вариант обеспечения ремонтного цеха рабочими местами для механиков.

3.8. В бухгалтерии предприятия имеются два кассира, каждый из которых может обслужить в среднем 30 сотрудников в час. Поток сотрудников, получающих заработную плату, – простейший, с интенсивностью, равной 40 сотрудников в час. Очередь в кассе не ограничена. Дисциплина очереди не регламентирована. Время обслуживания подчинено экспоненциальному закону распределения.

Вычислите вероятностные характеристики СМО в стационарном режиме и определите целесообразность приема третьего кассира на предприятие, работающего с такой же производительностью, как и первые два.

3.9. В инструментальном отделении сборочного цеха работают три кладовщика. В среднем за 1 мин. за инструментом приходят 0,8 рабочего ($\lambda = 0{,}8$). Обслуживание одного рабочего занимает у кладовщика $\overline{t} = 1{,}0$ мин. Очередь не имеет ограничения. Известно, что поток рабочих за инструментом – пуассоновский, а время обслуживания подчинено экспоненциальному закону распределения. Стоимость 1 мин. работы рабочего равна 30 д. е., а кладовщика – 15 д. е.

Найдите средние потери цеха при данной организации обслуживания в инструментальном отделении (стоимость простоя) при стационарном режиме работы.

3.10. Билетная касса работает без перерыва. Билеты продает один кассир. Среднее время обслуживания – 2 мин. на каждого человека. Среднее число пассажиров, желающих приобрести билеты в кассе в течение одного часа, равно $\lambda = 20$ пасс/час. Все потоки в системе простейшие.

Определите среднюю длину очереди, вероятность простоя кассира, среднее время нахождения пассажира в билетной кассе (в очереди и на обслуживании), среднее время ожидания в очереди в условиях стационарного режима работы кассы.

3.11. Пост диагностики автомобилей представляет собой одноканальную СМО с отказами. Заявка на диагностику, поступившая в момент, когда пост занят, получает отказ. Интенсивность потока заявок на диагностику $\lambda = 0{,}5$ автомобиля в час. Средняя продолжительность диагностики $\bar{t} = 1{,}2$ ч. Все потоки событий в системе простейшие.

Определите в установившемся режиме вероятностные характеристики системы.

3.12. Используйте условия задачи **3.11** ($\lambda = 0{,}5$; $\bar{t} = 1{,}2$ час.). Однако вместо одноканальной СМО ($n = 1$) рассматривается трехканальная ($n = 3$), т. е. число постов диагностики автомобилей увеличено до трех.

Найдите вероятностные характеристики СМО в установившемся режиме.

3.13. Автозаправочная станция представляет собой СМО с одним каналом обслуживания и одной колонкой. Площадка при АЗС допускает пребывание в очереди на заправку не более трех автомобилей одновременно. Если в очереди уже находится три автомобиля, очередной автомобиль, прибывший к станции, в очередь не становится, а проезжает мимо. Поток автомобилей, прибывающих для заправки, имеет интенсивность $\lambda = 0{,}7$ автомобиля в минуту. Процесс заправки продолжается в среднем 1,25 мин. Все потоки простейшие.

Определите вероятностные характеристики СМО в стационарном режиме.

3.14. Используйте условия задачи **3.13.** Однако ограничения на длину очереди сняты.

Найдите вероятностные характеристики СМО в стационарном режиме.

Определите, выгодно ли в данной ситуации снятие ограничения на длину очереди в предположении, что дополнительных финансовых ресурсов не требуется для расширения площадки при АЗС.

3.15. На железнодорожную сортировочную горку прибывают составы с интенсивностью $\lambda = 2$ состава в час. Среднее время, в течение которого горка обслуживает состав, равно 0,4 час. Составы, прибывающие в момент, когда горка занята, становятся в очередь и ожидают в парке прибытия, где имеется три запасных пути, на каждом из которых может ожидать один состав. Состав, прибывший в момент, когда все три запасных пути в парке прибытия заняты, становится в очередь на внешний путь. Все потоки событий простейшие.

При установившемся режиме *найдите*:

среднее число составов, ожидающих в очереди (как в парке прибытия, так и вне его);

среднее время ожидания в парке прибытия и на внешних путях;

среднее время ожидания состава в системе обслуживания;

вероятность того, что прибывший состав займет место на внешних путях.

3.16. Рассматривается работа АЗС, на которой имеются три заправочные колонки. Заправка одной машины длится в среднем 3 мин. В среднем на АЗС каждую минуту прибывает машина, нуждающаяся в заправке бензином. Число мест в очереди не ограничено. Все машины, вставшие в очередь на заправку, дожидаются своей очереди. Все потоки в системе простейшие.

Определите вероятностные характеристики работы АЗС в стационарном режиме.

3.17. На станцию технического обслуживания (СТО) автомобилей каждые два часа подъезжает в среднем одна машина. Станция имеет 6 постов обслуживания. Очередь автомобилей, ожидающих обслуживания, не ограничена. Среднее время обслуживания одной машины – 2 часа. Все потоки в системе простейшие.

Определите вероятностные характеристики станции технического обслуживания автомобилей.

3.18. Используйте условия задачи **3.17**, однако на СТО нет возможности организовать стоянку для автомобилей, ожидающих обслуживания. Каждый автомобиль, прибывающий в момент, когда все посты заняты, получает отказ в обслуживании.

Определите вероятностные характеристики СТО автомобилей.

3.19. В вычислительном центре работают 9 персональных компьютеров (ПК). Простейший поток неисправностей имеет интенсивность 0,3 отказа в день. Среднее время устранения одной неисправности одним инженером равно 1,5 час. Компьютеры обслуживают три инженера с одинаковой производительностью. Все потоки событий простейшие. Возможны следующие варианты организации обслуживания ПК:

три инженера обслуживают все 9 компьютеров, так что при отказе ПК его обслуживает один из свободных инженеров, в этом случае $R = 3$; $N = 9$;

каждый из трех инженеров обслуживает по три закрепленных за ним ПК. В этом случае $R = 1$; $N = 3$.

Необходимо *выбрать* наилучший вариант организации обслуживания ПК.

3.20. Малое транспортное предприятие эксплуатирует десять моделей автомобилей одной марки. Простейший поток отказов автомобилей имеет интенсивность $\lambda = 0{,}25$ отказа в день. Среднее время устранения одного отказа автомобиля одним механиком равно 2 час. Все потоки событий простейшие. Возможны два варианта обслуживания:

все автомобили обслуживают два механика с одинаковой производительностью;

все автомобили предприятия обслуживают три механика с одинаковой производительностью.

Необходимо *выбрать* наилучший вариант организации обслуживания автомобилей.

3.21. На вход телефонной станции, имеющей 9 каналов обслуживания, поступает в среднем 120 заявок в час. Заявка получает отказ, если все каналы заняты. Среднее время обслуживания в одном канале равно 4 мин. Все потоки в системе простейшие.

Определите вероятностные характеристики телефонной станции, выступающей в качестве СМО.

3.22. В магазине работает один продавец, который может обслужить в среднем 30 покупателей в час. Поток покупателей простейший с интенсивностью, равной 60 покупателям в час. Все покупатели «нетерпеливые» и уходят, если в очереди стоит 5 человек (помимо обслуживаемых). Все потоки событий простейшие.

Определите следующие вероятностные характеристики магазина для стационарного режима работы:

вероятность обслуживания покупателя;

абсолютную пропускную способность магазина;

среднюю длину очереди;

среднее время ожидания в очереди;

среднее время всего обслуживания;

вероятность простоя продавца.

3.23. Рассматривается работа АЗС, на которой имеется пять заправочных колонок. Заправка одной машины длится в среднем 4 мин. В среднем на АЗС каждую минуту прибывает машина, нуждающаяся в заправке бензином. Число мест в очереди не ограничено. Все машины, вставшие в очередь, дожидаются своей очереди. Все потоки событий простейшие.

Определите вероятностные характеристики АЗС для стационарного режима.

3.24. Имеется двухканальная простейшая СМО с отказами. На ее вход поступает поток заявок с интенсивностью $\lambda = 3$ заявки в час. Среднее время обслуживания одной заявки $\overline{t} = 0{,}5$ час. Каждая обслуженная заявка приносит доход 5 д. е. Содержание канала обходится 3 д. е./час.

Решите, выгодно ли в экономическом отношении увеличить число каналов СМО до трех.

3.25. *Подсчитайте* вероятностные характеристики для простейшей одноканальной СМО с тремя местами в очереди при условиях $\lambda = 4$ заявки/час; $\overline{t} = 0{,}5$ час.

Выясните, как эти характеристики изменятся, если увеличить число мест в очереди до четырех.

3.26. Как изменятся характеристики эффективности СМО в задаче **3.25**, если λ и μ остаются прежними, а ограничение на число мест в очереди снято.

3.27. Одноканальная СМО – ЭВМ, на которую поступают заявки (требования на расчеты). Поток заявок простейший со средним интервалом между заявками $\overline{t} = 10$ мин. Время обслуживания распределено по экспоненциальному закону с математическим ожиданием $\overline{t}_{\text{обсл}} = 8$ мин.

Определите среднее число заявок в СМО, среднее число заявок в очереди, среднее время пребывания заявки в системе и в очереди.

3.28. Система массового обслуживания – билетная касса с тремя окошками (с тремя кассирами) и неограниченной очередью. Пассажиров, желающих купить билет, приходит в среднем 5 человек за 20 мин. Поток пассажиров можно считать простейшим. Кассир в среднем обслуживает трех пассажиров за 10 мин. Время обслуживания подчинено показательному закону распределения.

Определите вероятностные характеристики СМО в стационарном режиме.

3.29. Технические устройства (ТУ) могут время от времени выходить из строя (отказывать). Поток отказов ТУ простейший с интенсивностью $\lambda = 1{,}6$ отказа в сутки. Время восстановления ТУ имеет экспоненциальное распределение. Математическое ожидание времени обслуживания $\overline{t} = 0{,}5$ суток. Количество каналов, выполняющих обслуживание ТУ, равно 5 ед. Количество заявок в очереди не ограничено.

Определите вероятностные характеристики СМО, выполняющие обслуживание ТУ в установившемся режиме.

3.30. Как изменятся вероятностные характеристики СМО задачи **3.29**, если λ и μ остаются прежними, но число каналов обслуживания уменьшится до двух?

Глава 4
Статистическое моделирование экономических систем

4.1. Теоретические основы метода

Метод статистического моделирования (или метод Монте-Карло) – это способ исследования поведения вероятностных систем (экономических, технических и т. д.) в условиях, когда не известны в полной мере внутренние взаимодействия в этих системах.

Этот метод заключается в воспроизведении исследуемого физического процесса при помощи вероятностной математической модели и вычислении характеристик этого процесса. Одно такое воспроизведение функционирования системы называют *реализацией*, или *испытанием*. После каждого испытания регистрируют совокупность параметров, характеризующих случайный исход реализации. Метод основан на многократных испытаниях построенной модели с последующей статистической обработкой полученных данных с целью определения числовых характеристик рассматриваемого процесса в виде статистических оценок его параметров. Процесс моделирования функционирования экономической системы сводится к машинной имитации изучаемого процесса, который как бы копируется на ЭВМ со всеми сопровождающими его случайностями.

Первые сведения о методе Монте-Карло были опубликованы в конце 1940-х гг. Авторами метода являются американские математики Дж. Нейман и С. Улам. В нашей стране первые работы были опубликованы в 1955–1956 гг. В.В. Чавчанидзе, Ю.А. Шрейдером и В.С. Владимировым.

Основным методом статистического моделирования является *закон больших чисел*. Закон больших чисел в теории вероятностей доказывает для различных условий сходимость по вероятности средних значений результатов большого числа наблюдений к некоторым постоянным величинам.

Под законом больших чисел понимают ряд теорем. Например, одна из *теорем П.Л. Чебышева* формулируется так:

> «При неограниченном увеличении числа независимых испытаний n среднее арифметическое свободных от систематических ошибок и равноточных результатов наблюдений ξ_i случайной величины ξ, имеющей конечную дисперсию $D(\xi)$, сходится по вероятности к математическому ожиданию $M(\xi)$ этой случайной величины». Это можно записать в следующем виде:

$$\lim_{n\to\infty} P\left\{\left|\frac{\sum_{i=1}^{n}\xi_i}{n} - M(\xi)\right| < \varepsilon\right\} = 1, \tag{4.1}$$

где ε — сколь угодно малая положительная величина.

Теорема Бернулли формулируется так: «При неограниченном увеличении числа независимых испытаний в одних и тех же условиях частота $P^*(A)$ наступления случайного события A сходится по вероятности к его вероятности P», т. е.

$$\lim_{n\to\infty} P\left\{\left|\frac{m_i^*}{n} - P\right| < \varepsilon\right\} = 1. \tag{4.2}$$

Согласно данной теореме для получения вероятности какого-либо события, например вероятности состояний некоторой системы $P_i(t)$, $i = \overline{0,k}$, вычисляют частоты $P_i^* = \dfrac{m_i^*}{n}$ для одной реализации (испытания), далее проводят подобные вычисления для числа реализаций, равного n. Результаты усредняют и этим самым с некоторым приближением получают искомые вероятности состояний системы. На основе вычисленных вероятностей определяют другие характеристики системы. Следует отметить, что чем больше число реализаций n, тем точнее результаты вычисления искомых величин (вероятностей состояний системы).

Последнее утверждение легко доказать. Предположим, что требуется найти неизвестную величину m. Подберем такую случайную величину ξ, чтобы $M(\xi) = m$ и $D(\xi) = b^2$. Рассмотрим n случайных величин $\xi_1, \xi_2, \xi_3, \ldots, \xi_n$, распределение которых совпадает с распределением ξ. Если n достаточно велико, то согласно центральной предельной теореме распределение суммы $\rho_n = \xi_1 + \xi_2 + \ldots + \xi_n$ будет приближенно нормальным с параметрами $a = nm$; $\sigma^2 = n\,b^2$.

Из правила «трех сигм»

$$P\{a - 3\sigma < \xi < a + 3\sigma\} = 0{,}997 \tag{4.3}$$

следует, что

$$P\left\{nm - 3b\sqrt{n} < \rho_n < nm + 3b\sqrt{n}\right\} = 0{,}997.$$

Разделим неравенство, стоящее в фигурной скобке, на n и получим эквивалентное неравенство с той же вероятностью:

$$P\left\{m-\frac{3b}{\sqrt{n}}<\frac{\rho_n}{n}<m+\frac{3b}{\sqrt{n}}\right\}=0{,}997.$$

Это соотношение можно записать в виде

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-m\right|<\frac{3b}{\sqrt{n}}\right\}=0{,}997. \tag{4.4}$$

Соотношение (4.4) определяет метод расчета m и оценку погрешности. В самом деле, найдем n значений случайной величины ξ. Из выражения (4.4) видно, что среднее арифметическое этих значений будет приближенно равно m. С вероятностью $P = 0{,}997$ ошибка такого приближения не превосходит величины $3b/\sqrt{n}$. Очевидно, эта ошибка стремится к нулю с ростом n, что и требовалось доказать.

Решение любой задачи методом статистического моделирования состоит в следующем:

- разработке и построении структурной схемы процесса, выявлении основных взаимосвязей;
- формальном описании процесса;
- моделировании случайных явлений (случайных событий, случайных величин, случайных функций), сопровождающих функционирование исследуемой системы;
- моделировании (с использованием данных, полученных на предыдущем этапе) функционирования системы – воспроизведении процесса в соответствии с разработанной структурной схемой и формальным описанием;
- накоплении результатов моделирования, их статистической обработке, анализе и обобщении.

В отличие от описанных ранее математических моделей, результаты которых отражали устойчивое во времени поведение системы, результаты, получаемые при статистическом моделировании, подвержены экспериментальным ошибкам. Это означает, что любое утверждение, касающееся характеристик моделируемой системы, должно основываться на результатах соответствующих статистических проверок.

Экспериментальные ошибки при статистическом моделировании в значительной степени зависят от точности моделирования случайных явлений, сопровождающих функционирование исследуемой системы.

Известно, что при изучении вероятностных систем случайные явления могут интерпретироваться в виде случайных событий, случайных величин и случайных функций. Следовательно, моделирование случайных явлений сводится к моделированию случайных событий, случайных величин и случайных функций. Так как случайные события и случайные функции могут быть представлены через случайные величины, то и моделирование случайных событий и случайных функций производится с помощью случайных величин. В связи с этим рассмотрим сначала способы моделирования случайных величин.

Моделирование случайных величин. Для моделирования случайной величины необходимо знать ее закон распределения. Наиболее *общим способом* получения последовательности случайных чисел, распределенных по произвольному закону, является способ, в основе которого лежит их формирование из исходной последовательности случайных чисел, распределенных в интервале [0,1] по равномерному закону.

Равномерно распределенные в интервале [0,1] *последовательности случайных чисел* можно получить тремя способами:

- использованием таблиц случайных чисел;
- применением генераторов случайных чисел;
- методом псевдослучайных чисел.

При решении задачи без применения ЭВМ чаще всего используют таблицы случайных чисел. В таблицах случайных чисел случайные цифры имитируют значения дискретной случайной величины с равномерным распределением:

x_i	0	1	2	3	...	9
p_i	0,1	0,1	0,1	0,1	...	0,1.

При составлении таких таблиц выполняется требование, чтобы каждая из этих цифр от 0; 1;...; 9 встречалась примерно одинаково часто и независимо от других с вероятностью $p_i = 0{,}1$.

Самая большая из опубликованных таблиц случайных чисел содержит 1 000 000 цифр. Таблицы случайных чисел составить не так просто. Они требуют тщательной проверки с помощью специальных статистических тестов.

При решении задач на ЭВМ для выработки случайных чисел, равномерно распределенных в интервале [0,1], могут применяться генераторы случайных чисел. Данные генераторы преобразуют результаты случайного физического процесса в двоичные числа. В качестве случайного физического процесса обычно используют собственные шумы (случайным образом меняющееся напряжение).

Недостатки данного способа получения случайных чисел следующие:

1) трудно проверить качество вырабатываемых чисел;

2) случайные числа не воспроизводимы (если их не запоминать), и, как следствие, нельзя повторить расчет на ЭВМ для исключения случайного сбоя.

Получение псевдослучайных чисел с равномерным законом распределения заключается в выработке псевдослучайных чисел. *Псевдослучайные числа* – это числа, полученные по какой-либо формуле и имитирующие значения случайной величины. Под словом «имитирующие» подразумевается, что эти числа удовлетворяют ряду тестов так, как если бы они были значениями этой случайной величины.

Первый алгоритм для получения псевдослучайных чисел предложил Дж. Нейман. Это так называемый *метод середины квадратов*, который заключается в следующем:

$$\gamma_0 = 0{,}9876,\ \gamma_0^2 = 0{,}97\underline{5353}76,$$
$$\gamma_1 = 0{,}5353,\ \gamma_1^2 = 0{,}28\underline{6546}09,$$
$$\gamma_2 = 0{,}6546 \text{ и т. д.}$$

Алгоритм себя не оправдал: получилось больше, чем нужно, малых значений γ_i – случайных чисел. В настоящее время разработано множество алгоритмов для получения псевдослучайных чисел.

Назовем *достоинства метода псевдослучайных чисел.*

1. На получение каждого случайного числа затрачивается несколько простых операций, так что скорость генерирования случайных чисел имеет тот же порядок, что и скорость работы ЭВМ.
2. Малый объем памяти ЭВМ для программирования.
3. Любое из чисел легко воспроизвести.
4. Качество генерируемых случайных чисел достаточно проверить один раз.

Подавляющее число расчетов по методу Монте-Карло осуществляется с использованием псевдослучайных чисел. От последовательности случайных чисел, равномерно распределенных в интервале [0,1], нетрудно перейти к последовательности случайных чисел с произвольно заданным законом распределения.

Существует основное соотношение, связывающее случайные числа с заданным законом распределения и случайные числа с равномерным законом распределения в интервале [0,1]. Суть его состоит в том, что для преобразования последовательности случайных чисел с равномерным законом распределения в интервале [0,1] в последовательность случайных чисел с заданной функцией распределения $F(x)$ необходимо из совокупности случайных чисел с равномерным законом распределения в интервале [0,1] выбрать случайное число ξ и решить уравнение:

$$F(x) = \xi \tag{4.5}$$

относительно x.

Решение уравнения представляет собой случайное число из совокупности случайных чисел, имеющих функцию распределения $F(x)$.

В случае когда вместо функции распределения $F(x)$ задана плотность вероятности $f(x)$, соотношение (4.5) принимает вид:

$$\int_{-\infty}^{x} f(x)dx = \xi. \tag{4.6}$$

Для ряда законов распределения, наиболее часто встречающихся в реальной экономике, получено аналитическое решение уравнения (4.6), результаты которого приведены в табл. 4.1.

Таблица 4.1

Формулы для моделирования случайных величин

Закон распределения случайной величины	Плотность распределения	Формула для моделирования случайной величины
Экспоненциальный	$f(x) = \lambda e^{-\lambda x}$	$x_i = -\frac{1}{\lambda}\ln\xi_i$
Вейбула	$f(x) = \frac{a}{b}\left(\frac{x}{a}\right)^{a-1}\exp\left[-\left(\frac{x}{b}\right)^{a}\right]$	$x_i = -b\ (\ln \xi_i)^{1/a}$
Гамма-распределение (η – целые числа)	$f(x) = \frac{\lambda^n}{\Gamma(\eta)}e^{-\lambda\cdot x}x^{\eta-1}$	$x_i = -\frac{1}{\lambda}\sum_{j=1}^{\eta}\ln(1-\xi_j)$
Нормальное	$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$	$x_i = \bar{x} + \sigma\left(\sum_{i=1}^{\eta}\xi_i - 6\right)$

Параметры *закона распределения Вейбула* выбираются по таблицам приложения.

Пример 4.1. В результате статистической обработки экспериментальных данных получены следующие значения характеристик случайной величины X: $\bar{x} = 40{,}7$ и $\sigma = 30{,}2$. Установлено, что величина X распределена в соответствии с законом Вейбула.

Определите параметры данного закона.

Решение

1. Вычислим коэффициент вариации случайной величины X:

$$V_X = \frac{\sigma}{\bar{x}} = \frac{30{,}2}{40{,}7} = 0{,}742.$$

2. Исходя из значения коэффициента вариации, определим по таблицам приложения параметры a и C_a. Величины параметров при $V = 0{,}742$ равны $a = 1{,}4$; $C_a = 0{,}659$.

3. Вычислим параметр b по формуле:

$$b = \frac{\sigma}{C_a} = \frac{30{,}2}{0{,}659} = 45{,}8. \tag{4.7}$$

Параметры гамма-распределения вычислим по следующим формулам:

$$\lambda = \frac{\bar{x}}{\sigma^2}; \tag{4.8}$$

$$\eta = \frac{(\bar{x})^2}{\sigma^2}. \tag{4.9}$$

Пример 4.2. Время обслуживания пассажира в кассе Аэрофлота подчинено гамма-распределению. При этом известно среднее значение времени обслуживания $\bar{t} = 42$ мин.; среднее квадратическое отклонение времени равно 14,8 мин.

Вычислите параметры закона распределения.

Решение

1. Вычислим параметр λ:

$$\lambda = \frac{\bar{x}}{\sigma^2} = \frac{42}{(14{,}8)^2} = 0{,}191746.$$

2. Величину параметра η определим по следующей формуле:

$$\eta = \frac{(\bar{x})^2}{\sigma^2} = \frac{42^2}{14{,}8^2} = 8{,}74826 \approx 9.$$

Пример 4.3. Для ПК интенсивность потока отказов λ = 1,2 отк/сутки.

Определите последовательность значений продолжительности интервалов между отказами ПК. Известно, что эти интервалы описываются показательным законом распределения.

Решение

Определим продолжительность интервала между отказами t_i, используя формулу для моделирования случайной величины, распределенной в соответствии с экспоненциальным законом:

$$t_i = x_i = -\frac{1}{\lambda} \cdot \ln \xi_i.$$

Значения ξ_i определим по таблицам случайных чисел.
Допустим $\xi_1 = 0{,}7182$; $\xi_2 = 0{,}4365$; $\xi_3 = 0{,}1548$; $\xi_4 = 0{,}8731$.
Тогда

$$t_1 = -\frac{1}{1{,}2} \ln 0{,}7182 = 6{,}6 \text{ суток};$$

$$t_2 = -\frac{1}{1{,}2} \ln 0{,}4365 = 16{,}6 \text{ суток};$$

$$t_3 = -\frac{1}{1{,}2} \ln 0{,}1548 = 37{,}3 \text{ суток};$$

$$t_4 = -\frac{1}{1{,}2} \ln 0{,}8731 = 2{,}7 \text{ суток и т. д.}$$

Моделирование случайных событий. Моделирование случайного события заключается в воспроизведении факта появления или непоявления случайного события в соответствии с заданной его вероятностью. Моделирование полной группы несовместных событий $A_1, A_2, ..., A_n$, вероятности которых $P(A_i) = P_i$; $i = \overline{1, n}$ известны, можно свести к моделированию дискретной случайной величины Y, имеющей закон распределения

$$P(y_i) = P_i,$$

где вероятности ее возможных значений

$$P(y_i) = P(A_i) = P_i.$$

Очевидно, что принятие в испытании дискретной случайной величиной Y возможного значения y_i равносильно появлению в испытании события A_i. При практической реализации данного способа на единичном отрезке числовой оси откладывают интервалы $\Delta i = P_i$ (рис. 4.1).

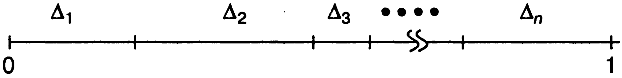


Рис. 4.1. Интервалы $\Delta_i = P_i$

Вырабатывают равномерно распределенное на интервале [0,1] случайное число ξ_j и проверяют условие

$$\sum_{i=1}^{k-1} P_i \le \xi_j < \sum_{i=1}^{k} P_i. \tag{4.10}$$

При выполнении условия (4.10) считают, что при испытании наступило событие A_k.

Нетрудно заметить, что моделирование факта появления одного события A, имеющего вероятность $P(A)$, сводится к моделированию полной группы двух несовместных событий, т. е. противоположных событий с вероятностями $P(A)$ и $P(\bar{A}) = 1 - P(A)$.

Пример 4.4. Вероятность появления события A в каждом испытании $P(A) = 0{,}75$.

Смоделируйте три испытания и *определите* последовательность реализации события A.

Решение

Отложим на единичном отрезке числовой оси точку $E = 0{,}75$ и будем считать, что если случайное число $\xi_i < E$, то в испытании наступило событие A. В противном случае при $\xi_i \ge E$ наступило событие $\bar{A}$, т. е. событие A не имело места.

Пусть из таблицы выбраны равномерно распределенные на интервале [0,1] случайные числа $\xi_1 = 0{,}925$; $\xi_2 = 0{,}135$; $\xi_3 = 0{,}088$. Тогда при трех испытаниях получим следующую последовательность реализации событий: $\bar{A}$; A; A.

Моделирование совместных (зависимых и независимых) событий можно выполнить двумя способами.

Первый способ. На первом этапе моделирования определяют все возможные исходы появления совместных событий в испытании (находят полную группу несовместных событий и вычисляют их вероятности). На последующем этапе работ поступают так же, как и при моделировании полной группы несовместных событий.

Пример 4.5. Пусть при испытании могут иметь место зависимые и совместные события A и B, при этом известно, что $P(A) = 0{,}7$; $P(B) = 0{,}5$; $P(AB) = 0{,}3$.

Смоделируйте появление событий A и B в двух испытаниях.

Решение

При каждом испытании возможны четыре несовместных исхода, т. е. наступление четырех событий:

1. $C_1 = AB$, при этом по условию $P(C_1) = P(AB) = 0{,}3$.
2. $C_2 = A\overline{B}$, при этом $P(C_2) = P(A\overline{B}) = P(A) - P(BA) =$
$= 0{,}7 - 0{,}3 = 0{,}4$.
3. $C_3 = A\overline{B}$, при этом $P(C_3) = P(A\overline{B}) = P(B) - P(AB) =$
$= 0{,}5 - 0{,}3 = 0{,}2$.
4. $C_4 = \overline{A}\,\overline{B}$, при этом $P(C_4) = 1 - [P(C_1) + P(C_2) + P(C_3)] =$
$= 1 - (0{,}3 + 0{,}4 + 0{,}2) = 0{,}1$.

Смоделируем полную группу событий C_1, C_2, C_3, C_4 в двух испытаниях. Предварительно на единичном отрезке числовой оси (рис. 4.2) откладываем интервалы $\Delta i = P(C_i)$, $i = \overline{1,4}$.

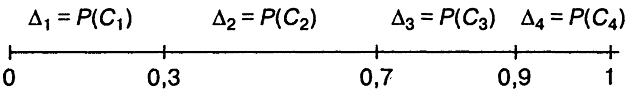


Рис. 4.2. Интервалы $\Delta_i = P(C_i)$

Пусть получены (взяты из таблицы) случайные числа $\xi_1 = 0{,}68$ и $\xi_2 = 0{,}95$. Случайное число ξ_1 принадлежит интервалу Δ_2, поэтому при первом испытании имело место событие A, а событие B не наступило. При втором испытании случайное число ξ_2 принадлежит интервалу Δ_4. Оба события A и B не имели места.

Второй способ. Моделирование совместных событий состоит в разыгрывании факта появления каждого из совместных событий отдельно, при этом, если события зависимые, необходимо предварительно определить условные вероятности.

Пример 4.6. Используя условия примера 4.5, смоделируйте раздельное появление событий A и B в одном испытании.

Решение

События A и B зависимы, поэтому предварительно находим условные вероятности $P(B/A)$ и $P(B/\bar{A})$:

$$P\left(\frac{B}{A}\right)=\frac{P(AB)}{P(A)}=\frac{0,3}{0,7}=\frac{3}{7};$$

$$P\left(\frac{B}{\bar{A}}\right)=\frac{P(\bar{A}B)}{P(\bar{A})}=\frac{0,2}{1-0,7}=\frac{P(B)-P(AB)}{1-P(A)}=\frac{0,5-0,3}{1-0,7}=\frac{0,2}{0,3}=\frac{2}{3}.$$

Для моделирования события A выработаем случайное число ξ_1. Пусть $\xi_1 = 0,96$, так как $\xi_1 > P(A)$. Событие A в испытании не наступило.

Теперь разыграем событие B при условии, что событие A в испытании не имело место. Пусть случайное число $\xi_2 = 0,22$, тогда, $\xi_2 < P(B/\bar{A})$, т. е. $0,22<\frac{2}{3}$. Событие B при испытании наступило.

Понятие о моделировании случайных функций. Для моделирования случайных функций используют два способа. В первом из них применяются специальные физические датчики, вырабатывающие непрерывные реализации случайной функции. Физические датчики с помощью специальных фильтров преобразуют собственные шумы в случайные функции с заданными характеристиками.

В основе второго способа моделирования случайных функций лежит использование случайных чисел. При этом получают значения реализации моделируемой случайной функции в изолированных точках. Сущность способа состоит в том, что воспроизведение реализации случайной функции сводится к моделированию системы коррелированных случайных величин.

4.2. Моделирование систем массового обслуживания с использованием метода Монте-Карло

Рассмотренные в гл. 3 аналитические методы анализа СМО исходят из предположения, что входящие и исходящие потоки требований являются простейшими. Зависимости, используемые в этих методах для определения показателей качества обслуживания, справедливы лишь для установившегося режима функционирования СМО. Однако в реальных условиях функционирования СМО имеются переходные режимы, а входящие и исходящие потоки требований являются далеко не простейшими. В этих условиях для оценки качества функционирования систем обслуживания широко используют *метод статистических испытаний (метод Монте-Карло).* Основой решения задачи исследования функционирования СМО в реальных условиях является статистическое моделирование входящего потока требований и процесса их обслуживания (исходящего потока требований).

Для решения задачи статистического моделирования функционирования СМО должны быть заданы следующие исходные данные:

- описание СМО (тип, параметры, критерии эффективности работы системы);
- параметры закона распределения периодичности поступления требований в систему;
- параметры закона распределения времени пребывания требования в очереди (для СМО с ожиданием);
- параметры закона распределения времени обслуживания требований в системе.

Решение задачи статистического моделирования функционирования СМО складывается из следующих этапов.

1. Вырабатывают равномерно распределенное случайное число ξ_i.

2. Равномерно распределенные случайные числа преобразуют в величины с заданным законом распределения:

- интервал времени между поступлениями требований в систему (Δt_{Ti});
- время ухода заявки из очереди (для СМО с ограниченной длиной очереди);
- длительность времени обслуживания требования каналами (Δt_{Oi}).

3. Определяют моменты наступления событий:

- поступление требования на обслуживание;
- уход требования из очереди;
- окончание обслуживания требования в каналах системы.

4. Моделируют функционирование СМО в целом и накапливают статистические данные о процессе обслуживания.

5. Устанавливают новый момент поступления требования в систему, и вычислительная процедура повторяется в соответствии с изложенным.

6. Определяют показатели качества функционирования СМО путем обработки результатов моделирования методами математической статистики.

Методику решения задачи рассмотрим на *примере моделирования СМО с отказами.*

Пусть система имеет два однотипных канала, работающих с отказами, причем моменты времени окончания обслуживания на первом канале обозначим через t_{1i}, на втором канале – через t_{2i}. Закон распределения интервала времени между смежными поступающими требованиями задан плотностью распределения $f_1(t_T)$. Продолжительность обслуживания также является случайной величиной с плотностью распределения $f_2(t_0)$.

Процедура решения задачи будет выглядеть следующим образом:

1. Вырабатывают равномерно распределенное случайное число ξ_i.

2. Равномерно распределенное случайное число преобразуют в величины с заданным законом распределения, используя формулы табл. 4.1. Определяют реализацию случайного интервала времени (Δt_{Ti}) между поступлениями требований в систему.

3. Вычисляют момент поступления заявки на обслуживание: $t_i = t_{i-1} + \Delta t_{Ti}$.

4. Сравнивают моменты окончания обслуживания предшествующих заявок на первом $t_{1(i-1)}$ и втором $t_{(2i-1)}$ каналах.

5. Сравнивают момент поступления заявки t_i с минимальным моментом окончания обслуживания (допустим, что $t_{1(i-1)} < t_{2(i-1)}$):

а) если $[t_i - t_{1(i-1)}] < 0$, то заявка получает отказ и вырабатывают новый момент поступления заявки описанным способом;

б) если $[t_i - t_{1(i-1)}] \geq 0$, то происходит обслуживание.

6. При выполнении условия 5 б) определяют время обслуживания i-й заявки на первом канале Δt_{1i} путем преобразования случайной величины ξ_i в величину (время обслуживания i-й заявки) с заданным законом распределения.

7. Вычисляют момент окончания обслуживания i-й заявки на первом канале $t_{1i} = [t_{1(i-1)} + \Delta t_{1i}]$.

8. Устанавливают новый момент поступления заявки, и вычислительная процедура повторяется в соответствии с изложенным.

9. В ходе моделирования СМО накапливаются статистические данные о процессе обслуживания.

10. Определяют показатели качества функционирования системы путем обработки накопленных результатов моделирования методами математической статистики.

4.3. Моделирование потоков отказов элементов сложных технических систем

Под *сложной технической системой* будем понимать систему, состоящую из элементов (два и более). Отказ одного из элементов системы приводит к отказу системы в целом.

Рассмотрим последовательность замен некоторого определенного элемента Z данного наименования. Эксплуатация каждого нового элемента начинается с момента окончания срока службы предыдущего. Первый элемент отрабатывает время Δt_1, второй – Δt_2, третий – Δt_3 и т. д.

Случайная ситуация, сложившаяся в k-м опыте (ситуации) для элемента Z, показана на рис. 4.3.

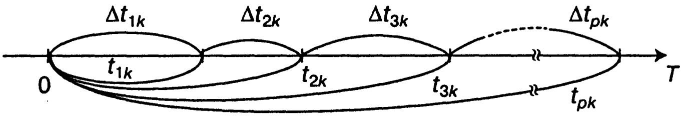


Рис. 4.3. Временная эпюра случайной ситуации при k-м опыте в случае мгновенного восстановления отказавшей системы путем замены элемента

На рис. 4.3 видно, что система начинает свою работу в момент времени $t = 0$ и, отработав случайное время Δt_{1k}, выходит из строя в момент $t_{1k} = \Delta t_{1k}$. В этот момент система мгновенно восстанавливается[1] (элемент заменяется) и снова работает случайное время Δt_{2k}. По истечении некоторого времени система (элемент) вновь выходит из строя в момент $t_{2k} = \Delta t_{1k} + \Delta t_{2k} = t_{1k} + \Delta t_{2k}$ и вновь мгновенно восстанавливается.

Считают, что интервалы времени между отказами Δt_{1k}, Δt_{2k}, ..., Δt_{pk} представляют собой систему взаимно независимых случайных величин с плотностями распределения наработок между отказами $f(\Delta t_1)$, $f(\Delta t_2)$ $f(\Delta t_3)$, ..., $f(\Delta t_p)$.

Моменты отказов или восстановлений образуют в каждом k-м опыте (испытании) ряд чисел по следующему правилу:

$$\begin{aligned}
&t_{1k} = \Delta t_{1k} \\
&t_{2k} = \Delta t_{1k} + \Delta t_{2k} = t_{1k} + \Delta t_{2k} \\
&t_{3k} = \Delta t_{1k} + \Delta t_{2k} + \Delta t_{3k} = t_{2k} + \Delta t_{3k} \qquad (4.11) \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&t_{pk} = \Delta t_{1k} + \Delta t_{2k} + \dots + \Delta t_{pk} = t_{(p-1)k} + \Delta t_{pk}
\end{aligned}$$

[1] Здесь можно считать, что восстановление происходит мгновенно.

или

$$t_{pk} = \sum_{i=1}^{p} \Delta t_{ik} = \sum_{i=1}^{p-1} \Delta t_{ik} + \Delta t_{pk}, \tag{4.12}$$

где t_{ik} – время работы (наработка) элемента до i-го отказа в k-м опыте, час, $i = \overline{1, p}$, $k = \overline{1, N}$;

Δt_{ik} – время работы (наработка) элемента между $(i-1)$-м и i-м отказами в k-й реализации, час, $i = \overline{1, p}$, $k = \overline{1, N}$.

Числа t_{1k}, t_{2k}, ..., t_{pk} образуют случайный поток, который называется *процессом восстановления.* Этот процесс является различным для различных элементов и продолжается до окончания срока службы системы. Изучением таких процессов занимается *теория восстановления.*

Из большого количества различных процессов восстановления для исследования надежности элементов технической системы (как неремонтируемых, так и ремонтируемых) используют три типа процессов:

- простой, при котором все функции распределения наработок до первого и между последующими отказами $F_i(t)$ равны;
- общий, при котором вид функции распределения наработки до первого отказа элемента, установленного в системе заводом-изготовителем, отличается от вида функций распределения наработок элементов при последующих заменах, т. е. $F_1(t) \neq F_i(t)$, $i = 2, 3, 4, ...$;
- сложный, при котором все функции распределения $F_i(t)$ различны.

Основной характеристикой процесса восстановления является функция восстановления $\Omega(t)$ и ее дифференциальная характеристика – плотность восстановления $\omega(t)$, определяемые по следующим формулам:

$$\Omega(t) = \sum_{n=1}^{\infty} F_n(t); \tag{4.13}$$

$$\omega(t) = \sum_{n=1}^{\infty} f_n(t), \tag{4.14}$$

где $f_n(t)$ и $F_n(t)$ – соответственно плотность и функция распределения наработки до n-го отказа.

В случае независимости наработок между отказами функции распределения $F_n(t)$ наработок до n-го отказа находятся путем последовательного применения *правила свертки для суммы двух случайных величин:*

$$F_n(t) = F_{n-1}(t) \cdot F(\Delta t_n) = \int_0^t F_{n-1}(t-\Delta t) \cdot dF(\Delta t); \qquad (4.15)$$

$$F_1(t) = F(t).$$

Анализ и классификация методов расчета параметра и ведущей функции потока отказов приведены в книге авторов[1]. Следует отметить, что сложность получения аналитических выражений для $\Omega(t)$ и $\omega(t)$ по формулам (4.13), (4.14) состоит в том, что свертка (4.15) лишь для некоторых законов распределения вычисляется в конечном виде. Использование аналитических методов расчета плотности $\omega(t)$ и функции восстановления $\Omega(t)$ ограничено из-за сложности математической формализации применяемых стратегий восстановления работоспособности технических систем и необходимости учета множества факторов, влияющих на замену элемента в системе. В этих условиях наиболее эффективным методом расчета $\Omega(t)$ и $\omega(t)$ является метод Монте-Карло.

Расчет ведущей функции и параметра потока отказов этим методом в случае простого, общего или сложного процессов производится в следующем порядке.

По известным законам распределения наработок элементов с использованием формул преобразования (табл.4.1) моделируются массивы случайных величин Δt_{ik} между $(i - 1)$-м и i-м отказами. Размерность каждого массива равна N.

Далее вычисляются значения наработок до i-го отказа t_{ik} по следующим формулам:

$$t_{ik} = t_{(i-1)k} + \Delta t_{ik}; \qquad (4.16)$$

$$t_{1k} = \Delta t_{1k}, \qquad (4.17)$$

где i – номер отказа, $i = \overline{1, p}$;

k – номер реализации при моделировании, $k = \overline{1, N}$;

p – максимальное число отказов элемента, получаемое в k-й реализации случайного процесса.

Затем полученные случайные величины наработок t_{ik} группируются по интервалам времени.

Номера интервалов, в которые попадают моменты возникновения отказов $t_{1k}, t_{2k}, \ldots, t_{ik}, \ldots, t_{pk}$, определяются по формуле

[1] *Бережной В. И., Бережная Е. В.* Методы и модели управления материальными потоками микрологистической системы автопредприятия. – Ставрополь: Интеллект-сервис, 1996.

$$\gamma = CEIL\left(\frac{t_{ik}}{\Delta t}\right), \tag{4.18}$$

где $CEIL\left(\frac{t_{ik}}{\Delta t}\right)$ – наименьшее целое число, не меньшее $\left(\frac{t_{ik}}{\Delta t}\right)$;

Δt – величина интервала времени.

Параметр и ведущая функция потока отказов в j-м интервале времени определяются по следующим формулам:

$$\omega_j(t) = \sum_{i=1}^{p} n_{ij} \Big/ \Delta t \cdot N = \frac{n_j}{\Delta t \cdot N}; \tag{4.19}$$

$$\Omega_j(t) = \sum_{j=1}^{h} \sum_{i=1}^{p} n_{ij} = \frac{S_j}{N}, \tag{4.20}$$

где n_{ij} – число попаданий случайной наработки до i-го отказа t_{ik} в j-й интервал времени ($j = \overline{1, h}$) за N реализаций.

$$n_j = \sum_{i=1}^{p} n_{ij}; \tag{4.21}$$

$$\sum_{j=1}^{h} \sum_{i=1}^{p} n_{ij} = \begin{cases} S_1 = n_1 \\ S_2 = n_1 + n_2 \\ \dots\dots\dots\dots\dots\dots \\ S_h = n_1 + n_2 + \ldots + n_j + \ldots + n_h \end{cases}, \tag{4.22}$$

где h – максимальное число интервалов времени.

Методика расчета параметра $\omega(t)$ и ведущей функции нестационарного потока отказов с использованием метода статистических испытаний подробно рассмотрена в книге авторов[1].

Пример 4.7. Законы распределения наработок элемента системы до первого и второго отказов и соответствующие параметры этих законов приведены в табл. 4.2.

---

[1] *Бережной В. И., Бережная Е. В.* Методы и модели управления материальными потоками микрологистической системы автопредприятия. – Ставрополь: Интеллект-сервис, 1996.

Таблица 4.2

Законы распределения наработок

№ отказа	Закон распределения	Параметры закона	
		$a(\lambda)$	b
1	Вейбула	1,4	45,8
2	Экспоненциальный	0,30	–

Определите номера временных интервалов, на которых произойдут первый и второй отказы в ходе первого опыта (испытания) (Δt = 1 час).

Решение

1. Выберем равномерно распределенное случайное число. Допустим $\xi_1 = 0{,}725$.

2. Вычислим случайные значения наработок на отказ элемента, используя формулы табл. 4.1.

$t_{11} = \Delta t_{11} = -\, b \cdot (\ln\xi_1)^{1/a} = -45{,}8 \cdot (\ln 0{,}725)^{1/1{,}4} = 20{,}37$ час.;
$t_{21} = t_{11} + \Delta t_{21}$;

$$\Delta t_{21} = -\frac{1}{\lambda} \cdot \ln\xi_1 = -\frac{1}{0{,}3} \cdot \ln 0{,}725 = 1{,}07 \text{ час.};$$

$t_{21} = t_{11} + \Delta t_{21} = 20{,}37 + 1{,}07 = 21{,}44$ час.

3. Определим номер временного интервала, на котором произойдут отказы

$$\gamma = CEIL\left(\frac{t_{ik}}{\Delta t}\right);$$

первый отказ

$$\gamma = CEIL\left(\frac{t_{11}}{\Delta t}\right) = CEIL\left(\frac{20{,}37}{1}\right) = 21;$$

второй отказ

$$\gamma = CEIL\left(\frac{t_{21}}{\Delta t}\right) = CEIL\left(\frac{21{,}44}{1}\right) = 22.$$

В ходе первой реализации элемент системы первый раз откажет на 21-м временном интервале, а второй отказ произойдет на 22-м временном интервале.

Задачи

4.1. Периодичность поступления заявок на обслуживание подчинена показательному закону распределения. Средний интервал между поступлениями заявок в систему равен $\bar{t} = 2$ час.

Определите последовательность значений продолжительности интервалов между поступлениями заявок. Число реализаций равно 10.

4.2. Время обслуживания работника предприятия кассой бухгалтерии является случайной величиной, распределенной в соответствии с законом Вейбула. Среднее время обслуживания $\bar{t} = 3$ мин., среднее квадратическое отклонение равно $\sigma = 2$ мин.

Требуется смоделировать случайную величину, отвечающую этим условиям. Число реализаций принять равным 10.

4.3. При обработке экспериментальных данных было установлено, что время, расходуемое на станции технического обслуживания автомобилей для замены двигателя, распределено по нормальному закону, параметры которого $\bar{x} = 2{,}8$ час. на один двигатель и $\sigma_{\bar{x}} = 0{,}6$ час.

Требуется смоделировать для отмеченных условий случайную величину – время X, расходуемое для замены двигателя. Число реализаций принять равным 5.

4.4. Время проверки приемки квартального отчета инспектором налоговой службы (t) величина случайная, распределенная в соответствии с законом Вейбула. Среднее время проверки и приемки равно $\bar{t} = 20$ мин. Коэффициент вариации величины t равен $V_t = 0{,}52$.

Требуется смоделировать для заданных условий случайные числа t (число реализаций принять равным 10).

4.5. Среднее число исправных станков в токарном цехе на заводе равно $\bar{x} = 6$. Среднее квадратическое отклонение $\sigma_{\bar{x}} = 2{,}2$.

Требуется смоделировать число исправных станков в цехе (число реализаций равно 5) при условии, что случайная величина X имеет гамма-распределение.

4.6. Вероятность замены неисправной детали на новую при ремонте автомобиля в каждом испытании $p = 0{,}63$.

Смоделировать пять испытаний и *определить* последовательность замены детали на новую или восстановленную.

4.7. При испытании могут иметь место зависимые и совместные три события: работает только первый кассир по выдаче заработной платы, работает только второй кассир, работают оба кассира. При этом известно, что вероятность работы первого кассира равна 0,6; вероятность работы второго кассира равна 0,7; вероятность работы двух кассиров равна 0,4.

Смоделировать возможность реализации двух событий: работает только первый кассир; работает только второй кассир в трех испытаниях.

4.8. Известны законы распределения наработок элемента системы до первого и второго отказов. Средние значения и средние квадратические отклонения наработок приведены в следующей таблице:

№ отказа	Закон распределения	Среднее значение наработки, час.	Среднее квадратическое отклонение наработки, час.
1	Вейбула	50	28
2	Гамма-распределение	40	15

Определите параметр и ведущую функцию потока отказов элемента по интервалам времени (Δt = 10 час.). Число реализаций $N = 10$.

4.9. Используя условия задачи 4.8, *определите* номер интервала, в который попадет максимальное количество отказов.

4.10. Система имеет два элемента. Средняя периодичность первого элемента $\overline{t}_1$ = 60 час., второго элемента – $\overline{t}_2$ = 85 час. Периодичности отказа первого и второго элементов – случайные величины, подчиненные экспоненциальному закону распределения.

Определите параметр и функцию распределения потока отказов системы по интервалам времени Δt = 8 час. Число реализаций $N = 10$.

4.11. Используя условия задачи 4.10, *определите* номера интервалов, в которые попадут максимальные количества отказов первого, второго элементов и в целом всей системы.

4.12. Пусть при испытании могут иметь место зависимые и совместные события A и D. Известно, что вероятности появления событий равны $P(A)$ = 0,6; $P(D)$ = 0,3, а также вероятность совместного появления событий A и D: $P(AD)$ = 0,4.

Смоделируйте появление событий A и D в пяти испытаниях.

4.13. Периодичность проверки предприятий налоговой инспекции – величина случайная (Δt), подчиняющаяся закону гамма-распределения. Средний интервал проверки $\overline{\Delta t}$ = 2,5 мес. Коэффициент вариации величины Δt равен V = 0,38.

Требуется смоделировать для заданных условий возможные моменты проверок предприятия налоговой инспекцией (число реализаций принять равным 10).

4.14. Используя условия задачи 4.13, *определите* количество проверок налоговой инспекцией за первый год работы предприятия.

4.15. Среднее число работающих машин на заводе $\bar{x} = 25$. Коэффициент вариации числа работающих $V = 0{,}6$.

Требуется смоделировать число работающих машин на заводе (число реализаций равно 10). Случайная величина X имеет распределение Вейбула.

4.16. После каждой проверки предприятия налоговой инспекцией вероятность появления необходимости аудиторской проверки данного предприятия $P = 0{,}72$.

Смоделируйте шесть испытаний.

Определите последовательность проведения различных проверок предприятия.

Глава 5
Методы и модели корреляционно-регрессионного анализа

5.1. Общие сведения

Большинство явлений и процессов в экономике находятся в постоянной взаимной и всеохватывающей объективной связи. Исследование зависимостей и взаимосвязей между объективно существующими явлениями и процессами играет большую роль в экономике. Оно дает возможность глубже понять сложный механизм причинно-следственных отношений между явлениями. Для исследования интенсивности, вида и формы зависимостей широко применяется корреляционно-регрессионный анализ, который является методическим инструментарием при решении задач прогнозирования, планирования и анализа хозяйственной деятельности предприятий.

Различают два вида зависимостей между экономическими явлениями и процессами:

- функциональную;
- стохастическую (вероятностную, статистическую).

В случае функциональной зависимости имеется однозначное отображение множества A на множество B. Множество A называют областью определения функции, а множество B – множеством значений функции.

Функциональная зависимость встречается редко. В большинстве случаев функция (Y) или аргумент (X) – случайные величины. X и Y подвержены действию различных случайных факторов, среди которых могут быть факторы, общие для двух случайных величин.

Если на случайную величину X действуют факторы Z_1, Z_2, ..., V_1, V_2, а на Y – Z_0, Z_2, V_1, V_3 ..., то наличие двух общих факторов

Z_2 и V_1 позволит говорить о вероятностной или статистической зависимости между X и Y.

Определение. *Статистической* называется зависимость между случайными величинами, при которой изменение одной из величин влечет за собой изменение закона распределения другой величины.

В частном случае статистическая зависимость проявляется в том, что при изменении одной из величин изменяется математическое ожидание другой. В этом случае говорят о корреляции или корреляционной зависимости.

Статистическая зависимость проявляется только в массовом процессе, при большом числе единиц совокупности.

При стохастической закономерности для заданных значений зависимой переменной можно указать ряд значений объясняющей переменной, случайно рассеянных в интервале. Каждому фиксированному значению аргумента соответствует определенное статистическое распределение значений функции. Это обусловливается тем, что зависимая переменная, кроме выделенной переменной, подвержена влиянию ряда неконтролируемых или неучтенных факторов. Поскольку значения зависимой переменной подвержены случайному разбросу, они не могут быть предсказаны с достаточной точностью, а только указаны с определенной вероятностью.

В экономике приходится иметь дело со многими явлениями, имеющими вероятностный характер. Например, к числу случайных величин можно отнести стоимость продукции, доходы предприятия, межремонтный пробег автомобилей, время ремонта оборудования и т. д.

Односторонняя вероятностная зависимость между случайными величинами есть регрессия. Она устанавливает соответствие между этими величинами.

Односторонняя стохастическая зависимость выражается с помощью функции, которая называется *регрессией*.

Виды регрессий.

1. Регрессия относительно числа переменных:

• простая регрессия – регрессия между двумя переменными;

• множественная регрессия – регрессия между зависимой переменной y и несколькими объясняющими переменными $x_1, x_2, ..., x_m$. Множественная линейная регрессия имеет следующий вид:

$$y = a_0 + a_1x_1 + a_2x_2 + ... + a_mx_m, \tag{5.1}$$

где y – функция регрессии;
$x_1, x_2, ..., x_m$ – независимые переменные;
$a_1, a_2, ..., a_m$ – коэффициенты регрессии;
a_0 – свободный член уравнения;
m – число факторов, включаемых в модель.

2. Регрессия относительно формы зависимости:

• линейная регрессия, выражаемая линейной функцией;

• нелинейная регрессия, выражаемая нелинейной функцией.

3. В зависимости от характера регрессии различаются следующие ее виды:

• положительная регрессия: она имеет место, если с увеличением (уменьшением) объясняющей переменной значения зависимой переменной также соответственно увеличиваются (уменьшаются);

• отрицательная регрессия: в этом случае с увеличением или уменьшением объясняющей переменной зависимая переменная уменьшается или увеличивается.

4. Относительно типа соединения явлений различаются:

• непосредственная регрессия: в этом случае зависимая и объясняющая переменные связаны непосредственно друг с другом;

• косвенная регрессия: в этом случае объясняющая переменная действует на зависимую через ряд других переменных;

• ложная регрессия: она возникает при формальном подходе к исследуемым явлениям без уяснения того, какие причины обусловливают данную связь.

Регрессия тесно связана с корреляцией. *Корреляция* в широком смысле слова означает связь, соотношение между объективно существующими явлениями. Связи между явлениями могут быть различны по силе. При измерении тесноты связи говорят о корреляции в узком смысле слова. Если случайные переменные причинно обусловлены и можно в вероятностном смысле высказаться об их связи, то имеется корреляция.

Понятия «корреляция» и «регрессия» тесно связаны между собой. В корреляционном анализе оценивается сила связи, а в регрессионном анализе исследуется ее форма. Корреляция в широком смысле объединяет корреляцию в узком смысле и регрессию.

Корреляция, как и регрессия, имеет различные виды, так различают:

1) относительно характера –

• положительную;

• отрицательную;

2) относительно числа переменных –

• простую;

• множественную;

• частную;

3) относительно формы связи –

• линейную;

• нелинейную;

4) относительно типа соединения –

• непосредственную;
• косвенную;
• ложную.

Любое причинное влияние может выражаться либо функциональной, либо корреляционной связью. Но не каждая функция или корреляция соответствует причинной зависимости между явлениями. Поэтому требуется обязательное исследование причинно-следственных связей.

Исследование корреляционных связей мы называем *корреляционным анализом*, а исследование односторонних стохастических зависимостей – *регрессионным анализом*. Корреляционный и регрессионный анализ имеют свои задачи.

Задачи корреляционного анализа.

1. Измерение степени связности (тесноты, силы) двух и более явлений. Здесь речь идет в основном о подтверждении уже известных связей.

2. Отбор факторов, оказывающих наиболее существенное влияние на результативный признак на основе измерения тесноты связи между явлениями.

3. Обнаружение неизвестных причинных связей. Корреляция непосредственно не выявляет причинных связей между явлениями, но устанавливает степень необходимости этих связей и достоверность суждений об их наличии. Причинный характер связей выясняется с помощью логически-профессиональных рассуждений, раскрывающих механизм связей.

Задачи регрессионного анализа.

1. Установление формы зависимости (линейная или нелинейная; положительная или отрицательная и т. д.).

2. Определение функции регрессии и установление влияния факторов на зависимую переменную. Важно не только определить форму регрессии, указать общую тенденцию изменения зависимой переменной, но и выяснить, каково было бы действие на зависимую переменную главных факторов, если бы прочие не изменялись и если бы были исключены случайные элементы. Для этого определяют функцию регрессии в виде математического уравнения того или иного типа.

3. Оценка неизвестных значений зависимой переменной, т. е. решение задач экстраполяции и интерполяции. В ходе экстраполяции распространяются тенденции, установленные в прошлом, на будущий период. Экстраполяция широко используется в прогнозировании. В ходе интерполяции определяют недостающие значения,

соответствующие моментам времени между известными моментами, т. е. определяют значения зависимой переменной внутри интервала заданных значений факторов.

Рассмотрим подробнее регрессию.

Выборочные уравнения регрессии

Условное математическое ожидание случайной величины Y: $M(Y/X)$ есть функция от X, которая называется *функцией регрессии* и равна $f(x)$, т. е.

$$M(Y/X) = f(x); \tag{5.2}$$

аналогично

$$M(X/Y) = \varphi(y). \tag{5.3}$$

Графическое изображение $f(x)$ или $\varphi(y)$ называется *линией регрессии*, а записанные уравнения (5.2) и (5.3) – *уравнениями регрессии.*

Поскольку условное математическое ожидание M случайной величины Y есть функция от (x), то его оценка $\bar{y}$, т. е. условная средняя, также является функцией от X. Обозначим эту функцию через

$$\bar{y}_x = f^*(x). \tag{5.4}$$

Уравнение (5.4) определяет выборочное уравнение регрессии y на x. Сама функция $f^*(x)$ называется *выборочной регрессией* Y на X, а график $f^*(x)$ – *выборочной регрессией*. Аналогично определяется для случайных величин X:

$$\bar{x}_y = \varphi^*(y). \tag{5.5}$$

Функция регрессии необратима, так как речь идет о средних величинах для некоторого конкретного значения фактора.

Функция регрессии формально устанавливает соответствие между переменными X и Y, хотя такой зависимости может и не быть в экономике (ложная регрессия).

Линейная регрессия

Пусть задана система случайных величин X и Y и случайные величины X и Y зависимы.

Представим одну из случайных величин как линейную функцию другой случайной величины X:

$$Y = g(x) = \alpha + \beta x, \tag{5.6}$$

где α, β – параметры, которые подлежат определению.

В общем случае эти параметры могут быть определены различными способами, наиболее часто используется метод наименьших квадратов (МНК).

Функцию $g(x)$ называют *наилучшим приближением* в смысле МНК, если математическое ожидание $M[Y - g(x)]^2$ принимает наименьшее возможное значение.

В этом случае функцию $g(x)$ называют *средней квадратической регрессией* Y на X. Можно доказать, что линейная средняя квадратическая регрессия имеет вид:

$$g(x) = \alpha + \beta x = m_y - r\frac{\sigma_y}{\sigma_x}m_x + r\frac{\sigma_y}{\sigma_x}x, \tag{5.7}$$

где m_x, m_y – математические ожидания случайных величин X, Y соответственно;

σ_x, σ_y – средние квадратические отклонения случайных величин X, Y соответственно;

r – коэффициент парной корреляции, который определяется по формуле

$$r = \frac{M_{xy}}{\sigma_x \sigma_y}, \tag{5.8}$$

где M_{xy} – ковариация.

$$M_{xy} = M[(X - m_x) \cdot (Y - m_y)], \tag{5.9}$$

тогда $\beta = r\frac{\sigma_y}{\sigma_x}$ – коэффициент регрессии. Возникает проблема определения параметров α и β на основе выборки.

Рассмотрим определение параметров выбранного уравнения прямой линии средней квадратической регрессии по несгруппированным данным. Пусть изучается система количественных признаков (X, Y), т. е. ведутся наблюдения за двухмерной случайной величиной (X, Y). Пусть в результате n наблюдений получено n пар чисел (x_1, y_1), (x_2, y_2), ..., (x_n, y_n).

Требуется по полученным данным найти выборочное уравнение прямой линии средней квадратической регрессии:

$$\overline{y}_x = kx + b.$$

Поскольку данные несгруппированные, т. е. каждая пара чисел встречается один раз, то можно перейти от условной средней к переменной y. Угловой коэффициент k обозначим через $k = \rho$ и назовем его выборочной оценкой коэффициента регрессии $\beta = r\frac{\sigma_y}{\sigma_x}$.

Итак, требуется найти:

$$y = \rho x + b. \qquad (5.10)$$

Очевидно, параметры ρ и b нужно подобрать так, чтобы точки (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), построенные по исходным данным, лежали как можно ближе к прямой (5.10) (рис. 5.1).

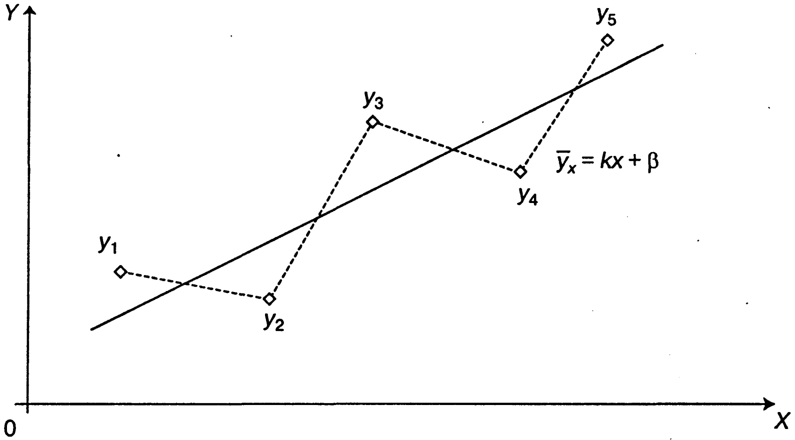


Рис. 5.1. Динамика изменения признака Y

Уточним смысл этого требования. Для этого введем следующее понятие. Назовем отклонением разность вида:

$$Y_i - y_i (i = 1, 2, ..., n),$$

где Y_i – вычисляется по уравнению (5.10) и соответствует наблюдаемому значению x_i;

y_i – наблюдаемая ордината, соответствующая x_i.

Подберем параметры ρ и b так, чтобы сумма квадратов указанных отклонений была наименьшей:

$$\sum_{i=1}^{n} (Y_i - y_i)^2 \to \min.$$

В этом состоит требование метода наименьших квадратов (МНК).

Эта сумма есть функция F отыскиваемых параметров ρ и b:

$$F(\rho, b) = \sum_{i=1}^{n} (Y_i - y_i)^2$$

или

$$F(\rho,b)=\sum_{i=1}^{n}(\rho x_i+b-y_i)^2.$$

Для отыскания min найдем частные производные и приравняем их к нулю:

$$\begin{cases}\dfrac{dF}{d\rho}=2\sum\limits_{i=1}^{n}(\rho x_i+b-y_i)\cdot x_i=0,\\ \dfrac{dF}{db}=2\sum\limits_{i=1}^{n}(\rho x_i+b-y_i)=0.\end{cases}$$

Далее запишем систему:

$$\begin{cases}\left(\sum\limits_{i=1}^{n}x_i^2\right)\rho+\left(\sum\limits_{i=1}^{n}x_i\right)b-\sum\limits_{i=1}^{n}x_iy_i=0,\\ \left(\sum\limits_{i=1}^{n}x_i\right)\rho+nb-\sum\limits_{i=1}^{n}y_i=0.\end{cases}$$

Для простоты вместо $\sum\limits_{i=1}^{n}x_i, \sum\limits_{i=1}^{n}x_i^2, \sum\limits_{i=1}^{n}x_iy_i, \sum\limits_{i=1}^{n}y_i$ будем писать Σx, Σx^2, Σxy, Σy (индекс i опускаем), тогда:

$$\begin{cases}\left(\sum x^2\right)\rho+\left(\sum x\right)b=\sum xy,\\ \left(\sum x\right)\rho+nb=\sum y.\end{cases}$$

Получили систему двух линейных уравнений относительно ρ и b. Решая эту систему, получим:

$$\rho=\frac{n\sum xy-\sum x\sum y}{x^2-\left(\sum x\right)^2}, \tag{5.11}$$

$$b=\frac{\sum y\sum x^2-\sum x\sum xy}{n\sum x^2-\left(\sum x\right)^2}. \tag{5.12}$$

Метод наименьших квадратов применяется и для нахождения параметров множественной регрессии. В этом случае число линейных уравнений возрастает, и такие системы уравнений решаются с помощью ЭВМ.

Основные понятия корреляционно-регрессионного анализа

1. Среднее значение переменной определяется по следующей формуле

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \tag{5.13}$$

где x_i – эмпирическое значение переменной x;
n – число наблюдений.

2. Дисперсия

$$\sigma_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}. \tag{5.14}$$

3. Ковариация

$$\mathrm{Cov}_{xy} = \sum_{i=1}^{n} \left[(x_i - \bar{x}) \cdot (y_i - \bar{y})\right]. \tag{5.15}$$

4. Коэффициент корреляции

$$r_{xy} = \frac{\sum_{i=1}^{n} \left[(x_i - \bar{x}) \cdot (y_i - \bar{y})\right]}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \cdot \sum_{i=1}^{n} (y_i - \bar{y})^2}}. \tag{5.16}$$

Коэффициент корреляции характеризует тесноту, или силу связи между переменными y и x. Значения, принимаемые r_{xy}, заключены в пределах от – 1 до + 1. При положительном значении r_{xy} имеет место положительная корреляция, т. е. с увеличением (уменьшением) значений одной переменной (x) значение другой (y) соответственно увеличивается (уменьшается). При отрицательном значении r_{xy} имеет место отрицательная корреляция, т. е. с увеличением (уменьшением) значений x значения y соответственно уменьшаются (увеличиваются). При изучении экономического явления, зависящего от многих факторов, строится множественная регрессионная зависимость. В этом случае для характеристики тесноты связи используется коэффициент множественной корреляции:

$$R=\sqrt{1-\frac{\sigma^2_{ост}}{\sigma^2_{общ}}},\qquad(5.17)$$

где $\sigma^2_{ост}$ – остаточная дисперсия зависимой переменной;
$\sigma^2_{общ}$ – общая дисперсия зависимой переменной.

5. Общая дисперсия определяется по формуле

$$\sigma^2_{общ}=\frac{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}{n-1}.\qquad(5.18)$$

Величина $\sigma^2_{общ}$ характеризует разброс наблюдений фактических значений от среднего значения $\bar{y}$.

6. Остаточная дисперсия определяется по следующей формуле

$$\sigma^2_{ост}=\frac{\sum_{i=1}^{n}\left(y_i-y_{iт}\right)^2}{n-1},\qquad(5.19)$$

где $y_{iт}$ – теоретические значения переменной y, полученные по уравнению регрессии (5.1) при подстановке в него наблюдаемых фактических значений x_i.

Остаточная дисперсия характеризует ту часть рассеяния переменной y, которая возникает из-за всякого рода случайностей и влияния неучтенных факторов.

7. Коэффициент детерминации служит для оценки точности регрессии, т. е. соответствия полученного уравнения регрессии имеющимся эмпирическим данным, и вычисляется по формуле

$$Д=1-\frac{\sigma^2_{ост}}{\sigma^2_{общ}}.\qquad(5.20)$$

Изменяется Д в пределах от 0 до 1, т. е.

$$0\le Д\le 1.$$

Модель считается тем точнее, чем ближе Д к 1, т. е. чем меньше $\sigma^2_{ост}$.

Стандартная ошибка оценки равна $\sqrt{\sigma^2_{ост}}$.

Если Д = 0, это значит отношение $\frac{\sigma^2_{ост}}{\sigma^2_{общ}}=1$, т. е. $\sigma^2_{ост}=\sigma^2_{общ}$, и, следовательно, $y_{iт}=\bar{y}$. В этом случае прямая регрессии будет параллельна оси X, корреляционно-регрессионная связь между X и Y отсутствует. Если Д = 1, значит, $\frac{\sigma^2_{ост}}{\sigma^2_{общ}} = 0$, т. е. $\sigma^2_{ост} = 0$. Отсюда $y_i = y_{iт}$, т. е. все наблюдаемые точки лежат на построенной прямой, следовательно, зависимость функциональная.

8. Корреляционное отношение используется для оценки тесноты связи между двумя явлениями, в частности для определения тесноты связи исходного ряда y_i с теоретическим рядом $y_{iт}$. Корреляционное отношение определяют по данным, сгруппированным по объясняющей переменной по следующей формуле

$$\eta = \frac{\sum_{i=1}^{n} (y_{iт} - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}. \tag{5.21}$$

5.2. Исходные предпосылки регрессионного анализа и свойства оценок

Применение метода наименьших квадратов для определения параметров регрессии предполагает выполнение некоторых предпосылок[1].

Отметим наиболее существенные из них.

Предпосылка 1. При нахождении оценок переменной y предполагается существование зависимости переменной y только от тех объясняющих переменных, которые вошли в модель (регрессию). Влияние прочих факторов и случайностей учитывается случайной возмущающей переменной z. При этом полагаем, что для фиксированных значений переменных $x_i (i = \overline{1,m})$ среднее значение переменной z равно нулю.

Предпосылка 2. Предполагается, что влияние неучтенных факторов постоянно. Так, при рассмотрении временны́х рядов в различные периоды эти неучтенные факторы оказывают одинаковое влияние.

Предпосылка 3. Отсутствует автокорреляция между возмущающими переменными z.

Предпосылка 4. Число наблюдений должно превышать число параметров регрессии, иначе невозможна оценка этих параметров.

---

[1] *Ферстер Э., Ренц Б.* Методы корреляционного и регрессионного анализа. – М.: Финансы и статистика, 1983.

Предпосылка 5. Предполагается односторонняя зависимость переменной y от факторов $x_i(i = \overline{1,m})$, отсутствие взаимосвязи.

Предпосылка 6. Зависимая переменная y и факторы $x_i(i = \overline{1,m})$ распределены нормально.

С помощью регрессионного анализа при указанных выше предпосылках находят оценки параметров, наиболее хорошо согласующиеся с опытными данными. Данные оценки должны обладать определенными свойствами. Рассмотрим некоторые из этих свойств (без доказательства).

1. Несмещенность оценок параметров регрессии. Оценка параметров регрессии называется несмещенной, если для любого фиксированного числа наблюдений выполняется равенство математического ожидания параметра и значения параметра регрессии. Надо отметить, что оценки, полученные методом наименьших квадратов, обладают свойством несмещенности.

2. Состоятельность оценок параметров регрессии. Данное свойство состоит в том, что с ростом объема выборки оценка параметра регрессии b сходится к теоретическому значению параметра β (вычисленного по всей генеральной совокупности), т. е. ошибка оценки стремится к нулю:

$$P\left(\lim_{n\to\infty} b\right) = \beta. \qquad (5.22)$$

3. Эффективность оценок параметров регрессии. Несмещенная оценка параметра регрессии называется несмещенной эффективной, если она среди всех прочих несмещенных оценок этого же параметра обладает наименьшей дисперсией.

4. Достаточность оценки. Если β представляет собой достаточную оценку параметра b, то не существует другой оценки этого параметра, которую можно получить по выборке из некоторой генеральной совокупности и которая дала бы дополнительную информацию о нем. Р. Фишер показал, что количество измеримой информации, содержащейся в некоторой оценке, равно обратной величине от ее дисперсии. Таким образом, понятие достаточности эквивалентно требованию минимальной дисперсии. Достаточная оценка с необходимостью должна быть эффективной и, следовательно, также состоятельной и несмещенной.

5.3. Этапы построения многофакторной корреляционно-регрессионной модели

Разработка модели и исследование экономических процессов должны выполняться по следующим этапам:

1) априорное исследование экономической проблемы;

2) формирование перечня факторов и их логический анализ;

3) сбор исходных данных и их первичная обработка;

4) спецификация функции регрессии;

5) оценка функции регрессии;

6) отбор главных факторов;

7) проверка адекватности модели;

8) экономическая интерпретация;

9) прогнозирование неизвестных значений зависимой переменной.

Рассмотрим подробнее содержание этапов.

А. Априорное исследование экономической проблемы. В соответствии с целью работы на основе знаний макро-и микроэкономики конкретизируются явления, процессы, зависимость между которыми подлежит оценке. При этом подразумевается прежде всего четкое определение экономических явлений, установление объектов и периода исследования.

На этом этапе исследования должны быть сформулированы экономически осмысленные и приемлемые гипотезы о зависимости экономических явлений.

Б. Формирование перечня факторов и их логический анализ. Для определения наиболее разумного числа переменных в регрессионной модели прежде всего ориентируются на соображения профессионально-теоретического характера. Исходя из физического смысла явления, производят классификацию переменных на зависимую и объясняющую.

В. Сбор исходных данных и их первичная обработка. При построении модели исходная информация может быть собрана в трех видах:

• динамические (временны́е) ряды;

• пространственная информация – информация о работе нескольких объектов в одном разрезе времени;

• сменная – табличная форма. Информация о работе нескольких объектов за разные периоды.

Объем выборки зависит от числа факторов, включаемых в модель с учетом свободного члена. Для получения статистически значимой модели требуется на один фактор объем выборки, равный $1 = 5 \div 8$ наблюдений. Например, если в модель включаются три фактора, то минимальный объем выборки

$$n_{\min} = 5 \cdot (m + n) = 5 \cdot (3 + 1) = 20,$$

где m – число факторов, включаемых в модель;

n – число свободных членов в уравнении.

Если в квартальном разрезе собирать данные, то надо их собирать за 5 лет [20/4].

Г. Спецификация функции регрессии. На данном этапе исследования дается конкретная формулировка гипотезы о форме связи (линейная или нелинейная, простая или множественная и т. д.). Для этого используются различные критерии для проверки состоятельности гипотетического вида зависимости. На этом этапе проверяются предпосылки корреляционно-регрессионного анализа.

Д. Оценка функции регрессии. Здесь определяются числовые значения параметров регрессии и вычисление ряда показателей, характеризующих точность регрессионного анализа.

Е. Отбор главных факторов. Выбор факторов – основа для построения многофакторной корреляционно-регрессионной модели.

На этапе формирования перечня факторов и их логического анализа собираются все возможные факторы, обычно более 20–30 факторов. Но это неудобно для анализа, и модель, включающая 20–30 факторов, будет неустойчива. Неустойчивость модели находит выражение в том, что в ней изменение некоторых факторов ведет к увеличению y вместо снижения y.

Мало факторов – тоже плохо. Это может привести к ошибкам при принятии решений в ходе анализа модели. Поэтому необходимо выбирать более рациональный перечень факторов. При этом проводят анализ факторов на мультиколлинеарность.

Анализ и способы снижения влияния мультиколлинеарности на значимость модели. Мультиколлинеарность – попарная корреляционная зависимость между факторами.

Мультиколлинеарная зависимость присутствует, если коэффициент парной корреляции $r_{ij} = \geq 0{,}70 \div 0{,}80$.

Отрицательное воздействие мультиколлинеарности состоит в следующем:

1) усложняется процедура выбора главных факторов;

2) искажается смысл коэффициента множественной корреляции (он предполагает независимость факторов);

3) усложняются вычисления при построении самой модели;

4) снижается точность оценки параметров регрессии, искажается оценка дисперсии.

Следствием снижения точности является ненадежность коэффициентов регрессии и отчасти неприемлемость их использования для интерпретации как меры воздействия соответствующей объясняющей переменной на зависимую переменную.

Оценки коэффициента становятся очень чувствительными к выборочным наблюдениям. Небольшое увеличение объема выборки может привести к очень сильным сдвигам в значениях оценок. Кроме того, стандартные ошибки оценок входят в формулы критерия значимости, поэтому применение самих критериев становится также ненадежным. Из сказанного ясно, что исследователь должен пытаться установить стохастическую мультиколлинеарность и по возможности устранить ее.

Для измерения мультиколлинеарности можно использовать коэффициент множественной детерминации

$$Д = R^2, \tag{5.23}$$

где R – коэффициент множественной корреляции.

При отсутствии мультиколлинеарности факторов

$$Д = \sum_{j=1}^{m} d_{yj}, \tag{5.24}$$

где d_{yj} – коэффициент парной детерминации, вычисляемый по формуле

$$d_{yj} = r^2_{yj}, \tag{5.25}$$

где r_{yj} – коэффициент парной корреляции между j-м фактором и зависимой переменной y.

При наличии мультиколлинеарности соотношение (5.24) не соблюдается. Поэтому в качестве меры мультиколлинеарности используется следующая разность:

$$M = Д - \sum_{j=1}^{m} d_{yj}. \tag{5.26}$$

Чем меньше эта разность, тем меньше мультиколлинеарность. Для устранения мультиколлинеарности используется *метод исключения переменных.* Этот метод заключается в том, что высоко коррелированные объясняющие переменные (факторы) устраняются из регрессии и она заново оценивается. Отбор переменных, подлежащих исключению, производится с помощью коэффициентов парной корреляции. Опыт показывает, что если $|r_{yj}| \geq 0{,}70$, то одну из переменных можно исключить, но какую переменную исключить из анализа, решают исходя из управляемости факторов на уровне предприятия.

Обычно в модели оставляют тот фактор, на который можно разработать мероприятие, обеспечивающее улучшение значения этого фактора в планируемом году. Возможна ситуация, когда оба мультиколлинеарных фактора управляемы на уровне предприятия. Решить вопрос об исключении того или иного фактора можно только в соответствии с процедурой отбора главных факторов.

Отбор факторов не самостоятельный процесс, он сопровождается построением модели. Принятие решения об исключении факторов производится на основе анализа значений специальных статистических характеристик и с учетом управляемости факторов на уровне предприятия.

Процедура отбора главных факторов. Эта процедура обязательно включает следующие этапы:

1. Анализ факторов на мультиколлинеарность и ее исключение. Здесь производится анализ значений коэффициентов парной корреляции r_{ij} между факторами x_i и x_j.

2. Анализ тесноты взаимосвязи факторов (x) с зависимой переменной (y).

Для анализа тесноты взаимосвязи x и y используются значения коэффициента парной корреляции между фактором и функцией (r_{x_iy}). Величина r_{x_iy} определяется на ЭВМ и представлена в корреляционной матрице вида (табл. 5.1).

Таблица 5.1

Корреляционная матрица

№ переменной	x_1	x_2	x_3	...	x_m	y
x_1	1	$r_{x_1x_2}$	$r_{x_1x_3}$	...	$r_{x_1x_m}$	r_{x_1y}
x_2	$r_{x_2x_1}$	1	$r_{x_2x_3}$	...	$r_{x_2x_m}$	r_{x_2y}
x_3	$r_{x_3x_1}$	$r_{x_3x_2}$	1	...	$r_{x_3x_m}$	r_{x_3y}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
x_m	$r_{x_mx_1}$	$r_{x_mx_2}$	$r_{x_mx_3}$	...	1	r_{x_my}
y	r_{yx_1}	r_{yx_2}	r_{yx_3}	...	r_{yx_m}	1

Факторы, для которых $r_{x_iy} = 0$, т. е. не связанные с y, подлежат исключению в первую очередь. Факторы, имеющие наименьшее значение r_{x_iy}, могут быть потенциально исключены из модели. Вопрос об их окончательном исключении решается в ходе анализа других статистических характеристик.

3. Анализ коэффициентов β факторов, которые потенциально могут быть исключены.

Коэффициент β учитывает влияние анализируемых факторов на y с учетом различий в уровне их колеблемости. Коэффициент β показывает, насколько сигм (средних квадратических отклонений) изменяется функция с изменением соответствующего аргумента на одну сигму при фиксированном значении остальных аргументов:

$$\beta_k = a_k \frac{\sigma_{x_k}}{\sigma_y}, \tag{5.27}$$

где β_k – коэффициент β k-го фактора;

σ_{x_k} – среднее квадратическое отклонение k-го фактора;

σ_y – среднее квадратическое отклонение функции;

a_k – коэффициент регрессии при k-м факторе.

Из двух факторов x_i и x_j может быть исключен тот фактор, который имеет меньшее значение β.

Допустим, исключению подлежит один из мультиколлинеарных факторов x_i или x_j. Оба фактора управляемы на уровне предприятия, коэффициенты регрессии a_i и a_j статистически значимы. Фактор x_i более тесно связан с y, т. е. $r_{x_i y} > r_{x_j y}$, но при этом $\beta_{x_i} < \beta_{x_j}$. В этом случае обычно исключению подлежит фактор x_j.

4. Проверка коэффициентов регрессии на статистическую значимость.

Проверка может быть произведена двумя способами:

• проверка статистической значимости a_k по *критерию Стьюдента* проводится по следующей формуле

$$t_k = \frac{a_k}{S_{a_k}}, \tag{5.28}$$

где a_k – коэффициент регрессии при k-м факторе;

S_{a_k} – стандартное отклонение оценки параметра a_k[1].

Число степеней свободы статистики t_k равно

$$f = n - m - 1,$$

где m – количество факторов, включенных в модель.

Значение t, вычисляемое по (5.28), сравнивают с критическим значением $t_{f,a}$, найденным по приложению 1 при заданном уровне значимости α и числе степеней свободы f (двухсторонняя критическая область).

Если $t_k \geq t_{f,\alpha}$, то a_k существенно больше 0, а фактор x_k оказывает существенное влияние на y. При этом фактор x_k оставляем в модели. Если $t_k < t_{f,\alpha}$, то фактор исключаем из модели;

• проверка статистической значимости a_k по *критерию Фишера* –

$$F_k = \left(\frac{a_k}{S_{a_k}}\right)^2 = t^2, \tag{5.29}$$

где t^2 – многомерный аналог критерия Стьюдента.

---

[1] *Ферстер Э., Ренц Б.* Методы корреляционного и регрессионного анализа. – М.: Финансы и статистика, 1983.

Число степеней свободы статистики F_k следующее: $f_1 = 1$, $f_2 = n - m - 1$. Значение F_k, вычисляемое по формуле (5.29), сравнивают с критическим значением $F_{f_1 f_2 \alpha}$, найденным по приложению 2, при заданных уровне значимости α и числе степеней свободы f_1, f_2.

Если $F_k \geq F_{f_1 f_2 \alpha}$, то α_k – существенно больше 0, а фактор x_k оказывает существенное влияние на y. При этом фактор x_k оставляем в модели. Если $F_k < F_{f_1 f_2 \alpha}$, то фактор исключаем из модели.

5. Анализ факторов на управляемость.

В ходе логического анализа на основе экономических знаний исследователь должен сделать вывод: можно ли разработать организационно-технические мероприятия, направленные на улучшение (изменение) выбранных факторов на уровне предприятия. Если это возможно, то данные факторы управляемы. Неуправляемые факторы на уровне предприятия могут быть исключены из модели. Например, из двух факторов x_1 – средняя техническая скорость автомобилей и x_2 – время погрузки-разгрузки на одну ездку при равенстве или близких по значению таких характеристик, как $r_{x_1 y}$ и $r_{x_2 y}$, β_{x_1} и β_{x_2}, исключению подлежит x_1. На уровне АТП практически невозможно повлиять на значение технической скорости, которая зависит в основном от климатических условий и величины транспортного потока.

6. Строится новая регрессионная модель без исключенных факторов. Для этой модели определяется коэффициент множественной детерминации *Д*.

7. Исследование целесообразности исключения факторов из модели с помощью коэффициента детерминации.

Прежде чем вынести решение об исключении переменных из анализа в силу их незначимого влияния на зависимую переменную, производят исследования с помощью коэффициента детерминации.

В первой регрессии содержится m объясняющих переменных, во второй – только часть из них, а именно m_1 объясняющих переменных. При этом $m = m_1 + m_2$, т. е. во вторую регрессию мы не включили m_2 объясняющих переменных. Теперь следует проверить, вносят ли совместно эти m_2 переменных существенную долю в объяснение вариации переменной y. Для этого используется статистика

$$F = \frac{(Д_m - Д_{m_1}) \cdot (n - m - 1)}{(m - m_1) \cdot (1 - Д_m)}, \qquad (5.30)$$

которая имеет F-распределение с $f_1 = m - m_1 = m_2$ и $f_2 = n - m - 1$ степенями свободы. Здесь $Д_m$ означает коэффициент детерминации регрессии с m объясняющими переменными, а $Д_{m_1}$ – коэффициент детерминации регрессии с m_1-факторами.

Разность $(Д_m - Д_{m_1})$ в числителе формулы является мерой дополнительного объяснения вариации переменной y за счет включения m_2 переменных.

Критическое значение $F_{f_1 f_2}$ находят по таблице F-распределения при заданном уровне значимости α и f_1 и f_2 степенях свободы. Если $F \leq F_{f_1 f_2 \alpha}$, то включение дополнительно объясняющих переменных совместно не оказывает значимого влияния на переменную y. Если $F > F_{f_1 f_2 \alpha}$, то m_2 объясняющих переменных совместно оказывают существенное влияние на вариацию переменной y, и, следовательно, в этом случае все m_2 переменные нельзя исключать из модели.

При реализации первой ситуации $(F \leq F_{f_1 f_2 \alpha})$ факторы окончательно исключаются из модели.

Ж. Проверка адекватности модели.

Данный этап анализа включает следующие процедуры:

• *оценку значимости коэффициента детерминации.* Данная оценка необходима для решения вопроса: оказывают ли выбранные факторы влияние на зависимую переменную? Оценку значимости Д следует проводить, так как может сложиться такая ситуация, когда величина коэффициента детерминации будет целиком обусловлена случайными колебаниями в выборке, на основе которой он вычислен. Это объясняется тем, что величина Д существенно зависит от объема выборки.

Для оценки значимости коэффициента множественной детерминации используется следующая статистика:

$$F = \frac{Д(n-m-1)}{m(1-Д)}, \tag{5.31}$$

которая имеет F-распределение с $f_1 = m$ и $f_2 = n - m - 1$ степенями свободы. Здесь $Д = R^2$, а m – количество учитываемых объясняющих переменных (факторов).

Значение статистики F, вычисленное по эмпирическим данным, сравнивается с табличным значением $F_{f_1 f_2 \alpha}$. Критическое значение определяется по приложению 2 по заданному α и степеням свободы f_1 и f_2. Если $F > F_{f_1 f_2 \alpha}$, то вычисленный коэффициент детерминации значимо отличается от 0 и, следовательно, включенные в регрессию переменные достаточно объясняют зависимую переменную, что позволяет говорить о значимости самой регрессии (модели);

• *проверку качества подбора теоретического уравнения.* Она проводится с использованием средней ошибки аппроксимации. Средняя ошибка аппроксимации регрессии определяется по формуле:

$$E = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \frac{y_i - y_{iт}}{y_{iт}} \right| \cdot 100\%; \tag{5.32}$$

• ***вычисление специальных показателей,*** которые применяются для характеристики воздействия отдельных факторов на результирующий показатель. Это коэффициент эластичности, который показывает, на сколько процентов в среднем изменяется функция с изменением аргумента на 1% при фиксированных значениях других аргументов:

$$Э_k = a_k \cdot \frac{\bar{x}_k}{\bar{y}}; \tag{5.33}$$

доля влияния каждого фактора x_j в отдельности на вариацию y[1]:

$$g_i = \beta_j^2, \tag{5.34}$$

где β_j – коэффициент бета фактора x_j.

Показатель g_j является мерой вариации результативного признака за счет изолированного влияния фактора x_j. Следует отметить, что система факторов, входящая в модель регрессии, – это не простая их сумма, так как система предполагает внутренние связи, взаимодействие составляющих ее элементов. Действие системы не равно арифметической сумме воздействий составляющих ее элементов. Поэтому необходимо определить показатель системного эффекта факторов η_s:

$$\eta_s = R^2 - \sum_{j=1}^{m} \beta_j^2.$$

На основе анализа специальных показателей и значений парной корреляции x с y делают вывод, какие из главных факторов оказывают наибольшее влияние на y. После этого переходят к разработке организационно-технических мероприятий, направленных на улучшение значений этих факторов, с целью повышения (снижения) результативного показателя y.

3. Экономическая интерпретация.

Результаты регрессионного анализа сравниваются с гипотезами, сформулированными на первом этапе исследования, и оценивается их правдоподобие с экономической точки зрения.

---

[1] *Елисеева И. И., Юзбашев М. М.* Общая теория статистики. – М.: Финансы и статистика, 1999.

И. Прогнозирование неизвестных значений зависимой переменной. Полученное уравнение регрессии находит практическое применение в прогностическом анализе. Прогноз получают путем подстановки в регрессию с численно оцененными параметрами значений факторов. Следует подчеркнуть, что прогнозирование результатов по регрессии лучше поддается содержательной интерпретации, чем простая экстраполяция тенденций, так как полнее учитывается природа исследуемого явления. Более подробно вопросы прогнозирования рассмотрены в следующей главе.

Глава 6
Методы и модели прогнозирования временных рядов экономических показателей

6.1. Основные положения и понятия в прогнозировании временных рядов

Среди большого разнообразия экономико-математических методов, используемых для решения задач управления предприятием, особое место занимают методы и модели прогнозирования.

Следует различать два понятия, связанных с прогнозированием, – предсказание и собственно прогнозирование.

Под *предсказанием* понимают суждение о будущем состоянии процесса, основанное на субъективном «взвешивании» большого числа факторов качественного и количественного характера. *Прогнозирование* – это исследовательский процесс, в результате которого получают прогноз о состоянии объекта. Прогноз является вероятностным суждением о возможном состоянии объекта или об альтернативных путях его достижения. Известно большое количество методов, методик и способов прогнозирования. Все они основаны на двух крайних подходах: эвристическом и математическом.

Эвристические методы базируются на использовании явлений или процессов, не поддающихся формализации.

Для математических методов прогнозирования характерен подбор и обоснование математической модели исследуемого процесса, а также способов определения ее неизвестных параметров. Задача прогнозирования при этом сводится к решению уравнений, описывающих данную модель для заданного момента времени.

Среди математических методов прогнозирования в особую группу выделяются методы экстраполяции, которые отличаются простотой, наглядностью и легко реализуются на ЭВМ. Методологическая предпосылка экстраполяции состоит в признании преимущественной связи между прошлым, настоящим и будущим. При этом развитие экономических явлений наиболее полно находит свое отражение во временных рядах, которые представляют собой упорядоченные во времени наборы измерений каких-либо характеристик исследуемого объекта, процесса. Поэтому независимая переменная для временного ряда, это, как правило, календарные равные отрезки времени (год, квартал, месяц и т. д.). Основной чертой, выделяющей временные ряды среди других видов статистических данных, является существенность порядка, в котором производятся наблюдения.

В ходе решения задачи прогнозирования пользуются ограниченным количеством информации об одномерном временном ряде конечной длины. При этом в экономике исследуются дискретные временные ряды, наблюдаемые в дискретные моменты времени.

Дискретный временной ряд можно рассматривать как последовательность значений $y_1, y_2, ..., y_n$ в моменты времени t, или сокращенно $y_t (t = 1, 2, ..., n)$.

Временной ряд может быть представлен в следующем виде:

$$y_t = x_t + \varepsilon_t, \quad (6.1)$$

где x_t – детерминированная неслучайная компонента процесса;

ε_t – стохастическая случайная компонента процесса.

Детерминированная компонента (тренд) x_t характеризует существующую динамику процесса в целом, основную, длительную тенденцию изменения изучаемого показателя. Стохастическая компонента ε_t отражает случайные колебания или шумы процесса. Задача прогнозирования, в частности, состоит в определении вида экстраполирующих функций x_t и ε_t на основе исходных эмпирических данных.

Отрезок времени ℓ от момента времени t, для которого имеются последние статистические данные об изучаемом процессе, до момента, к которому относится прогноз, называется *периодом упреждения* (периодом прогноза). В зависимости от длительности периода упреждения применительно к экономике различают три вида прогноза:

- краткосрочные – с периодом упреждения от нескольких дней до трех лет;
- среднесрочные – от трех до 5 лет;
- долгосрочные – от 5 лет и выше.

При прогнозировании, как правило, в точке прогноза оценивают математическое ожидание процесса (точечный прогноз) и величину интервала, в который с заданной вероятностью попадет прогнозируемое значение процесса (интервальный прогноз). Результаты экстраполяции наиболее надежны при кратко- и среднесрочном прогнозировании. При этом предполагается, что совокупность факторов, определявших тенденцию временного ряда в прошлом, в среднем сохранит свою силу и направление действия в течение прогнозируемого периода.

В экономической литературе предлагается широкий спектр методов экстраполяции. Остановимся на краткой характеристике основных методов данной группы.

6.2. Характеристика методов и моделей прогнозирования показателей работы предприятий

В настоящее время разработана большая группа экстраполяционных методов прогнозирования отдельных экономических показателей. В данной группе методов можно выделить следующие:

1) основанные на построении многофакторных корреляционно-регрессионных моделей;

2) авторегрессии, учитывающие взаимосвязь членов временного ряда;

3) основанные на разложении временного ряда на компоненты: главная тенденция (тренд), сезонные колебания и случайная составляющая;

4) позволяющие учесть неравнозначность исходных данных;

5) прямой экстраполяции, при этом используются разные трендовые модели.

Коротко охарактеризуем эти методы. При прогнозировании методом корреляционно-регрессионного анализа строится модель, включающая набор переменных, от которых зависит поведение функции. Основным недостатком этого подхода является то, что необходимы сбор и обработка больших массивов информации по группе однородных предприятий и прогнозирование самих объясняющих переменных. При этом остается открытым вопрос о прогнозировании показателей работы предприятий, не вошедших в группу однородных.

Для данных методов характерна невысокая точность прогноза для конкретного, отдельного предприятия.

Авторегрессионные модели чаще всего используются для прогнозирования тех экономических процессов, для которых внешний механизм их формирования четко не определен, и практически невозможно выделить стабильные во времени причинно-следственные связи. Применение этих моделей целесообразно и для сильно автокоррелированных динамических рядов.

Главная идея методов авторегрессии состоит в том, что будущие значения временного ряда не могут произвольно отклоняться в большую или меньшую сторону от предшествующих значений временного ряда, какими бы причинами ни были вызваны эти отклонения. Во временных рядах экономических показателей существует связь между недавно реализованными значениями и значением, реализующимся в близком будущем. Смысл этой связи таков, что если между близкими значениями временного ряда существует корреляция, то можно построить прогноз показателя. Модель авторегрессии имеет вид:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_m y_{t-m}, \tag{6.2}$$

где $a_0, a_1, a_2, \ldots, a_m$ – параметры уравнения авторегрессии;
y_t – значение динамического ряда показателя.

Считают, что при расчете необходимо проводить элиминирование мультиколлинеарности в матрице входных параметров (y_{t-1}, $y_{t-2}, \ldots, y_{t-m}$). Для улучшения прогнозирующих свойств модели (6.2) в нее можно ввести фактор времени в виде самостоятельной переменной. Подобный подход в ряде случаев существенно увеличивает точность прогноза, что объясняется учетом линейного тренда.

Методы, основанные на разложении временного ряда на компоненты – главная тенденция, сезонные колебания и случайная составляющая, – позволяют описать почти любой экономический процесс, независимо от его характера.

При аддитивной связи между компонентами модель имеет вид:

$$y_t = y_t' + y_t'', \tag{6.3}$$

где y_t' – составляющая, описывающая тренд.

Составляющая y_t'' определяется по формуле

$$y_t'' = y_t - y_t' \tag{6.4}$$

и может быть разложена в ряд Фурье:

$$y_t'' = a_0 + \sum_{i=1}^{m} (a_i \cdot \cos it_r + b_i \sin it_r) + \xi_t. \tag{6.5}$$

Формулы для определения параметров уравнения (6.5) приведены в работах Л. Г. Лабскера, Л. О. Бабешко «Теория массового обслуживания в экономической сфере» и А. А. Френкеля «Математические методы анализа динамики и прогнозирования производительности труда».

Модель «гармонический фильтр» с учетом статистически значимых гармоник аналогична выше описанной модели. Оценка значимости i-й гармоники рассчитывается по следующей формуле:

$$F_{y_i} = \frac{n \cdot S^2_{\xi t}}{n \cdot (a_i^2 + b_i^2)}, \tag{6.6}$$

где $S^2{}_{\xi t}$ – дисперсия остатков.

Несмещенная оценка $S^2{}_{\xi t}$ рассчитывается так:

$$S^2_{\xi t} = \frac{\sum\limits_{t=0}^{n-1} y_t'' - n \cdot a_0^2 - \frac{n}{2} \sum\limits_{i=1}^{m} (a_i^2 + b_i^2)}{n - 2m - 1} \tag{6.7}$$

и подчиняется примерно F-распределению Фишера с $\nu_1 = 2$ и $\nu_2 = n - 5$ степенями свободы.

Проведенные расчеты показали, что при использовании моделей гармонического фильтра необходимо с осторожностью подходить к выбору величины предпрогнозного периода, так как величина предпрогнозного периода в значительной степени определяет точность прогноза.

Рассмотренные выше методы не позволяют в достаточной степени учесть неравнозначность исходных данных. К числу методов, учитывающих неравнозначность данных, можно отнести:
- метод авторегрессии с последующей адаптацией коэффициентов уравнения;
- метод взвешенных отклонений.

Для адаптации коэффициентов модели авторегрессии может быть использован метод наискорейшего спуска[1]. Согласно данному методу процедура пересчета коэффициентов уравнения авторегрессии осуществляется следующим образом:

$$A = A_c - k - \mathrm{grad}(\ell^2{}_{t+\tau}), \tag{6.8}$$

где A – вектор новых коэффициентов;
A_c – вектор старых коэффициентов;
k – коэффициент, $k > 0$;
$\ell_{t+\tau}$ – ошибка прогноза в точке $(t + \tau)$.

[1] *Лукинский В. С., Зайцев Е. И., Бережной В. И.* Модели и алгоритмы управления обслуживанием и ремонтом автотранспортных средств. – Спб.: СпГИЭА, 1997.

После дифференцирования (формула 6.8) по коэффициентам a_i и соответствующих преобразований получим

$$\mathrm{grad}(\ell^2{}_{t+\tau}) = -2_{\ell_{t+\tau}} \cdot x_t, \tag{6.9}$$

где x_t – вектор значений входных переменных в точке t.

Таким образом, откорректированные оценки коэффициентов определяются по формуле

$$A_н = A_c + 2k \,.\, \ell_{t+\tau} \,.\, x_t. \tag{6.10}$$

В уравнении (6.10) неизвестным является коэффициент k, который может быть идентифицирован как коэффициент, определяющий скорость движения в направлении, обратном градиенту. Для определения значения коэффициента k используется итеративная процедура, описанная в работе В. С. Лукинского, Е. И. Зайцева, В. И. Бережного «Модели и алгоритмы управления обслуживанием и ремонтом автотранспортных средств». Заметим, что адаптация производится либо по последнему эмпирическому значению, либо по предыдущему производному значению.

Метод взвешенных отклонений достаточно подробно изложен в работе Л. Г. Лабскера, Л. О. Бабешко «Теория массового обслуживания в экономической сфере». Для сравнения точности прогнозных оценок, получаемых с использованием рассмотренных выше методов, был произведен ретроспективный прогноз показателей работы подвижного состава десяти транспортных предприятий. Объем выборки составил более 857 вариантов расчетов. Средние ошибки приведены в табл. 6.1.

Таблица 6.1

Модели прогнозирования показателей работы автомобилей

№ п/п	Наименование модели прогнозирования	Средняя ошибка прогноза, %
1	Авторегрессия без учета времени	1,89
2	Авторегрессия с учетом времени	2,94
3	«Гармонический фильтр» без учета значимости	1,9
4	«Гармонический фильтр» с учетом значимости	2,34
5	Метод взвешенных отклонений	2,24
6	Авторегрессия без учета времени и с последующей адаптацией параметров модели	1,82
7	Авторегрессия с учетом времени и последующей адаптацией параметров модели	1,94

Анализ ошибок (табл. 6.1) позволяет сделать следующие выводы.

Все модели прогнозирования обладают достаточно высокой точностью. Наиболее точным методом прогнозирования показателей работы транспортных предприятий является авторегрессия без учета фактора времени и с последующей адаптацией коэффициентов данной модели ($\bar{E}$=1,82%).

Модель авторегрессии без учета фактора времени и с последующей адаптацией коэффициентов данной модели в отдельных случаях может значительно уступать по точности другим моделям прогнозирования. Например, при прогнозировании показателя балансовой прибыли ошибка прогноза оказалась в 2,61 раза больше, чем ошибка прогноза, полученного с использованием метода взвешенных отклонений. Общее число случаев, когда модель авторегрессии без учета времени и с последующей адаптацией коэффициентов данной модели оказалась лучшей по точности, составило 67%. Поэтому для прогнозирования экономических показателей работы предприятий необходимо использовать комплекс моделей прогнозирования, приведенных в табл. 6.1.

Сложность математического аппарата моделей прогнозирования, представленных в табл. 6.1, не оправдывает себя. Для получения точных оценок прогнозирования в каждом случае необходимо использовать эти модели прогнозирования в комплексе, что значительно увеличивает время на получение прогноза.

Проведенные исследования показали, что при краткосрочном прогнозировании (на один год) показателей работы предприятий целесообразно использовать комплекс трендовых моделей табл. 6.2, который позволяет с достаточной точностью описать динамику показателей.

Таблица 6.2

Таблица кодов и моделей прогноза

№ п/п	Модель прогнозирования	№ п/п	Модель прогнозирования
1	$y = A + B\,x$	9	$y = -A\,x^{B}$
2	$y = 1/(A + B\,x)$	10	$y = A + B\ln(x)$
3	$y = A + B/x$	11	$y = A + B\log(x)$
4	$y = x/(A + B\,x)$	12	$y = A/(B + x)$
5	$y = A\,B^{x}$	13	$y = A\,x/(B + x)$
6	$y = A\,exp(B\,x)$	14	$y = A\,exp\,(B/x)$
7	$y = 10^{(B\,x)}$	15	$y = A\,10^{(B/x)}$
8	$y = 1/(A + B\,exp\,(-x))$	16	$y = A + B\,(x^{N})$

В связи с вышеизложенным остановимся более детально на процедуре прогнозирования с помощью прямой экстраполяции.

6.3. Прогнозирование с помощью методов экстраполяции

Прогнозирование с помощью методов экстраполяции должно включать следующие этапы работ.

А. Установление цели и задачи исследования, анализ объекта прогнозирования.

Прогнозирование развития любой системы (предприятия, фирмы и т. д.) предъявляет специфические требования к параметрам (объектам), характеризующим и определяющим ее развитие. Поэтому необходимо на первом этапе работ провести детальное логическое изучение системы: зависимость рассматриваемого объекта (параметра, показателя) от других систем одного уровня и субсистемы (системы более высокого уровня); взаимосвязь между данным объектом и другими объектами системы; установление характера предоставления статистических данных об объекте.

Б. Подготовка исходных данных.

Работы по этому этапу начинаются с проверки временного ряда, в результате которой устанавливаются полнота ряда (наличие данных за каждый год (месяц, квартал) ретроспективного периода), сопоставимость данных и, в случае необходимости, проверка методики приведения данных к сопоставимому виду. Если временной ряд представлен не полностью, то необходимо недостающие данные определить с помощью тех или иных методов интерполяции в зависимости от характера протекания процесса.

Наряду с этим осуществляется также формирование массива функций, который в последующем будет использован для выбора вида математической модели.

В. Фильтрация исходного временно́го ряда.

В результате этой процедуры устраняются случайные возмущения (флуктуации), возникающие под воздействием неучтенных факторов или ошибок измерения относительно наиболее вероятного протекания процесса, и тем самым исключается искажающее влияние случайных колебаний на выбор вида регрессии. Фильтрация исходного динамического ряда включает его сглаживание и выравнивание.

Сглаживание применяется для устранения случайных отклонений (шума) из экспериментальных значений исходного ряда. Сглаживание производится с помощью многочленов, приближающих (обычно по методу наименьших квадратов) группы опытных точек.

Наилучшее сглаживание получается для средних точек группы, поэтому желательно выбирать нечетное количество точек в сглаживаемой группе. Обычно их выбирают три или пять. Например, по первым трем точкам (y_1, y_2, y_3) сглаживают среднюю y_2, затем по следующей тройке (y_2, y_3, y_4) сглаживают y_3 и т. д. Крайние точки сглаживают по специальным формулам.

Чаще всего для сглаживания применяют линейную зависимость. Тогда формулы сглаживания для групп из трех точек имеют вид:

$$\hat{y}_0 = \frac{1}{3}(y_{-1} + y_0 + y_{+1}); \tag{6.11}$$

$$\hat{y}_{-1} = \frac{1}{6}(5y_{-1} + 2y_0 + y_{+1}); \tag{6.12}$$

$$\hat{y}_{+1} = \frac{1}{6}(-y_{-1} + 2y_0 + y_{+1}); \tag{6.13}$$

где y_0, $\hat{y}_0$ — значения исходной и сглаженной функций в средней точке группы;
y_{-1}, $\hat{y}_{-1}$ — значения исходной и сглаженной функций в левой точке группы;
y_{+1}, $\hat{y}_{+1}$ — значения исходной и сглаженной функций в правой точке группы.

Формулы (6.12), (6.13) применяются для сглаживания крайних точек ряда. Для сглаживания по пяти точкам формулы имеют вид:

$$\hat{y}_0 = \frac{1}{5}(y_{-2} + y_{-1} + y_0 + y_{+1} + y_{+2}); \tag{6.14}$$

$$\hat{y}_{-1} = \frac{1}{10}(4y_{-2} + 3y_{-1} + 2y_0 + y_{+1}); \tag{6.15}$$

$$\hat{y}_{+1} = \frac{1}{10}(y_{-1} + 2y_0 + 3y_{+1} + 4y_{+2}); \tag{6.16}$$

$$\hat{y}_{-2} = \frac{1}{5}(3y_{-2} + 2y_0 + y_{+1} - y_{+2}); \tag{6.17}$$

$$\hat{y}_{+2} = \frac{1}{5}(-y_{-2} + y_0 + 2y_{+1} + 3y_{+2}). \tag{6.18}$$

Сглаживание (даже в простом линейном варианте) является во многих случаях эффективным средством выявления тренда при наличии в экспериментальных точках случайных помех и ошибок измерения.

Выравнивание применяется для более удобного представления исходного ряда без изменения его числовых значений. Выравниванием называется приведение исходной эмпирической формулы

$$y = f(t,\ a,\ b), \tag{6.19}$$

где t — время,
a, b — параметры,

к виду

$$y = a_1 T + b_0. \tag{6.20}$$

Использование двухпараметрической зависимости (6.19) объясняется ее наибольшим распространением в практике прогнозирования и сравнительно простыми способами получения выравниваемых формул. Функции с бóльшим (чем 2) числом параметров выравниваются не всегда, и формулы имеют громоздкий вид.

Наиболее распространенными способами выравнивания являются логарифмирование и замена переменных.

Пример 6.1. Дана исходная функция $y = at^b$.

Логарифмируя, получим $\lg y = \lg a + b \cdot \lg t$.

Вводя замену переменных, имеем: $T = \lg t$; $Y = \lg y$; $Y = a_1 T + b_1$, где $a_1 = b$; $b_1 = \lg a$.

Перестроив исходные данные (точки) на логарифмической бумаге, получим линейную зависимость, с которой легче работать и определять коэффициенты. Затем нужно пересчитать результаты по формулам, обратным исходному преобразованию.

Пример 6.2. Дана исходная формула $y = a \cdot e^{bt}$.

Выравнивание $\lg y = \lg a + b \cdot t \cdot \lg e$; $a_1 = b \cdot \lg e$; $b_1 = \lg a$, тогда

$$Y = \lg y = b_1 + a_1 \cdot t.$$

Пример 6.3. Дана исходная формула $y = \dfrac{1}{at + b}$.

Выравнивание $Y = \dfrac{1}{y} = a^t + b$.

Пример 6.4. Дана исходная формула $y = \dfrac{1}{a + b \cdot e^{-t}}$.

Выравнивание $Y = \dfrac{1}{y}$; $T = e^{-t}$; $a_1 = b$; $b_1 = a$; $Y = b_1 + a_1 \cdot T$.

Можно рассматривать выравнивание не как метод представления исходного динамического ряда, а как метод непосредственно-го приближенного определения параметров аппроксимирующей функции, что часто и делается на практике.

Г. Логический отбор видов аппроксимирующей функции.

На основе изучения статистических данных и логического анализа протекания изучаемого процесса из заданного массива функций отбираются наиболее приемлемые виды уравнений связи. Этот этап необходим, так как позволяет при отборе функций учесть основные условия протекания рассматриваемого процесса и требования, предъявляемые к математической модели. На этом этапе должны быть решены следующие вопросы:

- является ли исследуемый показатель величиной, монотонно возрастающей (убывающей), стабильной, периодической, имеющей один или несколько экстремумов;
- ограничен ли показатель сверху или снизу каким-либо пределом;
- имеет ли функция, определяющая процесс, точку перегиба;
- обладает ли анализируемая функция свойством симметричности;
- имеет ли процесс четкое ограничение развития во времени.

Рассмотрим те функции, которые предпочтительно использовать в прогнозной экстраполяции.

В качестве аппроксимирующих функций чаще всего используются различные полиномы с ограничением числа членов (степени полинома). Это

степенной полином

$$y(t) = a_0 + \sum_{i=1}^{n} a_i t^i; \tag{6.21}$$

экспоненциальный полином

$$y(t) = \exp\left[a + \sum_{i=1}^{n} a_i t^i\right]; \tag{6.22}$$

гиперболический полином

$$y(t) = a_0 + \sum_{i=1}^{n} \frac{1}{a_i t^i}, \tag{6.23}$$

где y – прогнозируемый показатель;
t – время;
$a_0, a_1, ..., a_n$ – параметры (коэффициенты), подлежащие определению.

Опыт применения аппроксимирующих функций для целей прогнозирования показывает, что наиболее простыми (математически) и чаще всего используемыми являются следующие функции:

- линейная $y(t) = a + bt;$ (6.24)
- квадратичная $y(t) = a + bt + ct^2;$ (6.25)
- степенная $y(t) = at^b;$ (6.26)
- экспоненциальная $y(t) = a\exp(bt);$ (6.27)
- модифицированная экспонента $y(t) = k - ae^{-bt};$ (6.28)
- гиперболическая $$y(t) = a + \frac{b}{c+t}; \quad (6.29)$$
- логистическая кривая $$y(t) = \frac{k}{1+be^{-et}}, \quad (6.30)$$

где a, b, c, k – параметры.

Когда это возможно, при выборе вида аппроксимирующей функции прибегают к графическому способу подбора по виду точек временного ряда, расположенных на плоскости $y0t$. Если по графику подобрать функцию трудно, иногда прибегают к анализу производных от соответствующих видов функций аппроксимации (или разностей Δ_1, Δ_2, Δ_3, ...) соответствующего порядка.

Выбирают ту функцию для прогноза, арифметическая средняя для разностного ряда которого будет равна нулю или близка к нулю по абсолютной величине.

Окончательное решение о виде аппроксимирующей функции может быть принято после определения ее параметров и верификации прогноза по ретроспективному ряду. Поэтому для прогнозирования используют несколько подходящих аппроксимирующих функций, с тем чтобы после оценки точности выбрать наиболее подходящую.

Д. Оценка математической модели прогнозирования.

На этом этапе исследования определяются параметры различных видов аппроксимирующих функций. Наиболее распространенными методами оценки параметров аппроксимирующих зависимостей являются метод наименьших квадратов (МНК) и его модификации, метод экспоненциального сглаживания, метод вероятностного моделирования, метод адаптивного сглаживания.

Рассмотрим для примера МНК и метод экспоненциального сглаживания.

Метод наименьших квадратов состоит в определении параметров модели тренда, минимизирующих ее отклонение от точек исходного временного ряда:

$$S = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \to \min, \tag{6.31}$$

где $\hat{y}_i$ – расчетные (теоретические) значения исходного ряда;
y_i – фактические значения исходного ряда;
n – число наблюдений.

Классический метод наименьших квадратов предполагает равноценность исходной информации в модели. В реальной же практике будущее поведение процесса в большей степени определяется поздними наблюдениями, чем ранними. Речь идет о дисконтировании, т. е. уменьшении ценности более ранней информации.

Дисконтирование учитывают путем введения в модель (6.31) некоторых весов $\beta_i < 1$. Тогда

$$S = \sum_{i=1}^{n} \beta_i (\hat{y}_i - y_i)^2 \to \min.$$

Коэффициенты β_i могут быть заданы в числовой форме или в виде функциональной зависимости таким образом, чтобы по мере продвижения в прошлое веса́ убывали.

Метод наименьших квадратов широко применяется при прогнозировании, что объясняется его простотой и легкостью реализации на ЭВМ. К недостаткам МНК можно отнести следующее.

Во-первых, модель тренда жестко фиксируется, и с помощью МНК можно получить прогноз на небольшой период упреждения. Поэтому МНК относят к методам краткосрочного прогнозирования.

Во-вторых, значительную трудность представляет правильный выбор вида модели, а также обоснование и выбор весов во взвешенном методе наименьших квадратов.

Наконец, МНК очень просто реализуется только для линейных и линеаризуемых зависимостей, когда для получения оценок коэффициентов моделей решается система линейных уравнений. Задача значительно усложняется, если для прогноза используется функциональная зависимость, не сводимая к линейной.

Метод экспоненциального сглаживания является эффективным и надежным методом среднесрочного прогнозирования.

Здесь следует остановиться более подробно на учете важности ретроспективной информации.

Практически большее значение для построения прогноза имеет информация, описывающая процесс в моменты времени, стоящие ближе к настоящему (нулевому) моменту времени. Чем дальше мы углубляемся в ретроспекцию, тем менее ценной для прогноза становится информация. Это можно учесть, придавая членам исходного динамического ряда некоторые веса, тем бо́льшие, чем ближе находится точка к началу периода прогноза.

Это положение лежит в основе метода экспоненциального сглаживания. Сущность метода заключается в сглаживании исходного динамического ряда взвешенной скользящей средней, веса́ которой (ω_i) подчиняются экспоненциальному закону (рис. 6.1).

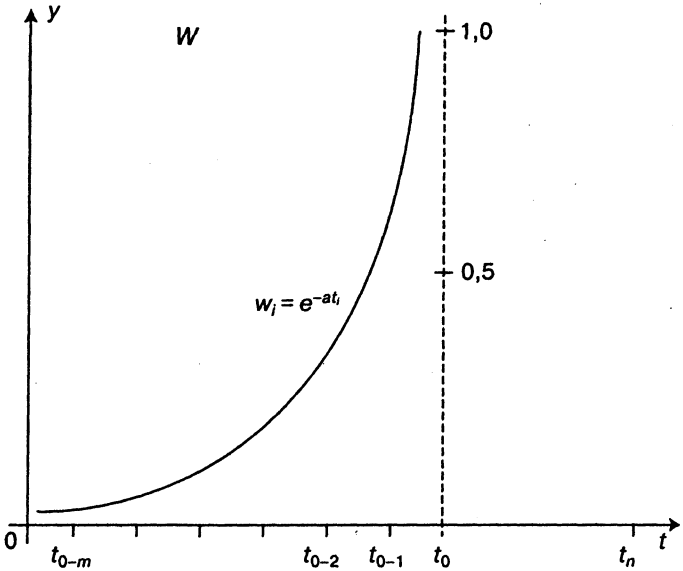


Рис. 6.1. Коэффициент экспоненциального сглаживания

Пусть исходный динамический ряд описывается полиномом следующего вида:

$$y_t = b_0 + b_1 t + \frac{b_2}{2!}t^2 + \ldots + \frac{b_p}{p!}t^p + \varepsilon_t = \sum_{j=0}^{p} \frac{b_j}{j!}t^j + \varepsilon_t, \tag{6.32}$$

где $b_0, b_1, b_2, \ldots, b_p$ – коэффициенты;
p – порядок полинома;
ε_t – случайная ошибка.

Метод экспоненциального сглаживания позволяет построить такое описание процесса (6.32), при котором более поздним наблюдениям придаются бо́льшие веса́ по сравнению с ранними наблюдениями, причем веса́ наблюдений убывают по экспоненте.

Выражение

$$S_t^{[k]}(y)=\alpha\sum_{i=0}^{n}(1-\alpha)^t S_{t-1}^{[k-1]}(y) \qquad (6.33)$$

называется *экспоненциальной средней k-го порядка* для ряда y_t, где α – параметр сглаживания.

В расчетах экспоненциальную среднюю определяют, пользуясь рекуррентной формулой, полученной Брауном:

$$S_t^{[k]}(y) = \alpha S_t^{[k-1]}(y) + (1-\alpha)S_{t-1}^{[k]}(y). \qquad (6.34)$$

Использование соотношения (6.34) предполагает задание начальных условий $S_0^{[1]}$, $S_0^{[2]}$, $S_0^{[k]}$, которые могут быть определены по формуле Брауна–Мейера:

$$S_t^{[k]}(y)=\sum_{p=0}^{n}(-1)^p\cdot\frac{\hat{b}_p}{p!}\cdot\frac{\alpha(1-\alpha)}{(k-1)!}\cdot\sum_{j=0}^{\infty}j^p(1-\alpha)^j\frac{(p-1+j)!}{j!}, \qquad (6.35)$$

где $p = 1, 2, ..., n + 1$;

$\hat{b}$ – оценки коэффициентов.

Оценки коэффициентов прогнозирующего полинома определяют через экспоненциальные средние по фундаментальной теореме Брауна–Мейера. В этом случае коэффициенты b_j находят решением системы $(p - 1)$ уравнений с $(p + 1)$ неизвестными, связывающей параметры полинома с исходной информацией.

Рассмотрим применение метода экспоненциального сглаживания для двух наиболее употребительных случаев, когда тренд описывается линейной функцией и параболой.

Линейная модель Брауна

$$y_t = b_0 + b_1 t + \varepsilon_t. \qquad (6.36)$$

Начальные приближения для случая линейного тренда равны (по формуле (6.35)):

экспоненциальная средняя 1-го порядка:

$$S_0^{[1]}(y)=b_0-\frac{1-\alpha}{\alpha}b_1; \qquad (6.37)$$

экспоненциальная средняя 2-го порядка:

$$S_0^{[2]}(y)=b_0-\frac{2(1-\alpha)}{\alpha}b_1. \qquad (6.38)$$

Зная начальные условия $S_0^{[1]}$ и $S_0^{[2]}$ и значение параметра α, вычисляют экспоненциальные средние 1-го и 2-го порядка:

$$S_t^{[1]}(y) = \alpha y_t + (1-\alpha)S_{t-1}^{[1]}(y); \tag{6.39}$$

$$S_t^{[2]}(y) = \alpha S_t^{[1]} + (1-\alpha)S_{t-1}^{[2]}(y). \tag{6.40}$$

Оценки коэффициентов линейного тренда:

$$\hat{b}_0 = 2S_t^{[1]}(y) - S_t^{[2]}(y); \tag{6.41}$$

$$\hat{b}_1 = \frac{\alpha}{1-\alpha}\left[S_t^{[1]}(y) - S_t^{[2]}(y)\right]. \tag{6.42}$$

Прогноз на l шагов (на время t_1) равен: $y_t = \hat{b}_0 + \hat{b}_1 \cdot t$.
Ошибка прогноза

$$\sigma = \sigma_{\varepsilon_t}\sqrt{\frac{\alpha}{(2-\alpha)^3}\left[1+4(1-\alpha)+5(1-\alpha)^2+2\alpha(4-3\alpha)t_\ell+2\alpha^2 t_\ell^2\right]}. \tag{6.43}$$

Параболический тренд

$$y_t = b_0 + b_1 t + \frac{1}{2}b_2 t^2 + \varepsilon_t. \tag{6.44}$$

Начальные приближения

$$S_0^{[1]}(y) = b_0 - \frac{1-\alpha}{\alpha}b_1 + \frac{(1-\alpha)(2-\alpha)}{2\alpha^2}b_2; \tag{6.45}$$

$$S_0^{[2]}(y) = b_0 - \frac{2(1-\alpha)}{\alpha}b_1 + \frac{(1-\alpha)(3-2\alpha)}{2\alpha^2}b_2; \tag{6.46}$$

$$S_0^{[3]}(y) = b_0 - \frac{3(1-\alpha)}{\alpha}b_1 + \frac{(1-\alpha)(4-3\alpha)}{2\alpha^2}b_2. \tag{6.47}$$

Экспоненциальные средние

$$S_t^{[1]}(y) = \alpha y_t + (1-\alpha)S_{t-1}^{[1]}(y); \tag{6.48}$$

$$S_t^{[2]}(y) = \alpha S_t^{[1]}(y) + (1-\alpha)S_{t-1}^{[2]}(y); \tag{6.49}$$

$$S_t^{[3]}(y) = \alpha S_t^{[2]}(y) + (1-\alpha)S_{t-1}^{[3]}(y). \tag{6.50}$$

Оценки коэффициентов параболической зависимости для тренда

$$\hat{b}_0 = 3[S_t^{[1]}(y) - S_t^{[2]}(y)] + S_t^{[3]}(y); \tag{6.51}$$

$$b_1 = \frac{\alpha}{(1-\alpha)^2}\left[(6-5\alpha)S_t^{[1]}(y) - 2(5-4\alpha)S_t^{[2]}(y) + (4-3\alpha)S_t^{[3]}(y)\right]; \tag{6.52}$$

$$\hat{b}_2 = \frac{\alpha}{(1-\alpha)^2}\left[S_t^{[1]}(y) - 2S_t^{[2]}(y) + S_t^{[3]}(y)\right]. \tag{6.53}$$

Прогноз на момент t_1

$$y_{t_\ell}^* = \hat{b}_0 + \hat{b}_1 t_\ell + \frac{1}{2}\hat{b}_2 \hat{t}_\ell^2. \tag{6.54}$$

Ошибка прогноза

$$\sigma_y^* \approx \sigma_{\varepsilon_t}\sqrt{2\alpha + 3\alpha^2 + 3\alpha^3 t_1}. \tag{6.55}$$

Для метода экспоненциального сглаживания основным и наиболее трудным моментом является выбор параметра сглаживания α, начальных условий и степени прогнозирующего полинома.

Параметр сглаживания α определяет оценки коэффициентов модели, а следовательно, результаты прогноза.

В зависимости от величины параметра прогнозные оценки поразному учитывают влияние исходного ряда наблюдений: чем больше α, тем больше вклад последних наблюдений в формирование тренда, а влияние начальных условий убывает быстро. При малом α прогнозные оценки учитывают все наблюдения, при этом уменьшение влияния более ранней информации происходит медленно.

Для приближенной оценки α известны два основных соотношения:

- *соотношение Брауна,* выведенное из условия равенства скользящей и экспоненциальной средней

$$\alpha = \frac{2}{N+1}, \tag{6.56}$$

где N – число точек ряда, для которых динамика ряда считается однородной и устойчивой (число точек в интервале сглаживания).

Иногда $\alpha = \frac{2}{n+1}$, где n – число наблюдений (точек) в ретроспективном динамическом ряду;

• *соотношение Мейера*

$$\alpha = \frac{\sigma_n}{\sigma_\varepsilon}, \tag{6.57}$$

где σ_n – средняя квадратическая ошибка модели;
σ_ε – средняя квадратическая ошибка исходного ряда.

Однако достоверно определить σ_n и σ_ε из исходной информации очень сложно, поэтому использование соотношения (6.57) затруднено.

Очевидно, что выбор параметра α нужно связывать с точностью прогноза, поэтому для более обоснованного выбора α можно использовать процедуру обобщенного сглаживания.

В этом случае получаются следующие соотношения, связывающие дисперсию прогноза $\sigma^2_{\hat{x}\tau}$ и параметр сглаживания α:

для линейной модели

$$\sigma^2_{\hat{x}\tau} = \frac{\alpha}{(1+\beta)^2}\left[1 + 4\beta + 5\beta^2 + 2\alpha(1+3\beta)\tau + 2\alpha^2\tau^3\right]\sigma^2_\varepsilon, \tag{6.58}$$

для квадратичной модели

$$\sigma^2_{\hat{x}\tau} = [2\alpha + 3\alpha^3 + 3\alpha^2\tau]\sigma_\varepsilon^2, \tag{6.59}$$

где $\beta = 1 - \alpha$;
τ – период прогноза;
σ_ε – средняя квадратическая ошибка аппроксимации исходного динамического ряда.

При $\alpha = 0$ выражения, стоящие в правых частях формул (6.58), (6.59), обращаются в нуль, следовательно, для уменьшения ошибки прогноза необходимо выбирать минимальное α.

В то же время, чем меньше α, тем ниже точность определения начальных условий, а следовательно, ухудшается и качество прогноза.

Таким образом, использование формул (6.58), (6.59) приводит к противоречию при определении параметра сглаживания. В ряде случаев параметр α выбирают так, чтобы минимизировать ошибку прогноза, рассчитанного по ретроспективной информации.

Качество прогноза во многом зависит от выбора порядка прогнозирующего полинома. Известно, что превышение второго порядка модели не приводит к существенному увеличению точности прогноза, но значительно усложняет расчет.

Отметим в заключение, что метод экспоненциального сглаживания является одним из наиболее эффективных, надежных и широко применяемых методов прогнозирования. Он позволяет получить оценку параметров тренда, характеризующих не средний уровень процесса, а тенденцию, сложившуюся к моменту последнего наблюдения, и при этом отличается простотой вычислительных операций.

Пример 6.5. Пусть задан временной ряд y_t:

Год	1997	1998	1999	2000
t y_t	1 40	2 43	3 46	4 48

Можно считать, что аппроксимирующая функция (тренд) описывается линейной функцией (рис. 6.2).

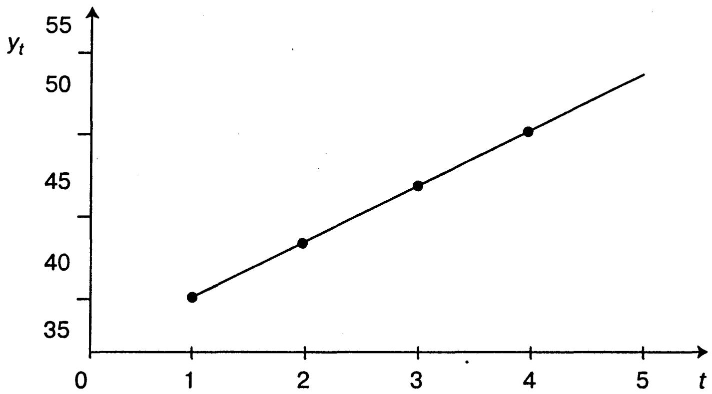


Рис. 6.2. Динамика изменения показателя y_t

Таблица 6.3

Расчетная таблица

Год	Период, t	Фактическое значение y_t	Расчетные значения			
			t^2	t_{y_t}	$y = 37{,}5 + 2{,}7 \cdot t$	$\Delta y = y - y_t$
1997	1	40	1	40	40,2	0,2
1998	2	43	4	86	42,9	–0,1
1999	3	46	9	138	45,6	–0,4
2000	4	48	16	192	48,3	0,3
Итого	10	177	30	456	–	–

1. Определим коэффициенты прямой $y = a_0 + a_1 t$ по методу наименьших квадратов. Для этого вычислим ряд промежуточных значений и их суммы. Результаты занесем в табл. 6.3.

Далее найдем

$$D_0 = \begin{vmatrix} n & \sum t_i \\ \sum t_i & \sum t_i^2 \end{vmatrix} = \begin{vmatrix} 4 & 10 \\ 10 & 30 \end{vmatrix} = 120 - 100 = 20;$$

$$D_1 = \begin{vmatrix} \sum y_i & \sum t_i \\ \sum y_i t_i & \sum t_i^2 \end{vmatrix} = \begin{vmatrix} 177 & 10 \\ 456 & 30 \end{vmatrix} = 750;$$

$$D_2 = \begin{vmatrix} n & \sum t_i \\ \sum t_i & \sum t_i y_i \end{vmatrix} = \begin{vmatrix} 4 & 177 \\ 10 & 456 \end{vmatrix} = 54;$$

$$a_0 = \frac{D_1}{D_0} = \frac{750}{20} = 37{,}5;$$

$$a_1 = \frac{D_2}{D_0} = \frac{54}{20} = 2{,}7.$$

Окончательно уравнение прямой имеет вид

$$y = 37{,}5 + 2{,}7t.$$

Подставив в него значения $t = 1, 2, 3, 4$, получим расчетные значения тренда (табл. 6.3).

Основная ошибка

$$\sigma_t = \sqrt{\frac{0{,}2^2 + (-0{,}1)^2 + (-0{,}4)^2 + 0{,}3^2}{3}} = 0{,}3.$$

2. Параметр сглаживания

$$\alpha = \frac{2}{N+1} = \frac{2}{4+1} = 0{,}4.$$

3. Начальные условия

$$S_0^{[1]} = 37{,}5 - \frac{1-0{,}4}{0{,}4} \cdot 2{,}7 = 33{,}45;$$

$$S_0^{[2]} = 37{,}5 - \frac{2(1-0{,}4)}{0{,}4} \cdot 2{,}7 = 29{,}4.$$

4. Для $t = 2$ вычислим экспоненциальные средние:

$$S_2^{[1]} = 0{,}4 \cdot 40 + 0{,}6 \cdot 33{,}45 = 36;$$

$$S_2^{[2]} = 0{,}4 \cdot 36 + 0{,}6 \cdot 29{,}4 = 32;$$

значения коэффициентов:

$$\hat{a}_0 = 2 \cdot 36 - 32 = 40,$$

$$\hat{a}_1 = \frac{0{,}4}{1-0{,}4} \cdot (36-32) = 2{,}6;$$

прогнозируемые значения:

$$y_2^* = 40 + 2{,}6 \cdot 1 = 42{,}6;$$

отклонения от фактического значения:

$$\Delta y_2^* = 42{,}6 - 43 = -0{,}4.$$

Аналогичные вычисления выполним для $t = 3$,

$t = 4$, $t = 5$.

Результаты представим в табл. 6.4.

Таблица 6.4

Типовая таблица для построения прогноза по методу экспоненциального сглаживания

Год	Период, t	Фактическое значение y_t	Расчетные значения					
			$S_t^{[1]}$	$S_t^{[2]}$	$\hat{a}_0$	$\hat{a}_1$	y_t^*	$\Delta y^* = y_t^* - y_t$
1997	1	40						
1998	2	43	36	32	40	2,6	42,6	−0,4
1999	3	46	38,6	34,6	42,6	2,7	45,3	−0,7
2000	4	48	41,6	37,4	45,8	2,8	48,6	0,6
2001	$\ell = 1$	–	44,2	40,1	48,3	2,7	51	–

Для $t = 3$:

$$S_3^{[1]} = 0{,}4 \cdot 43 + 0{,}6 \cdot 33 = 38{,}6;$$
$$S_3^{[2]} = 0{,}4 \cdot 38{,}6 + 0{,}6 \cdot 32 = 34{,}6;$$
$$\hat{a}_0 = 42{,}6;\ \hat{a}_1 = 2{,}7;$$
$$y_3^* = 42{,}6 + 2{,}7 \cdot 1 = 45{,}3;$$
$$\Delta y^* = -\ 0{,}7.$$

Для $t = 4$:

$$S_4^{[1]} = 0{,}4 \cdot 46 + 0{,}6 \cdot 38{,}6 = 41{,}6;$$
$$S_4^{[2]} = 0{,}4 \cdot 41{,}6 + 0{,}6 \cdot 34{,}6 = 37{,}4;$$
$$\hat{a}_0 = 45{,}8;\ \hat{a}_1 = 2{,}8;$$
$$y_4^* = 45{,}8 + 2{,}8 \cdot 1 = 48{,}6;$$
$$\Delta y_4^* = 0{,}6.$$

Для построения модели прогноза ($\ell = 1$) определим

$$S_5^{[1]} = 0{,}4 \cdot 48 + 0{,}6 \cdot 41{,}6 = 44{,}2;$$
$$S_5^{[2]} = 0{,}4 \cdot 44{,}2 + 0{,}6 \cdot 37{,}4 = 40{,}1;$$
$$\hat{a}_0 = 48{,}3;\ \hat{a}_1 = 2{,}7.$$

Окончательная модель прогноза имеет вид

$$y_{t+\ell} = 48{,}3 + 2{,}7 \cdot \ell,$$

где $\ell = 1, 2, \ldots$ (что соответствует $t = 5$, $t = 6$ и т. д.).

Ошибка прогноза

$$\sigma_{y^*} = 0{,}3\sqrt{\frac{0{,}4}{1{,}6^3}\left[1 + 4 \cdot 0{,}6 + 50 \cdot 0{,}6^2 + 0{,}8 \cdot 2{,}8 \cdot 1 + 0{,}32 \cdot 1^2\right]} = 0{,}3 \cdot 0{,}87 = 0{,}46.$$

Е. Выбор математической модели прогнозирования.

Выбор моделей прогнозирования базируется на оценке их качества. Независимо от метода оценки параметров моделей экстраполяции (прогнозирования) их качество определяется на основе исследования свойств остаточной компоненты: $(y_i - y_{\text{т}i})$, $i = \overline{1,n}$, т. е. величины расхождений на участке аппроксимации (построения модели) между фактическими уровнями и их расчетными значениями.

Качество модели определяется ее адекватностью исследуемому процессу и точностью. Адекватность характеризуется наличием и учетом определенных статистических свойств, а точность – степенью близости к фактическим данным. Модель прогнозирования будет считаться лучшей со статистической точки зрения, если она является адекватной и более точно описывает исходный динамический ряд.

Модель прогнозирования считается адекватной, если она учитывает существенную закономерность исследуемого процесса. В ином случае ее нельзя применять для анализа и прогнозирования.

Закономерность исследуемого процесса находит отражение в наличии определенных статистических свойств остаточной компоненты, а именно: независимости уровней, их случайности, соответствия нормальному закону распределения и равенства нулю средней ошибки.

Независимость остаточной компоненты означает отсутствие автокорреляции между остатками $(y_i - y_{\text{т}i})$.

Перечислим *последствия, вызываемые автокорреляцией остатков*[1]:

1) недооценка дисперсии остатков функции регрессии;

2) наличие ошибки при оценке выборочной дисперсии параметров регрессии.

Ошибки в вычислении дисперсий – препятствие к корректному применению метода наименьших квадратов при построении модели исходного динамического ряда.

Очевидно, важно иметь критерий, позволяющий устанавливать наличие автокорреляции. Таким критерием является *критерий Дарбина–Уотсона*, в соответствии с которым вычисляется статистика d:

$$d = \frac{\sum_{i=2}^{n}\left[(y_i - y_{\text{т}i}) - (y_{i-1} - y_{\text{т}i-1}\right]^2}{\sum_{i=1}^{n}(y_i - y_{\text{т}i})^2}, \qquad (6.60)$$

где y_i; y_{i-1} – уровни фактического динамического ряда;
$y_{\text{т}i}$; $y_{\text{т}i-1}$ – теоретические (прогнозные) уровни динамического ряда;
n – объем выборки.

Возможные значения статистики лежат в интервале $0 \le d \le 4$. Согласно методу Дарбина и Уотсона существует верхний $d_в$ и ниж-

[1] См.: *Ферстер Э., Ренц Б.* Методы корреляционного и регрессионного анализа. – М.: Финансы и статистика, 1983.

ний $d_н$ пределы значений статистики d. Эти критические значения зависят от уровня значимости α, объема выборки n и числа объясняющих переменных m (для трендовых моделей $m = 1$). В приложении 4 приведены значения $d_в$ и $d_н$ для 5 %-го уровня значимости при n от 15 до 100 и числе объясняющих переменных от 1 до 5.

Вычисленное по (6.60) значение d сравнивается с $d_в$ и $d_н$, найденными по приложению 4. При этом руководствуются следующими правилами:

1) $d_в \le d \le 4 - d_в$	принимается гипотеза: автокорреляция отсутствует;
2) $0 \le d \le d_н$	принимается гипотеза о существовании положительной автокорреляции остатков;
3) $d_н \le d \le d_в$ и $4 - d_в \le d \le 4 - d_н$	при выбранном уровне значимости нельзя прийти к определенному выводу;
4) $4 - d_н \le d \le 4$	принимается гипотеза о существовании отрицательной автокорреляции остатков.

Критерий Дарбина-Уотсона обладает двумя недостатками. Первый из них – наличие области неопределенности, в которой с помощью данного критерия нельзя прийти ни к какому решению. Второй недостаток заключается в том, что при объеме выборки меньше 15 для d не существует критических значений $d_н$ и $d_в$. В этом случае для оценки независимости уровней ряда можно использовать коэффициент автокорреляции r_a. Данный показатель приближенно можно вычислить по формуле

$$r_a = 1 - \frac{d}{2}, \tag{6.61}$$

где d – статистика Дарбина-Уотсона.

Расчетное значение r_a сравнивают с табличным $r_{aт}$ (приложение 3). Критическое значение коэффициента автокорреляции $r_{aт}$ имеет одну степень свободы,т. е. $f = n$. Если $r_a \le r_{aт}$, то уровни динамического ряда независимы.

Для проверки случайности уровней ряда можно использовать критерий поворотных точек, который называется также критерием «пиков» и «впадин». В соответствии с этим критерием каждый уровень ряда сравнивается с двумя соединенными с ним. Если он больше или меньше их, то эта точка считается поворотной. Далее подсчитывается сумма поворотных точек K.

В случайном ряду чисел должно выполняться строгое неравенство:

$$K > \left[\frac{(2n-4)}{3} - 2\sqrt{\left(\frac{16n-29}{90} \right)} \right]. \qquad (6.62)$$

Соответствие ряда остатков нормальному закону распределения важно с точки зрения правомерности построения интервалов прогноза. Основными свойствами ряда остатков являются их симметричность относительно тренда и преобладание малых по абсолютной величине ошибок над бо́льшими. В этой связи определяется близость к соответствующим параметрам нормального закона распределения коэффициентов асимметрии – A_c (мера «скошенности») и эксцесса – $Э_k$ (мера «скученности») наблюдений около модели, т. е.

$$A_c = \frac{\left[\frac{1}{n} \sum_{i=1}^{n} (y_i - y_{\text{т}i})^3 \right]}{\sqrt{\left[\frac{1}{n} \sum_{i=1}^{n} (y_i - y_{\text{т}i}) \right]^3}}; \qquad (6.63)$$

$$Э_k = \frac{\left(\frac{1}{n} \sum_{i=1}^{n} (y_i - y_{\text{т}i})^4 \right)}{\left(\frac{1}{n} \sum_{i=1}^{n} (y_i - y_{\text{т}i})^2 \right)^2} - 3. \qquad (6.64)$$

Если эти коэффициенты близки к нулю или равны нулю, то ряд остатков распределен в соответствии с нормальным законом. Для оценки степени их близости к нулю вычисляют средние квадратические отклонения:

$$S_a = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}};$$

$$S_э = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}. \qquad (6.65)$$

Если выполняются соотношения:

$$|A_c| \le 1{,}5 S_a;$$

$$|Э_k| \le 1{,}5 S_э,$$

то считается, что распределение ряда остатков не противоречит нормальному закону. В случае когда

$$|A_c| > 2S_a \text{ или } |Э_k| > 2S_э,$$

то распределение ряда не соответствует нормальному закону распределения, и построение доверительных интервалов прогноза неправомочно. В случае попадания A_c и $Э_k$ в зону неопределенности (между полутора и двумя среднеквадратическими отклонениями) может быть использован RS-критерий:

$$RS = (E_{\max} - E_{\min})/S, \tag{6.66}$$

где $E_{\max}$ – максимальный уровень ряда остатков $(y_i - y_{\text{т}i})$, $i = \overline{1, n}$;
$E_{\min}$ – минимальный уровень ряда остатков $(y_i - y_{\text{т}i})$, $i = \overline{1, n}$;
S – среднее квадратическое отклонение остатков.

Если значение этого критерия попадает между табулированными границами с заданным уровнем значимости (приложение 5), то гипотеза о нормальном распределении ряда остатков принимается.

Равенство нулю средней ошибки (математическое ожидание случайной последовательности) проверяют с помощью t-критерия Стьюдента:

$$t_p = \left| \frac{1}{n} \sum_{i=1}^{n} (y_i - y_{\text{т}i}) \right| \frac{\sqrt{n}}{S}. \tag{6.67}$$

Гипотеза равенства нулю средней ошибки отклоняется, если t_p больше табличного уровня t-критерия с $f_1 = (n - 1)$ степенями свободы и выбранным уровнем значимости α (приложение 1).

После проверки всех моделей прогнозирования из выбранного массива на адекватность необходимо выполнить оценку их точности.

В статистическом анализе известно большое число характеристик точности[1].

Наиболее часто в практической работе встречаются следующие характеристики.

---

[1] *Чуев Ю. В., Михайлов Ю. Б., Кузьмин В. И.* Прогнозирование количественных характеристик процессов. – М.: Советское радио, 1975.

1. Оценка стандартной ошибки:

$$S_{1,f(x)} = \sqrt{\frac{\sum_{i=1}^{n}\left[y_i - f(x_i)\right]^2}{n-p}}, \tag{6.68}$$

где n – число наблюдений;

p – число определяемых коэффициентов модели.

2. Средняя относительная ошибка оценки

$$\overline{m}_\alpha = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i - f(x_i)}{f(x_i)}\cdot 100\%. \tag{6.69}$$

3. Среднее линейное отклонение

$$B = \frac{\sum_{i=1}^{n}\left|y_i - f(x_i)\right|}{\sqrt{n(n-1)}}. \tag{6.70}$$

4. Ширина доверительного интервала в точке прогноза.

Для получения данной статистической оценки определим доверительный интервал в прогнозируемом периоде, т.е. возможные отклонения прогноза от основной тенденции протекания рассматриваемого процесса. Для решения этой задачи построим интервальные оценки параметров регрессии a_0 и a_1 в формах:

$$a_0 = \hat{a}_0 \pm t_p\ .\ \sigma_{\hat{a}_0},\ a_1 = \hat{a}_1 \pm t_p\ .\ \sigma_{\hat{a}_1}. \tag{6.71}$$

Здесь серединами интервалов являются точечные оценки $\hat{a}_0$ и $\hat{a}_1$, рассчитанные с помощью метода наименьших квадратов. Величина t_p – теоретическое значение критерия Стьюдента при уровне значимости, равном 5%, и числе степеней свободы, равном $V_1 = n - m - 1$ (приложение 1).

Стандартные ошибки коэффициентов регрессии $\sigma_{\hat{a}_0}$ и $\sigma_{\hat{a}_1}$ вычисляются по следующим формулам:

$$\sigma_{\hat{a}_0} = \sqrt{\frac{\sigma^2}{n}};$$

$$\sigma_{\hat{a}_1} = \sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}}. \tag{6.72}$$

Несмещенная оценка дисперсии случайной составляющей

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - y_{\text{т}i})^2, \tag{6.73}$$

где x_i, y_i – фактические значения динамических рядов x и y;
$y_{\text{т}i}$ – теоретическое значение, рассчитанное по уравнению регрессии;
$\bar{x}$ – среднее значение фактора x.

Верхняя $Y^в$ и нижняя $Y^н$ границы доверительного интервала в точке прогноза будут равны:

$$Y^в = f(a_0^в;\ a_1^в;\ x_n);$$
$$Y^н = f(a_0^н;\ a_1^н;\ x_n), \tag{6.74}$$

где $a_0^в$; $a_0^н$ – верхнее и нижнее значения параметра a_0 модели прогноза;
$a_1^в$; a_1 – верхнее и нижнее значения параметра a_1 модели прогноза;
x_n – значение фактора времени в точке прогноза.

Ширина доверительного интервала в точке прогноза

$$\Delta = Y^в - Y^н. \tag{6.75}$$

Надо отметить, что ширина доверительного интервала зависит:

- от числа степеней свободы и тем самым от объема выборки, т. е. чем больше объем выборки, тем меньше (при прочих равных условиях) значение критерия t и, следовательно, уже доверительный интервал;
- от величины стандартной ошибки оценки параметра регрессии ($\sigma_{\hat{a}_1}$ и $\sigma_{\hat{a}_0}$). Чем меньше $\sigma_{\hat{a}_1}$ и $\sigma_{\hat{a}_0}$, тем меньше при равных условиях ширина доверительного интервала.

Лучшей по точности считается та модель, у которой все перечисленные характеристики имеют меньшую величину. Однако эти показатели по-разному отражают степень точности модели и поэтому нередко дают противоречивые выводы. Для однозначного выбора лучшей модели исследователь должен воспользоваться либо одним основным показателем, либо обобщенным критерием.

Итогом работ по выбору вида математической модели прогноза является формирование ее обобщенных характеристик. В обобщенную характеристику должны быть включены вид уравнения регрессии, значения его параметров, оценки точности и адекватности модели и сами прогнозные оценки, точечные и интервальные.

6.4. Прогнозирование на основе временных рядов с использованием пакета программ для персональных ЭВМ

Как видно из вышеизложенного, вычислительные процедуры рассмотренных методов прогнозирования громоздки и трудоемки. Задача исследователя значительно облегчается при использовании пакета прикладных программ для ПЭВМ, в основу которых положены рассмотренные в подразд. 6.3 алгоритмы прогнозирования методами наименьших квадратов и экспоненциального сглаживания.

Исходная информация предоставляется в виде временного ряда значений прогнозируемого показателя y_t, где $t = 1, 2, 3, ..., N$.

Прогноз осуществляется по специально подобранной зависимости на ℓ шагов вперед, т. е. определяется y_{t+1}, где $\ell = 1, 2, 3..., M$.

Программы используют следующие виды прогнозов:

- экстраполяцию с использованием 16 двухпараметрических зависимостей;
- экспоненциальное сглаживание;
- комбинированный прогноз, включающий прогнозы на основе экстраполяции и экспоненциального сглаживания.

Прогноз с использованием метода экстраполяции. Прогноз осуществляется по одной из моделей, приведенных в табл. 6.2. Параметры моделей определяются с использованием метода наименьших квадратов. Для каждой модели рассчитывается корреляционное отношение. По максимальной величине корреляционного отношения производится выбор модели при комбинированном прогнозе.

Прогноз с использованием экспоненциального сглаживания. Прогноз осуществляется по линейным и параболическим зависимостям. Параметры моделей рассчитываются по формулам, предложенным Брауном. Выбор параметра сглаживания α производится в пределах от 0,1 до 0,9 с шагом 0,1. При комбинированном прогнозе выбор параметра производится по формуле $\alpha = 2/(N + 1)$, где N – число точек временного ряда.

Комбинированный прогноз. Комбинированный прогноз включает следующие этапы:

- оценку принадлежности вариантов прогноза к одной совокупности с использованием критериев Фишера и Стьюдента;
- определение весовых коэффициентов μ_1, μ_2;

• расчет параметров комбинированного прогноза по формулам дисперсии:

среднего значения:

$$\overline{y}^*_{t+1} = \mu_1 \overline{y}^*_{1(t+1)} + \mu_2 \overline{y}^*_{2(t+1)}; \tag{6.76}$$

дисперсии

$$D^*_{t+1} = \frac{D^*_{1(t+1)} D^*_{2(t+1)}}{\left[D^*_{1(t+1)} + D^*_{2(t+1)}\right]},$$

где $\overline{y}^*_{1(t+1)}$, $\overline{y}^*_{2(t+1)}$ – средние значения прогнозируемых параметров с использованием экстраполяции и экспоненциального сглаживания;

$D^*_{1(t+1)}$, $D^*_{2(t+1)}$ – соответственно дисперсии прогнозируемых параметров.

В состав пакета программ входят:

• исполняемый файл (основное меню) MENU1.EXE;

• файл прогнозирования методом экстраполяции TREND.TBC;

• файл прогнозирования методом экспоненциального сглаживания ALFA.TBC;

• файл комбинированного прогноза KOMBI.TBC.

РАЗДЕЛ II

Оптимизационные методы и модели в управлении экономическими системами

Глава 7
Линейное программирование

Оптимизационная задача – это экономико-математическая задача, которая состоит в нахождении оптимального (максимального или минимального) значения целевой функции, причем значения переменных должны принадлежать некоторой области допустимых значений.

В самом общем виде задача математически записывается так:

$$U = f(X) \to \max;\ X \in W, \tag{7.1}$$

где $X = (x_1, x_2, ..., x_n)$;
W – область допустимых значений переменных $x_1, x_2, ..., x_n$;
$f(X)$ – целевая функция.

Для того чтобы решить задачу оптимизации, достаточно найти ее оптимальное решение, т. е. указать $X_0 \in W$ такое, что $f(X_0) \geq f(X)$ при любом $X \in W$, или для случая минимизации – $f(X_0) \leq f(X)$ при любом $X \in W$.

Оптимизационная задача является неразрешимой, если она не имеет оптимального решения. В частности, задача максимизации будет неразрешима, если целевая функция $f(X)$ не ограничена сверху на допустимом множестве W.

Методы решения оптимизационных задач зависят как от вида целевой функции $f(X)$, так и от строения допустимого множества W. Если целевая функция в задаче является функцией n переменных, то методы решения называют *методами математического программирования*.

В математическом программировании принято выделять *следующие основные задачи в зависимости от вида целевой функции $f(X)$ и от области W*:

• задачи линейного программирования, если $f(X)$ и W линейны;

• задачи целочисленного программирования, если ставится условие целочисленности переменных $x_1, x_2, ..., x_n$;

• задачи нелинейного программирования, если форма $f(X)$ носит нелинейный характер.

7.1. Задачи линейного программирования

Задачей линейного программирования называется задача исследования операций, математическая модель которой имеет вид:

$$f(X)=\sum_{j=1}^{n} c_j x_j \to \max(\min); \tag{7.2}$$

$$\sum_{j=1}^{n} a_{ij}\, x_j = b_i, \quad i \in I, \quad I \subseteq M=\{1,2,...m\}; \tag{7.3}$$

$$\sum_{j=1}^{n} a_{ij}\, x_j \le b_i, \quad i \in M; \tag{7.4}$$

$$x_j \ge 0, j \in J, J \subseteq N = \{1, 2, ..., n\}. \tag{7.5}$$

При этом система линейных уравнений (7.3) и неравенств (7.4), (7.5), определяющая допустимое множество решений задачи W, называется *системой ограничений задачи линейного программирования*, а линейная функция $f(X)$ называется *целевой функцией*, или *критерием оптимальности.*

В частном случае, если $I = \varnothing$, то система (7.3) – (7.4) состоит только из линейных неравенств, а если $I = M$, то – из линейных уравнений.

Если математическая модель задачи линейного программирования имеет вид:

$$f(X)=\sum_{j=1}^{n} c_j \cdot x_j \to \min; \tag{7.6}$$

$$\sum_{j=1}^{n} a_{ij}\, x_j = b_i, \; i=\overline{1,m}; \tag{7.7}$$

$$b_i \ge 0;$$

$$x_j \ge 0, j = \overline{1,n}, \tag{7.8}$$

то говорят, что задача представлена в *канонической форме.*

Любую задачу линейного программирования можно свести к задаче линейного программирования в канонической форме. Для этого в общем случае нужно уметь сводить задачу максимизации к задаче минимизации; переходить от ограничений неравенств к ограничениям равенств и заменять переменные, которые не подчиняются условию неотрицательности. Максимизация некоторой функции эквивалентна минимизации той же функции, взятой с противоположным знаком, и наоборот.

Правило приведения задачи линейного программирования к каноническому виду состоит в следующем:

1) если в исходной задаче требуется определить максимум линейной функции, то следует изменить знак и искать минимум этой функции;

2) если в ограничениях правая часть отрицательна, то следует умножить это ограничение на -1;

3) если среди ограничений имеются неравенства, то путем введения дополнительных неотрицательных переменных они преобразуются в равенства;

4) если некоторая переменная x_k не имеет ограничений по знаку, то она заменяется (в целевой функции и во всех ограничениях) разностью между двумя новыми неотрицательными переменными: $x_k = x'_k - x_\ell$, где ℓ – свободный индекс, $x'_k \geq 0$, $x_\ell \geq 0$.

Пример 7.1. Приведение к канонической форме задачи линейного программирования:

$$\begin{aligned}
&\min L = 2x_1 + x_2 - x_3;\\
&2x_2 - x_3 \leq 5;\\
&x_1 + x_2 - x_3 \geq -1;\\
&2x_1 - x_2 \leq -3;\\
&x_1 \leq 0,\ x_2 \geq 0,\ x_3 \geq 0.
\end{aligned}$$

Решение

Введем в каждое уравнение системы ограничений выравнивающие переменные x_4, x_5, x_6. Система запишется в виде равенств, причем в первое и третье уравнения системы ограничений переменные x_4, x_6 вводятся в левую часть со знаком «+», а во второе уравнение вводится x_5 со знаком «–».

Итак:

$$\begin{aligned}
&2x_2 - x_3 + x_4 = 5;\\
&x_1 + x_2 - x_3 - x_5 = -1;\\
&2x_1 - x_2 + x_6 = -3;\\
&x_4 \geq 0,\quad x_5 \geq 0,\quad x_6 \geq 0.
\end{aligned}$$

Свободные члены в канонической форме должны быть положительными, для этого два последних уравнения умножим на – 1:

$$\begin{aligned} &2x_2 - x_3 + x_4 = 5; \\ &-x_1 - x_2 + x_3 + x_5 = 1; \\ &-2x_1 + x_2 - x_6 = 3; \end{aligned}$$

В канонической форме записи задач линейного программирования все переменные, входящие в систему ограничений, должны быть неотрицательными. Допустим, что

$$x_1 = x'_1 - x_7, \text{ где } x'_1 \geq 0,\ x_7 \geq 0.$$

Подставляя данное выражение в систему ограничений и целевую функцию и записывая переменные в порядке возрастания индекса, получим задачу линейного программирования, представленную в канонической форме:

$$\begin{aligned} &2x_2 - x_3 + x_4 = 5; \\ &-x'_1 - x_2 + x_3 + x_5 + x_7 = 1; \\ &-2x'_1 + x_2 - x_6 + 2x_7 = 3; \\ &L\min = 2x'_1 + x_2 - x_3 - 2x_7; \\ &x'_1 \geq 0;\ x_i \geq 0,\ i = \overline{2,7}. \end{aligned}$$

7.2. Построение экономико-математических моделей задач линейного программирования

Рассмотрим процесс построения математических моделей задач линейного программирования на примерах.

Пример 7.2. Определение оптимального ассортимента продукции.

Предприятие изготавливает два вида продукции – $П_1$ и $П_2$, которая поступает в оптовую продажу. Для производства продукции используются два вида сырья – А и В. Максимально возможные запасы сырья в сутки составляют 9 и 13 единиц соответственно. Расход сырья на единицу продукции вида $П_1$ и вида $П_2$ дан в табл. 7.1.

Опыт работы показал, что суточный спрос на продукцию $П_1$ никогда не превышает спроса на продукцию $П_2$ более чем на 1 ед. Кроме того, известно, что спрос на продукцию $П_2$ никогда не превышает 2 ед. в сутки.

Таблица 7.1

Расход сырья продукции

Сырье	Расход сырья на 1 ед. продукции		Запас сырья, ед.
	$П_1$	$П_2$	
A	2	3	9
B	3	2	13

Оптовые цены единицы продукции равны: 3 д. е. – для $П_1$ и 4 д. е. для $П_2$.

Какое количество продукции каждого вида должно производить предприятие, чтобы доход от реализации продукции был максимальным?

Процесс построения математической модели для решения поставленной задачи начинается с ответов на следующие вопросы[1].

1. Для определения каких величин должна быть построена модель, т. е. как идентифицировать *переменные* данной задачи?

2. Какие *ограничения* должны быть наложены на переменные, чтобы выполнялись условия, характерные для моделируемой системы?

3. В чем состоит *цель задачи*, для достижения которой из всех допустимых значений переменных нужно выбрать те, которые будут соответствовать оптимальному (наилучшему) решению задачи?

Ответы на вышеперечисленные вопросы могут быть сформулированы для данной задачи так: фирме требуется определить объемы производства каждого вида продукции в тоннах, максимизирующие доход в д. е. от реализации продукции, с учетом ограничений на спрос и расход исходных продуктов.

Для построения математической модели остается только идентифицировать переменные и представить цель и ограничения в виде математических функций этих переменных.

Предположим, что предприятие изготовит x_1 единиц продукции $П_1$ и x_2 единиц продукции $П_2$. Поскольку производство продукции $П_1$ и $П_2$ ограничено имеющимися в распоряжении предприятия сырьем каждого вида и спросом на данную продукцию, а также учитывая, что количество изготовляемых изделий не может быть отрицательным, должны выполняться следующие неравенства:

[1] *Таха Х.* Введение в исследование операций: В 2-х кн. – М.: Мир, 1985.

$$2x_1 + 3x_2 \le 9;$$
$$3x_1 + 2x_2 \le 13;$$
$$x_1 - x_2 \le 1;$$
$$x_2 \le 2;$$
$$x_1 \ge 0;$$
$$x_2 \ge 0.$$

Доход от реализации x_1 единиц продукции $П_1$ и x_2 единиц продукции $П_2$ составит $F = 3x_1 + 4x_2$.

Таким образом, мы приходим к следующей математической задаче: среди всех неотрицательных решений данной системы линейных неравенств требуется найти такое, при котором функция F принимает максимальное значения F_{max}.

Рассмотренная задача относится к разряду типовых задач оптимизации производственной программы предприятия. В качестве критериев оптимальности в этих задачах могут быть также использованы: прибыль, себестоимость, номенклатура производимой продукции и затраты станочного времени.

Пример 7.3. Использование мощностей оборудования.

Предприятие имеет m моделей машин различных мощностей. Задан план по времени и номенклатуре: T – время работы каждой машины; продукции j-го вида должно быть выпущено не менее N_j единиц.

Необходимо составить такой план работы оборудования, чтобы обеспечить минимальные затраты на производство, если известны производительность каждой i-й машины по выпуску j-го вида продукции b_{ij} и стоимость единицы времени, затрачиваемого i-й машиной на выпуск j-го вида продукции c_{ij}.

Другими словами, задача для предприятия состоит в следующем: требуется определить время работы i-й машины по выпуску j-го вида продукции x_{ij}, обеспечивающее минимальные затраты на производство при соблюдении ограничений по общему времени работы машин T и заданному количеству продукции N_j.

По условию задачи машины работают заданное время T, поэтому данное ограничение можно представить в следующем виде:

$$\sum_{j=1}^{n} x_{ij} = T, \quad i = \overline{1,m}. \tag{7.9}$$

Ограничение по заданному количеству продукции выглядит следующим образом:

$$\sum_{i=1}^{m} b_{ij} x_{ij} \ge N_j, \; j = \overline{1,n}. \tag{7.10}$$

Задача решается на минимум затрат на производство:

$$Z_{\min} = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}. \qquad (7.11)$$

Необходимо также учесть неотрицательность переменных $x_{ij} \geq 0$.

Задача поставлена так, чтобы израсходовать все отведенное время работы машины, т. е. обеспечить полную загрузку машины. При этом количество выпускаемой продукции каждого вида должно быть по крайней мере не менее N_j. Однако в некоторых случаях не допускается превышение плана по номенклатуре, тогда ограничения математической модели изменяются следующим образом:

$$\sum_{j=1}^{n} x_{ij} \leq T, \quad i = \overline{1,m}; \qquad (7.12)$$

$$\sum_{i=1}^{m} b_{ij} x_{ij} = N_j, \quad j = \overline{1,n}; \qquad (7.13)$$

$$x_{ij} \geq 0;$$

$$Z_{\min} = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}. \qquad (7.14)$$

Пример 7.4. Минимизация дисбаланса на линии сборки.

Промышленная фирма производит изделие, представляющее собой сборку из m различных узлов. Эти узлы изготавливаются на n заводах.

Из-за различий в составе технологического оборудования производительность заводов по выпуску j-го узла неодинакова и равна b_{ij}. Каждый i-й завод располагает максимальным суммарным ресурсом времени в течение недели для производства m узлов, равного величине T_i.

Задача состоит в максимизации выпуска изделий, что по существу эквивалентно минимизации дисбаланса, возникающего вследствие некомплектности поставки по одному или по нескольким видам узлов.

В данной задаче требуется определить еженедельные затраты времени (в часах) на производство j-го узла на i-м заводе, не превышающие в сумме временны́е ресурсы i-го завода и обеспечивающие максимальный выпуск изделий.

Пусть x_{ij} – недельный фонд времени (в часах), выделяемый на заводе i для производства узла j. Тогда объемы производства узла j будут следующими:

$$\sum_{i=1}^{n} b_{ij} x_{ij}, \quad j = \overline{1,n}. \tag{7.15}$$

Так как в конечной сборке каждый из комплектующих узлов представлен в одном экземпляре, количество конечных изделий должно быть равно количеству комплектующих узлов, объем производства которых минимален:

$$\min\left(\sum_{i=1}^{n} b_{ij} x_{ij},\ j = \overline{1,m}\right). \tag{7.16}$$

Условие рассматриваемой задачи устанавливает ограничение на фонд времени, которым располагает завод i.

Таким образом, математическая модель может быть представлена в следующем виде.

Максимизируем

$$Z = \min\left(\sum_{i=1}^{n} b_{ij} x_{ij},\ j = \overline{1,m}\right); \tag{7.17}$$

$$\sum_{j=1}^{m} x_{ij} \le T_i, \quad i = \overline{1,n}; \tag{7.18}$$

$x_{ij} \ge 0$ для всех i и j.

Эта модель не является линейной, но ее можно привести к линейной форме с помощью простого преобразования. Пусть Y – количество изделий:

$$Y = \min\left(\sum_{i=1}^{n} b_{ij} x_{ij} = N_j,\ j = \overline{1,m}\right). \tag{7.19}$$

Этому выражению с математической точки зрения эквивалентна следующая формулировка: максимизировать $Z = Y$ при ограничениях

$$\sum_{i=1}^{n} b_{ij} x_{ij} - Y \ge 0, \quad j = \overline{1,m}; \tag{7.20}$$

$$\sum_{j=1}^{m} x_{ij} \le T_i, \quad i = \overline{1,n}; \tag{7.21}$$

$x_{ij} \ge 0$ для всех i и j; $Y \ge 0$.

Пример 7.5. Задача составления кормовой смеси, или задача о диете[1].

Пусть крупная фирма (условно назовем ее «Суперрацион») имеет возможность покупать m различных видов сырья и приготавливать различные виды смесей (продуктов). Каждый вид сырья содержит разное количество питательных компонентов (ингредиентов).

Лабораторией фирмы установлено, что продукция должна удовлетворять по крайней мере некоторым минимальным требованиям с точки зрения питательности (полезности). Перед руководством фирмы стоит задача определить количество каждого i-го сырья, образующего смесь минимальной стоимости при соблюдении требований к общему расходу смеси и ее питательности.

Решение

Введем условные обозначения:

x_i – количество i-го сырья в смеси;

m – количество видов сырья;

n – количество ингредиентов в сырье;

a_{ij} – количество ингредиента j, содержащегося в единице i-го вида сырья;

b_j – минимальное количество ингредиента j, содержащегося в единице смеси;

c_i – стоимость единицы сырья i;

q – минимальный общий вес смеси, используемый фирмой.

Задача может быть представлена в виде

$$Z_{\min} = \sum_{i=1}^{m} c_i x_i, \tag{7.22}$$

при следующих ограничениях:

на общий расход смеси:

$$\sum_{i=1}^{m} x_i \geq q; \tag{7.23}$$

на питательность смеси:

$$\sum_{i=1}^{m} a_{ij} x_i \geq b_j \sum_{i=1}^{m} x_i, \quad j = \overline{1,n}; \tag{7.24}$$

на неотрицательность переменных:

$$x_i \geq 0,\ i = \overline{1,m}. \tag{7.25}$$

---

[1] *Вентцель Е. С., Овчаров Л. А.* Прикладные задачи теории вероятностей. – М.: Радио и связь, 1983.

Пример 7.6. Задача составления жидких смесей.

Еще один класс моделей, аналогичных рассмотренным выше, возникает при решении экономической проблемы, связанной с изготовлением смесей различных жидкостей с целью получения пользующихся спросом готовых продуктов.

Представим себе фирму, торгующую различного рода химическими продуктами, каждый из которых является смесью нескольких компонентов. Предположим, что эта фирма планирует изготовление смесей m-видов. Обозначим подлежащее определению количество литров i-го химического компонента, используемого для получения j-го продукта через x_{ij}. Будем предполагать, что

$$x_{ij} \geq 0,\ i = \overline{1,n},\ j = \overline{1,m}.$$

Первая группа ограничений относится к объемам потребляемых химических компонентов:

$$\sum_{i=1}^{m} x_{ij} \leq S_i, \quad i = \overline{1,n}, \tag{7.26}$$

где S_i – объем i-го химического компонента, которым располагает фирма в начале планируемого периода.

Вторая группа ограничений отражает требование, заключающееся в том, чтобы запланированный выпуск продукции хотя бы в минимальной степени удовлетворял имеющийся спрос на каждый из химических продуктов, т. е.

$$\sum_{i=1}^{n} x_{ij} \geq D_j, \quad j = \overline{1,m}, \tag{7.27}$$

где D_j – минимальный спрос на продукцию j в течение планируемого периода.

Третья группа ограничений связана с технологическими особенностями, которые необходимо принимать во внимание при приготовлении смеси, например простое ограничение, определяемое некоторыми минимально допустимыми значениями, отношения между объемами двух химических компонентов в процессе получения продукта j:

$$\frac{x_{ij}}{x_{i+1\,j}} \geq r \text{ или } x_{ij} - r\,.\,x_{i+1j} \geq 0,$$

где r – некоторая заданная константа.

Обозначив через P_{ij} доход с единицы продукции x_{ij}, запишем целевую функцию:

$$Z_{\max} = \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} x_{ij}. \qquad (7.28)$$

Пример 7.7. Задача о раскрое или о минимизации обрезков.

Данная задача состоит в разработке таких технологических планов раскроя, при которых получается необходимый комплекс заготовок, а отходы (по длине, площади, объему, массе или стоимости) сводятся к минимуму.

Например, продукция бумажной фирмы выпускается в виде бумажных рулонов стандартной ширины L. По специальным заказам потребителей фирма поставляет рулоны других размеров, для этого производится разрезание стандартных рулонов. Типичные заказы на рулоны нестандартных размеров могут включать m видов шириной $\ell_i (i = \overline{1,m})$. Известна потребность в нестандартных рулонах каждого вида, она равна b_i. Возможны n различных вариантов построения технологической карты раскроя рулонов стандартной ширины L на рулоны длиной ℓ_i.

Обозначим через a_{ij} количество рулонов i-го вида, получаемых при раскрое единицы стандартного рулона по j-му варианту. При каждом варианте раскроя на каждый стандартный рулон возможны потери, равные P_j. К потерям следует относить также избыточные рулоны нестандартной длины l_i, получаемые при различных вариантах раскроя y_{ij}, $j = \overline{1,n}$.

В качестве переменных следует идентифицировать количество стандартных рулонов, которые должны быть разрезаны при j-м варианте раскроя. Определим переменную следующим образом: x_j – количество стандартных рулонов, разрезаемых по варианту j, $j = \overline{1,n}$.

Целевая функция – минимум отходов при раскрое

$$Z_{\min} = \sum_{j=1}^{n} P_j x_j + \sum_{i=1}^{m} \sum_{j=1}^{n} y_{ij}. \qquad (7.29)$$

Ограничение на удовлетворение спроса потребителя

$$b_i = \sum_{j=1}^{n} a_{ij} x_j - \sum_{j=1}^{n} y_{ij}, \quad x_j \ge 0, \quad y_{ij} \ge 0. \qquad (7.30)$$

Пример 7.8. Многосторонний коммерческий арбитраж[1].

В сфере деятельности, связанной с валютными и биржевыми операциями, а также коммерческими сделками контрактного характера, возможны различного рода трансакции, позволяющие извлекать прибыль на разнице в курсе валют. Такого рода трансакции называются ***коммерческим арбитражем***.

Представим себе коммерсанта (условно назовем его N), имеющего возможность реализовать многосторонний коммерческий арбитраж. Предположим, что число валютных рынков, вовлеченных в трансакционную деятельность коммерсанта N, равняется шести, а максимальное число возможных трансакций равняется девяти. Подробные данные, характеризующие рассматриваемую задачу, приведены в табл. 7.2.

Таблица 7.2

Многосторонний коммерческий арбитраж

Валютный номинал	Тип трансакции									Возможность рынка
	1	2	3	4	5	6	7	8	9	
I	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}	-1	-1			≥ 0
II	-1					r_{26}			r_{29}	≥ 0
III		-1					r_{37}	r_{38}		≥ 0
IV			-1					-1		≥ 0
V				-1				r_{58}		≥ 0
VI					-1		r_{67}		-1	≥ 0
Размер трансакции	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	

При трансакции x_1 продажа единицы валютного номинала (ценных бумаг) II позволяет приобрести r_{11} единиц валютного номинала I. При трансакции x_7 взамен единицы валютного номинала I можно получить r_{37} единиц валютного номинала III и r_{67} единиц валютного номинала VI. Остальные трансакции расшифровываются аналогично. Значения r_{ij} могут быть дробными. Заметим, что при любой трансакции x_i ($i = 1, 2, 3, 4, 5$) каждый из валютных номиналов можно обменять на валютный номинал I. Следует обратить внимание на правило выбора знака перед показателями в табл. 7.2. Чтобы отличать куплю от продажи, будем соответственно использовать знаки «плюс» и «минус» перед показателями, характеризующими данную трансакцию.

---

[1] *Таха Х.* Введение в исследование операций: В 2-х кн. – М.: Мир, 1985.

Рассмотрим идеализированный случай, когда все трансакции коммерсанта N выполняются одновременно. Ограничения определяются единственным требованием – трансакция возможна лишь при условии, если коммерсант N располагает наличными ценными бумагами. Другими словами, количество проданных ценных бумаг не должно превышать количество приобретенных. Данные ограничения имеют вид

$$
\begin{aligned}
& r_{11}x_1 + r_{12}x_2 + r_{13}x_3 + r_{14}x_4 + r_{15}x_5 - x_6 - x_7 \geq 0; \\
& -x_1 + r_{26}x_6 + r_{29}x_9 \geq 0; \\
& -x_2 + r_{37}x_7 + r_{38}x_8 \geq 0; \\
& -x_3 - x_8 + r_{49}x_9 \geq 0; \\
& -x_4 + r_{58}x_8 \geq 0; \\
& -x_5 + r_{67}x_7 - x_9 \geq 0; \\
& x_j \geq 0, j = \overline{1, 9}.
\end{aligned}
$$

Пусть целевая функция представляет собой чистый доход, выраженный в единицах валютного номинала I, т. е. задача состоит в том, чтобы

$$Z_{\max} = r_{11}x_1 + r_{12}x_2 + r_{13}x_3 + r_{14}x_4 + r_{15}x_5 - x_6 - x_7.$$

Пример 7.9. Транспортная задача.

Имеется три поставщика и четыре потребителя однородной продукции. Известны затраты на перевозку груза от каждого поставщика каждому потребителю. Обозначим их c_{ij}, $i = \overline{1,3}$; $j = \overline{1,4}$. Запасы грузов у поставщиков равны a_i, $i = \overline{1,3}$. Известны потребности каждого потребителя b_j, $j = \overline{1,4}$. Будем считать, что суммарные потребности равны суммарным запасам:

$$\sum_{i=1}^{3} a_i = \sum_{j=1}^{4} b_j.$$

Требуется составить такой план перевозок, чтобы обеспечить минимальные суммарные затраты при полном удовлетворении потребностей.

Введем переменные x_{ij} – количество груза, перевозимого от i-го поставщика j-му потребителю.

Ограничения задачи выглядят следующим образом:

• потребности всех потребителей должны быть удовлетворены полностью:

$$\begin{cases} x_{11} + x_{21} + x_{31} = b_1; \\ x_{12} + x_{22} + x_{32} = b_2; \\ x_{13} + x_{23} + x_{33} = b_3; \\ x_{14} + x_{24} + x_{34} = b_4; \end{cases} \tag{7.31}$$

или в общем виде:

$$\sum_{i=1}^{3} x_{ij} = b_j, \quad j = \overline{1,4};$$

• груз от поставщика должен быть вывезен полностью:

$$\begin{cases} x_{11} + x_{12} + x_{13} + x_{14} = a_1; \\ x_{21} + x_{22} + x_{23} + x_{24} = a_2; \\ x_{31} + x_{32} + x_{33} + x_{34} = a_3; \end{cases} \tag{7.32}$$

или в общем виде:

$$\sum_{j=1}^{4} x_{ij} = a_i, \quad i = \overline{1,3};$$

• условие неотрицательности переменных:

$$x_{ij} \geq 0, \ i = \overline{1,3}, \ j = \overline{1,4}.$$

Целевая функция должна минимизировать суммарные затраты на перевозку:

$$Z_{\min} = \sum_{i=1}^{3} \sum_{j=1}^{4} c_{ij} x_{ij}. \tag{7.33}$$

Количество поставщиков и потребителей в общем случае может быть произвольным (≥ 2).

Мы рассмотрели девять примеров типовых задач линейного программирования. Обобщая их, можно сделать следующие *выводы*.

1. Ограничения в задачах линейного программирования могут быть выражены как равенствами, так и неравенствами.

2. Линейная функция может стремиться как к максимуму, так и к минимуму.

3. Переменные в задачах всегда неотрицательны.

Напомним, что от любой из вышеперечисленных задач можно перейти к канонической (основной) задаче линейного программирования.

7.3. Графическое решение задачи линейного программирования

Графический способ решения задач линейного программирования целесообразно использовать для:

- решения задач с двумя переменными, когда ограничения выражены неравенствами;
- решения задач со многими переменными при условии, что в их канонической записи содержится не более двух свободных переменных.

Запишем задачу линейного программирования с двумя переменными:

целевая функция:

$$Z_{\max} = c_1x_1 + c_2x_2; \tag{7.34}$$

ограничения:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 \le b_1; \\ a_{21}x_1 + a_{22}x_2 \le b_2; \\ \dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 \le b_m; \end{cases} \tag{7.35}$$

$$x_1 \ge 0;\ x_2 \ge 0. \tag{7.36}$$

Каждое из неравенств (7.35) – (7.36) системы ограничений задачи геометрически определяет полуплоскость соответственно с граничными прямыми $a_{i1}x_1 + a_{i2}x_2 = b_i$; $(i = \overline{1,m})$; $x_1 = 0$; $x_2 = 0$. В том случае, если система неравенств (7.35) – (7.36) совместна, область ее решений есть множество точек, принадлежащих всем указанным полуплоскостям. Так как множество точек пересечения данных полуплоскостей – выпуклое, то областью допустимых решений является выпуклое множество, которое называется *многоугольником* решений. Стороны этого многоугольника лежат на прямых, уравнения которых получаются из исходной системы ограничений заменой знаков неравенств на знаки равенств.

Областью допустимых решений системы неравенств (7.35) – (7.36) может быть:

- выпуклый многоугольник;
- выпуклая многоугольная неограниченная область;
- пустая область;
- луч;
- отрезок;
- единственная точка.

Целевая функция (7.34) определяет на плоскости семейство параллельных прямых, каждой из которых соответствует определенное значение Z.

Вектор $\overline{C} = (c_1; c_2)$ с координатами c_1 и c_2, перпендикулярный этим прямым, указывает направление наискорейшего возрастания Z, а противоположный вектор – направление убывания Z.

Если в одной и той же системе координат изобразить область допустимых решений системы неравенств (7.35) – (7.36) и семейство параллельных прямых (7.34), то задача определения максимума функции Z сведется к нахождению в допустимой области точки, через которую проходит прямая из семейства Z = const, и которая соответствует наибольшему значению параметра Z.

Эта точка существует тогда, когда многоугольник решений не пуст и на нем целевая функция ограничена сверху. При указанных условиях в одной из вершин многоугольника решений целевая функция принимает максимальное значение.

Для определения данной вершины построим линию уровня $Z = c_1x_1 + c_2x_2 = 0$, проходящую через начало координат и перпендикулярную вектору $\overline{C} = (c_1;c_2)$, и будем передвигать ее в направлении вектора $\overline{C} = (c_1;c_2)$ до тех пор, пока она не коснется последней крайней (угловой) точки многоугольника решений. Координаты указанной точки и определяют оптимальный план данной задачи.

Заканчивая рассмотрение геометрической интерпретации задачи (7.34) – (7.36), отметим, что при нахождении ее решения могут встретиться случаи, изображенные на рис. 7.1 – 7.4. Рис. 7.1 харак-

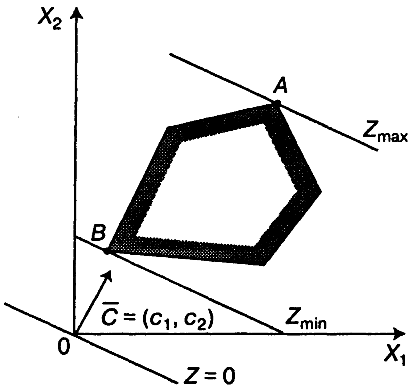


Рис. 7.1. Оптимум функции Z достижим в точке A

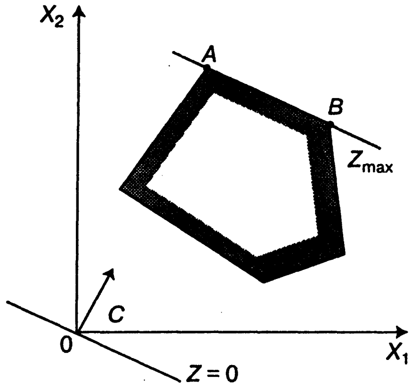


Рис. 7.2. Оптимум функции Z достигается в любой точке [AB]

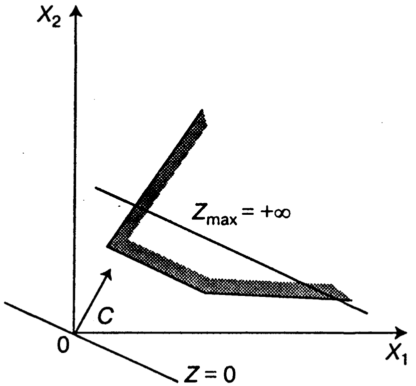


Рис. 7.3. Оптимум функции Z недостижим

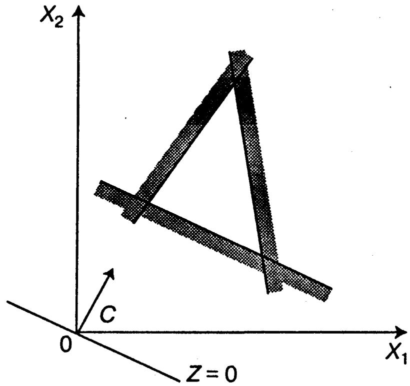


Рис. 7.4. Область допустимых решений — пустая область

теризует такой случай, когда целевая функция принимает максимальное значение в единственной точке А. Из рис. 7.2 видно, что максимальное значение целевая функция принимает в любой точке отрезка AB.

На рис. 7.3 изображен случай, когда максимум недостижим, а на рис. 7.4 – случай, когда система ограничений задачи несовместна. Отметим, что нахождение минимального значения Z при данной системе ограничений отличается от нахождения ее максимального значения при тех же ограничениях лишь тем, что линия уровня Z передвигается не в направлении вектора $\overline{C} = (c_1;c_2)$, а в противоположном направлении. Таким образом, отмеченные выше случаи, встречающиеся при нахождении максимального значения целевой функции, имеют место и при определении ее минимального значения.

Для практического решения задачи линейного программирования (7.34) – (7.36) на основе ее геометрической интерпретации необходимо следующее.

1. Построить прямые, уравнения которых получаются в результате замены в ограничениях (7.35) – (7.36) знаков неравенств на знаки равенств.

2. Найти полуплоскости, определяемые каждым из ограничений задачи.

3. Определить многоугольник решений.

4. Построить вектор $\overline{C} = (c_1;c_2)$.

5. Построить прямую $Z = c_1x_1 + c_2x_2 = 0$, проходящую через начало координат и перпендикулярную вектору $\overline{C}$.

6. Передвигать прямую $Z = c_1x_1 + c_2x_2$ в направлении вектора $\overline{C}$, в результате чего либо находят точку (точки), в которой целевая функция принимает максимальное значение, либо устанавливают неограниченность функции сверху на множестве планов.

7. Определить координаты точки максимума функции и вычислить значение целевой функции в этой точке.

Пример 7.10. Рассмотрим решение задачи об ассортименте продукции (пример 7.2) геометрическим способом.

Решение

Построим многоугольник решений (рис. 7.5). Для этого в системе координат X_10X_2 на плоскости изобразим граничные прямые:

$$\begin{aligned} 2x_1 + 3x_2 &= 9 \quad (L_1); \\ 3x_1 + 2x_2 &= 13 \quad (L_2); \\ x_1 - x_2 &= 1 \quad (L_3); \\ x_2 &= 2 \quad (L_4). \end{aligned}$$

Взяв какую-либо точку, например начало координат, установим, какую полуплоскость определяет соответствующее неравенство. Полуплоскости, определяемые неравенствами, на рис. 7.5 показаны стрелками. Областью решений является многоугольник *OABCD*.

Для построения прямой $Z = 3x_1 + 4x_2 = 0$ строим вектор-градиент $\overline{C} = (3;4)$ и через точку 0 проводим прямую, перпендикулярную ему. Построенную прямую $Z = 0$ перемещаем параллельно самой себе в направлении вектора $\overline{C}$. Из рис. 7.5 следует, что по отношению к многоугольнику решений опорной эта прямая становится в точке C, где функция принимает максимальное значение. Точка C лежит на пересечении прямых L_1 и L_3. Для определения ее координат решим систему уравнений:

$$\begin{cases} 2x_1 + 3x_2 = 9; \\ x_1 - x_2 = 1. \end{cases}$$

Оптимальный план задачи $x_1 = 2{,}4$; $x_2 = 1{,}4$. Подставляя значения x_1 и x_2 в линейную функцию, получим:

$$Z_{\max} = 3 \cdot 2{,}4 + 4 \cdot 1{,}4 = 12{,}8.$$

Полученное решение означает, что объем производства продукции $П_1$ должен быть равен 2,4 ед., а продукции $П_2$ – 1,4 ед. Доход, получаемый в этом случае, составит: $Z = 12{,}8$ д. е.

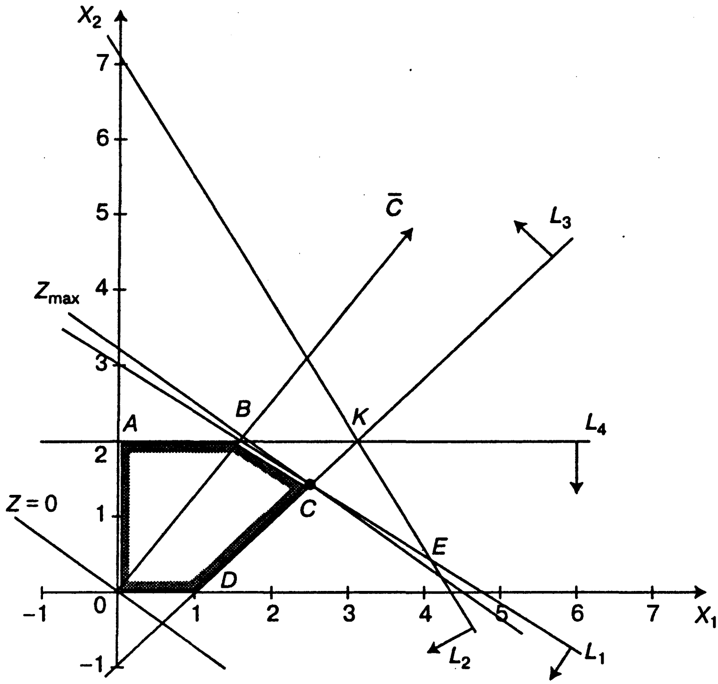


Рис. 7.5. Решение задачи линейного программирования геометрическим способом

Геометрическим способом можно также решать задачи линейного программирования с числом переменных более двух. Для этого исходную задачу преобразуют *методом Жордана–Гаусса* (приложение 11).

Пример 7.11.

$$\begin{cases} x_1 - x_2 + 4x_3 - 2x_4 = 2 \\ 3x_1 + 2x_2 - x_3 + 4x_4 = 3; \\ x_i \geq 0, \ i = \overline{1,4} \end{cases}$$

$$Z_{max} = 4x_1 - 2x_2 + x_3 - x_4.$$

Используя метод Жордана–Гаусса, произведем два полных исключения x_1 и x_4:

$$\left|\begin{array}{cccc|c|c} \boxed{1} & -1 & 4 & -2 & 2 & 4 \\ 3 & 2 & -1 & 4 & 3 & 11 \end{array}\right| \Rightarrow \left|\begin{array}{cccc|c|c} 1 & -1 & 4 & -2 & 2 & 4 \\ 0 & 5 & -13 & \boxed{10} & -3 & -1 \end{array}\right| \Rightarrow$$

$$\Rightarrow\left|\begin{array}{cccc|c|c} 1 & 0 & 1{,}4 & 0 & 1{,}4 & 3{,}2 \\ 0 & 0{,}5 & -1{,}3 & 1 & -0{,}3 & -0{,}1 \end{array}\right|.$$

В результате приходим к системе

$$\begin{cases} x_1 + 1{,}4x_3 = 1{,}4; \\ 0{,}5x_2 - 1{,}3x_3 + x_4 = -0{,}3, \end{cases}$$

откуда

$$\begin{cases} x_1 = 1{,}4 - 1{,}4x_3; \\ x_4 = -0{,}3 - 0{,}5x_2 + 1{,}3x_3. \end{cases}$$

Подставляя эти значения в линейную функцию Z и отбрасывая в последней системе базисные переменные, получим задачу, выра-

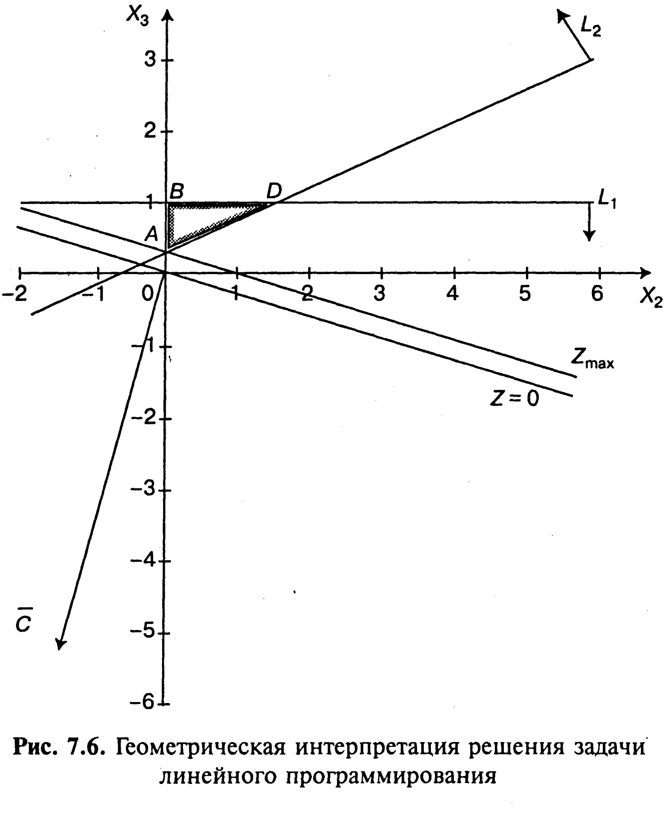


Рис. 7.6. Геометрическая интерпретация решения задачи линейного программирования

женную только через свободные неизвестные x_2 и x_3. Найдем максимальное значение линейной функции

$$Z = 5{,}9 - 5{,}9x_3 - 1{,}5x_2$$

при следующих ограничениях:

$$x_3 \le 1;\ 0{,}5x_2 - 1{,}3x_3 \le 0{,}3.$$

Построим многоугольник решений и линейную функцию в системе координат $X_2 0 X_3$ (рис. 7.6). Согласно рис. 7.6 линейная функция принимает максимальное значение в точке A, которая лежит на пересечении прямых L_2 и $X_2 = 0$. Ее координаты (0;0,23).

Максимальное значение функции

$$Z_{\max} = 5{,}9 - 5{,}9 \cdot 0{,}23 - 1{,}5 \cdot 0 \approx 4{,}54.$$

Для отыскания оптимального плана исходной задачи подставляем в преобразованную систему x_2 и x_3.Окончательно получаем $X = (1{,}078;\ 0;\ 0{,}23;\ 0{,}001)$.

7.4. Анализ моделей на чувствительность

Анализ моделей на чувствительность – это процесс, реализуемый после получения оптимального решения. В рамках такого анализа выявляется чувствительность оптимального решения к определенным изменениям исходной модели. В задаче об ассортименте продукции (пример 7.2) может представлять интерес вопрос о том, как повлияет на оптимальное решение увеличение и уменьшение спроса на продукцию или запасов исходного сырья. Возможно, также потребуется анализ влияния рыночных цен на оптимальное решение.

При таком анализе всегда рассматривается комплекс линейных оптимизационных моделей. Это придает модели определенную динамичность, позволяющую исследователю проанализировать влияние возможных изменений исходных условий на полученное ранее оптимальное решение. Динамические характеристики моделей фактически отображают аналогичные характеристики, свойственные реальным процессам. Отсутствие методов, позволяющих выявлять влияние возможных изменений параметров модели на оптимальное решение, может привести к тому, что полученное (статическое) решение устареет еще до своей реализации. Для проведения анализа модели на чувствительность с успехом могут быть использованы графические методы[1].

Рассмотрим *основные задачи анализа на чувствительность* на примере 7.10.

Задача 1. Анализ изменений запасов ресурсов.

После нахождения оптимального решения представляется вполне логичным выяснить, как отразится на оптимальном решении изменение запасов ресурсов. Для этого необходимо ответить на два вопроса:

1. На сколько можно увеличить запас некоторого ресурса для улучшения полученного оптимального значения целевой функции Z?

2. На сколько можно снизить запас некоторого ресурса при сохранении полученного оптимального значения целевой функции Z?

Прежде чем ответить на поставленные вопросы, классифицируем ограничение линейной модели как связывающие (активные) и несвязывающие (неактивные) ограничения. Прямая, представляющая связывающее ограничение, должна проходить через оптимальную точку, в противном случае, соответствующее ограничение будет несвязывающим. На рис. 7.5 связывающими ограничениями являются ограничения (1) и (3), представленные прямыми L_1 и L_3 соответственно, т. е. те, которые определяют запасы исходных ресурсов. Ограничение (1) определяет запасы сырья A. Ограничение (3) определяет соотношение спроса на выпускаемую продукцию.

Если некоторое ограничение является связывающим, то соответствующий ресурс относят к разряду дефицитных ресурсов, так как он используется полностью. Ресурс, с которым ассоциировано несвязывающее ограничение, следует отнести к разряду недефицитных ресурсов (т. е. имеющихся в некотором избытке). В нашем примере несвязывающими ограничениями являются (2) и (4). Следовательно, ресурс – сырье B – недефицитный, т. е. имеется в избытке, а спрос на продукцию $П_2$ не будет удовлетворен полностью (в таблице – ресурсы 2 и 4).

При анализе модели на чувствительность к правым частям ограничений определяются:

1) предельно допустимое увеличение запаса дефицитного ресурса, позволяющее улучшить найденное оптимальное решение;

2) предельно допустимое снижение запаса недефицитного ресурса, не изменяющее найденное ранее оптимальное значение целевой функции.

---

[1] *Таха Х.* Введение в исследование операций: В 2-х т. – М.: Мир, 1985.

В нашем примере сырье A и соотношение спроса на выпускаемую продукцию $П_1$ и $П_2$ являются дефицитными ресурсами (в табл. 7.3 – ресурсы 1, 3).

Рассмотрим сначала ресурс – сырье A. На рис. 7.7 при увеличении запаса этого ресурса прямая L_1 перемещается вверх, параллельно самой себе, до точки K, в которой пересекаются линии ограничений L_2, L_3 и L_4. В точке K ограничения (2), (3) и (4) становятся связывающими; оптимальному решению при этом соответствует точка K, а пространством (допустимых) решений становится многоугольник $AKД0$. В точке K ограничение (1) (для ресурса A) становится избыточным, так как любой дальнейший рост запаса соответствующего ресурса не влияет ни на пространство решений, ни на оптимальное решение.

Таким образом, объем ресурса A не следует увеличивать сверх того предела, когда соответствующее ему ограничение (1) становится избыточным, т. е. прямая (1) проходит через новую оптимальную точку K. Этот предельный уровень определяется следую-

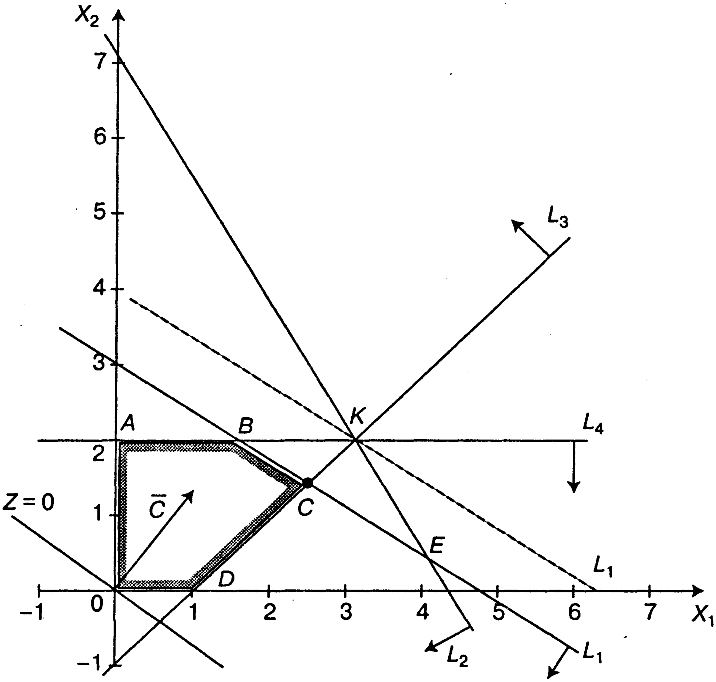


Рис. 7.7. Геометрическая интерпретация решения задачи линейного программирования (изменение ресурса A)

щим образом. Устанавливаются координаты точки K, в которой пересекаются прямые L_2, L_3 и L_4, т. е. находится решение системы уравнений

$$\begin{cases} 3x_1 + 2x_2 = 13; \\ x_1 - x_2 = 1; \\ x_2 = 2. \end{cases}$$

В результате получается $x_1 = 3$ и $x_2 = 2$. Затем, путем подстановки координат точки K в левую часть ограничения (1), определяется максимально допустимый запас ресурса A:

$$2x_1 + 3x_2 = 2 \cdot 3 + 3 \cdot 2 = 12.$$

Рис. 7.8 иллюстрирует ситуацию, когда рассматривается вопрос об изменении соотношения спроса на продукцию $П_1$ и $П_2$.

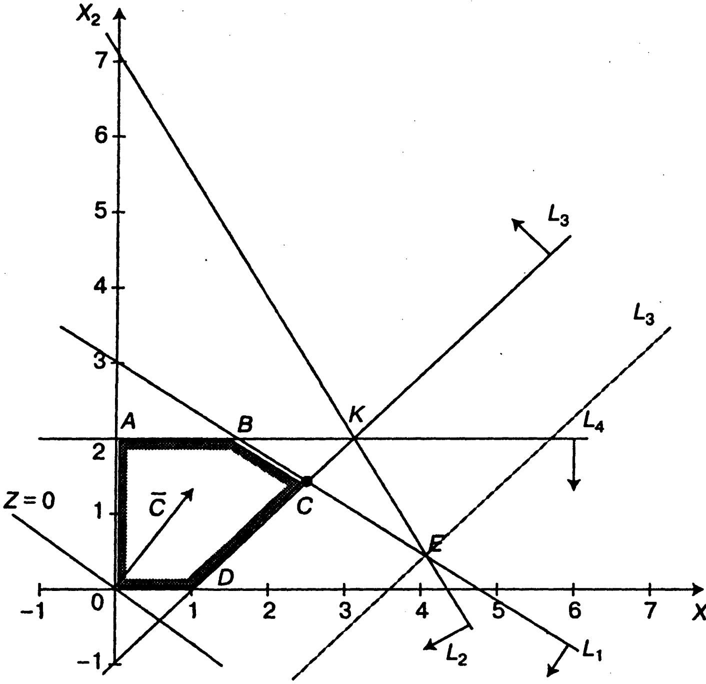


Рис. 7.8. Геометрическая интерпретация решения задачи линейного программирования (изменение спроса на продукцию)

Новой оптимальной точкой становится точка E, где пересекаются прямые L_1 и L_2. Координаты данной точки находятся путем решения системы уравнений (1) и (2) следующим образом:

$$\begin{cases} 2x_1 + 3x_2 = 9; \\ 3x_1 + 2x_2 = 13. \end{cases}$$

В результате получается $x_1 = 4{,}2$; $x_2 = 0{,}2$, причем суточный спрос на продукцию $П_1$ не должен превышать спрос на продукцию $П_2$ на величину $x_1 - x_2 = 4{,}2 - 0{,}2 = 4$ ед.

Дальнейшее увеличение разрыва в спросе на продукцию $П_1$ и $П_2$ не будет влиять на оптимальное решение.

Рассмотрим вопрос об уменьшении правой части несвязывающих ограничений. Ограничение (4) $x_2 \leq 2$ фиксирует предельный уровень спроса на продукцию $П_2$. Из рис. 7.5 следует, что, не изменяя оптимального решения, прямую L_4 (AB) можно опускать вниз до пересечения с оптимальной точкой C. Так как точка C имеет координаты $x_1 = 2{,}4$; $x_2 = 1{,}4$, уменьшение спроса на продукцию $П_2$ до величины $x_2 = 1{,}4$ никак не повлияет на оптимальность ранее полученного решения.

Рассмотрим ограничение (2) $3x_1 + 2x_2 \leq 13$, которое представляет собой ограничение на недефицитный ресурс – сырье B. И в этом случае правую часть – запасы сырья B – можно уменьшать до тех пор, пока прямая L_2 не достигнет точки C. При этом правая часть ограничения (2) станет равной $3x_1 + 2x_2 = 3 \cdot 2{,}4 + 2 \cdot 1{,}4 = 10{,}0$, что позволяет записать это ограничение в виде: $3x_1 + 2x_2 \leq 10$. Этот результат показывает, что ранее полученное оптимальное решение не изменится, если суточный запас ресурса B уменьшить на 3 ед.

Результаты проведенного анализа можно свести в табл. 7.3:

Таблица 7.3

Ресурс	Тип ресурса	Максимальное изменение запаса ресурса, ед.	Максимальное увеличение дохода от изменения ресурса, у. д. е.
1 (A)	Дефицитный	12 – 9 = +3	17 – 12,8 = +4,2
2 (B)	Недефицитный	10 – 13 = –3	12,8 – 12,8 = 0
3	Дефицитный	4 – 1 = +3	13,4 – 12,8 = +0,6
4	Недефицитный	1,4 – 2 = –0,6	12,8 – 12,8 = 0

Задача 2. Определение наиболее выгодного ресурса.

В задаче 1 анализа на чувствительность мы исследовали влияние на оптимум увеличения объема дефицитных ресурсов. При ограничениях, связанных с дополнительным привлечением ресурсов, естественно задать вопрос: какому из ресурсов следует отдать предпочтение при вложении дополнительных средств? Для этого вводится характеристика ценности каждой дополнительной единицы дефицитного ресурса, выражаемая через соответствующее приращение оптимального значения целевой функции. Такую характеристику для рассматриваемого примера можно получить непосредственно из таблицы, в которой приведены результаты решения задачи 1 на чувствительность. Обозначим ценность дополнительной единицы ресурса i через y_i. Величина y_i определяется из соотношения

$$y_i = \frac{\text{Максимальное приращение } Z}{\text{Максимально допустимый прирост ресурса } i}.$$

Результаты расчета ценности единицы каждого из ресурсов представлены в табл. 7.4:

Таблица 7.4

Ресурс i	Тип ресурса	Значение y_i
1 (A)	Дефицитный	4,2/3 = 1,4
2 (B)	Недефицитный	0/(–3) = 0
3	Дефицитный	0,6/3 = 0,2
4	Недефицитный	0/(–0,6) = 0

Полученные результаты свидетельствуют о том, что дополнительные вложения в первую очередь следует направить на увеличение ресурса A и лишь затем – на формирование соотношения спроса на продукцию $П_1$ и продукцию $П_2$. Что касается недефицитных ресурсов, то, как и следовало ожидать, их объем увеличивать не следует.

Задача 3. Определение пределов изменения коэффициентов целевой функции.

Изменение коэффициентов целевой функции оказывает влияние на наклон прямой, которая представляет эту функцию в принятой системе координат. Вариация коэффициентов целевой функции может привести к изменению совокупности связывающих ограничений и, следовательно, статуса того или иного ресурса (т. е. сделать недефицитный ресурс дефицитным, и наоборот).

При анализе модели на чувствительность рассмотрение коэффициентов целевой функции необходимо дополнить исследованием следующих вопросов:

1) каков диапазон изменения того или иного коэффициента целевой функции, при котором не происходит изменения оптимального решения?

2) на сколько следует изменить тот или иной коэффициент целевой функции, чтобы сделать некоторый недефицитный ресурс дефицитным, и, наоборот, дефицитный ресурс сделать недефицитным?

Ответим на поставленные вопросы на нашем *примере*.

Рассматривая первый вопрос, обозначим через c_1 и c_2 доходы предприятия от продажи единицы продукции $П_1$ и $П_2$ соответственно. Тогда целевую функцию можно представить в следующем виде:

$$Z = c_1x_1 + c_2x_2.$$

На рис. 7.5 видно, что при увеличении c_1 или уменьшении c_2 прямая, представляющая целевую функцию Z, вращается (вокруг точки C) по часовой стрелке. Если же c_1 уменьшается или c_2 увеличивается, эта прямая вращается в противоположном направлении – против часовой стрелки. Таким образом, точка C будет оставаться оптимальной точкой до тех пор, пока наклон прямой не выйдет за пределы, определяемые наклонами прямых для ограничений (1) и (3).

Когда наклон прямой Z станет равным наклону прямой L_1, получим две альтернативные оптимальные угловые точки – C и B. Аналогично, если наклон прямой Z станет равным наклону прямой для ограничения (3), будем иметь альтернативные оптимальные угловые точки C и D. Наличие альтернативных оптимумов свидетельствует о том, что одно и то же оптимальное значение Z может достигаться при различных значениях переменных x_1 и x_2. Как только наклон прямой выйдет за пределы указанного выше интервала c_1, получим некоторое новое оптимальное решение.

Рассмотрим на нашем примере, каким образом можно найти допустимый интервал изменения c_1, при котором точка C остается оптимальной. Исходное значение коэффициента $c_2 = 4$ оставим неизменным. На рис. 7.5 видно, что значение c_1 можно уменьшать до тех пор, пока прямая Z не совпадет с прямой L_1 (отрезок BC).

Это крайнее минимальное значение коэффициента c_1 можно определить из равенства углов наклонов прямой Z и прямой L_1. Так как тангенс угла наклона для прямой Z равен $\frac{c_1}{4}$, а для прямой (1) равен $\frac{2}{3}$, то минимальное значение c_1 определим из равенства $\frac{c_1}{4}=\frac{2}{3}$, откуда min $c_1=\frac{8}{3}$. На рис 7.5 видно, что значение c_1 можно увеличивать беспредельно, так как прямая Z при $c_2 = 4$ и $c_1 \to +\infty$ не совпадает с прямой L_3 (отрезок DC) и, следовательно, точка C при всех значениях коэффициента $c_1 \geq \frac{8}{3}$ будет единственной оптимальной.

Интервал изменения c_1, в котором точка C по-прежнему остается единственной оптимальной точкой, определяется неравенством $\frac{8}{3} < C_1 < +\infty$. При $c_1 = \frac{8}{3}$ оптимальными угловыми точками будут как точка C, так и точка B. Как только коэффициент c_1 становится меньше $\frac{8}{3}$, оптимум смещается в точку B.

Можно заметить, что, как только коэффициент c_1 оказывается меньше $\frac{8}{3}$, ресурс 3 становится недефицитным, а ресурс 4 – дефицитным. Для предприятия это означает следующее: если доход от продажи единицы продукции $П_1$ станет меньше $\frac{8}{3}$ д. е., то наиболее выгодная производственная программа предприятия должна предусматривать выпуск максимально допустимого количества продукции $П_2$ (полностью удовлетворять спрос на продукцию $П_2$).

При этом соотношение спроса на продукцию $П_1$ и $П_2$ не будет лимитировать объемы производства, что обусловит недефицитность ресурса (3). Увеличение коэффициента c_1 свыше $\frac{8}{3}$ д. е. не снимает проблему дефицита ресурсов (1) и (3). Точка C – точка пересечения прямых L_1 и L_3 – остается все время оптимальной.

7.5. Симплекс-метод

Общая идея симплекс-метода

Для начала работы требуется, чтобы заданная система ограничений выражалась равенствами, причем в этой системе ограничений должны быть выделены базисные неизвестные. Решение задачи при помощи симплекс-метода распадается на ряд шагов. На каждом шаге от данного базиса $Б$ переходят к другому, новому базису $Б^1$ с таким расчетом, чтобы значение функции Z уменьшалось, т. е. $Z_{Б^1} \le Z_Б$. Для перехода к новому базису из старого базиса удаляется одна из переменных и вместо нее вводится другая из числа свободных. После конечного числа шагов находится некоторый базис $Б^{(k)}$, для которого $Z_{Б^{(k)}}$ есть искомый минимум для линейной функции Z, а соответствующее базисное решение является оптимальным либо выясняется, что задача не имеет решения.

Алгоритм симплекс-метода

Рассмотрим систему ограничений и линейную форму вида:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2; \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m; \end{cases} \tag{7.37}$$

$$Z_{\min} = c_0 + c_1x_1 + c_2x_2 + \ldots + c_nx_n; \tag{7.38}$$

$$x_i \ge 0,\ i = \overline{1, n}. \tag{7.39}$$

Используя метод Жордана—Гаусса (приложение 11), приведем записанную систему к виду, где выделены базисные переменные.

Введем условные обозначения:

$x_1, x_2, \ldots, x_r$ – базисные переменные;

$x_{r+1}, x_{r+2}, \ldots, x_n$ – свободные переменные.

$$\begin{cases} x_1 = \beta_1 - (\alpha_{1r+1}x_{r+1} + \alpha_{1r+2}x_{r+2} + \ldots + \alpha_{1n}x_n); \\ x_2 = \beta_2 - (\alpha_{2r+1}x_{r+1} + \alpha_{2r+2}x_{r+2} + \ldots + \alpha_{2n}x_n); \\ \vdots \\ x_r = \beta_r - (\alpha_{rr+1}x_{r+1} + \alpha_{rr+2}x_{r+2} + \ldots + \alpha_{rn}x_n); \end{cases} \tag{7.40}$$

$$Z_{\min} = \gamma_0 - (\gamma_{r+1}x_{r+1} + \gamma_{r+2}x_{r+2} + \ldots + \gamma_nx_n). \tag{7.41}$$

По последней системе ограничений и целевой функции Z построим табл. 7.5.

Данная таблица называется симплекс-таблицей. Все дальнейшие преобразования связаны с изменением содержания этой таблицы.

Алгоритм симплекс-метода сводится к следующему.

1. В последней строке симплекс-таблицы находят наименьший положительный элемент, не считая свободного члена. Столбец, соответствующий этому элементу, считается разрешающим.

Таблица 7.5

Симплекс-таблица

Свободные неизвестные / Базисные неизвестные	Свободный член	x_{r+1}	x_{r+2}	...	x_n
x_1	β_1	α_{1r+1}	α_{1r+2}	...	α_{1n}
x_2	β_2	α_{2r+1}	α_{2r+2}	...	α_{2n}
...	...	...	...	...	...
x_r	β_r	α_{rr+1}	α_{rr+1}	...	α_{rn}
$Z_{\min}$	γ_0	γ_{r+1}	γ_{r+2}	...	γ_n

2. Вычисляют отношение свободных членов к положительным элементам разрешающего столбца (симплекс-отношение). Находят наименьшее из этих симплекс-отношений, оно соответствует разрешающей строке.

3. На пересечении разрешающей строки и разрешающего столбца находится разрешающий элемент.

4. Если имеется несколько одинаковых по величине симплекс-отношений, то выбирают любое из них. То же самое относится к положительным элементам последней строки симлекс-таблицы.

5. После нахождения разрешающего элемента переходят к следующей таблице. Неизвестные переменные, соответствующие разрешающей строке и столбцу, меняют местами. При этом базисная переменная становится свободной переменной, и наоборот. Симплекс-таблица преобразована следующим образом (табл. 7.6).

6. Элемент табл. 7.6, соответствующий разрешающему элементу табл. 7.5, равен обратной величине разрешающего элемента.

7. Элементы строки табл. 7.6, соответствующие элементам разрешающей строки табл. 7.5, получаются путем деления соответствующих элементов табл. 7.5 на разрешающий элемент.

8. Элементы столбца табл. 7.6, соответствующие элементам разрешающего столбца табл. 7.5, получаются путем деления соответствующих элементов табл. 7.5 на разрешающий элемент и берутся с противоположным знаком.

9. Остальные элементы вычисляются по правилу прямоугольника: мысленно вычерчиваем прямоугольник в табл. 7.6, одна вер-

Таблица 7.6

Преобразование симплекс-таблицы

Базисные неизвестные \ Свободные неизвестные	Свободный член	x_{r+1}	x_1	...	x_n
x_{r+2}	$\frac{\beta_1}{\alpha_{1r+2}}$	$\frac{\alpha_{1r+1}}{\alpha_{1r+2}}$	$\frac{1}{\alpha_{1r+2}}$	...	$\frac{\alpha_{1n}}{\alpha_{1r+2}}$
x_2			$-\frac{\alpha_{2r+2}}{\alpha_{1r+2}}$	...	
...	...	...	...	...	...
x_r			$-\frac{\alpha_{rr+2}}{\alpha_{1r+2}}$	...	
$Z_{\min}$			$-\frac{\gamma_{r+2}}{\alpha_{1r+2}}$	...	

шина которого совпадает с разрешающим элементом, а другая – с элементом, образ которого мы ищем; остальные две вершины определяются однозначно. Тогда искомый элемент из табл. 7.6 будет равен соответствующему элементу табл. 7.5 минус дробь, в знаменателе которой стоит разрешающий элемент, а в числителе – произведение элементов из двух неиспользованных вершин прямоугольника.

10. Как только получится таблица, в которой в последней строке все элементы отрицательны, считается, что минимум найден. Минимальное значение функции равно свободному члену в строке целевой функции, а оптимальное решение определяется свободными членами при базисных переменных. Все свободные переменные в этом случае равны нулю.

11. Если в разрешающем столбце все элементы отрицательны, то задача не имеет решений (минимум не достигается).

Пример 7.12. Решение задачи симплекс-методом:

$$\begin{cases} -x_1 + x_2 + x_3 = 1; \\ x_1 - x_2 + x_4 = 1; \\ x_1 + x_2 + x_5 = 2. \end{cases}$$

$$Z_{max} = 2x_1 - x_2 + 3x_3 - 2x_4 + x_5.$$

Приведем задачу к виду, допускающему применение симплекс-алгоритма:

$$x_3 = 1 - (-x_1 + x_2);$$
$$x_4 = 1 - (x_1 - x_2);$$
$$x_5 = 2 - (x_1 + x_2).$$

Подставим в выражение Z_{max} величины x_3, x_4, x_5:

$$Z_{max} = 6x_1 - 7x_2 + 3.$$

По алгоритму целевая функция должна стремиться к минимуму:

$$Z_{min} = -Z_{max} = -6x_1 + 7x_2 - 3 = -3 - (6x_1 - 7x_2).$$

Составим симплекс-таблицу (табл. 7.7).

Таблица 7.7

Базисные неизвестные \ Свободные неизвестные	Свободный член	x_1	x_2	
x_3	1	−1	1	
x_4	1	$\boxed{1}$	−1	←
x_5	2	1	1	
Z_{min}	−3	6	−7	
		↑		

Разыскиваем в последней строке наименьший положительный элемент, в нашем примере он равен + 6, первый столбец коэффициентов будет разрешающим. Определим отношение свободных членов к положительным элементам разрешающего столбца. Минимальное симплекс-отношение равно 1. Разрешающий элемент находится на пересечении строки переменной x_4 и столбца – x_1.

Переходим к следующей таблице, используя правило прямоугольника (табл. 7.8):

Таблица 7.8

Свободные неизвестные / Базисные неизвестные	Свободный член	x_1	x_2
x_3	2	1	0
x_4	1	1	–1
x_5	1	–1	2
$Z_{\min}$	–9	–6	–1

В последней строке нет положительных элементов, следовательно, оптимальное решение найдено:

$$Z_{\min} = -9;\ X = (0;\ 0;\ 2;\ 1;\ 1);\ Z_{\max} = -Z_{\min} = 9.$$

7.6. Методы нахождения опорного решения задачи линейного программирования

Метод искусственного базиса

Сформированный выше алгоритм симплекс-метода можно применять лишь в том случае, если выделено первое допустимое решение, т. е. исходная задача линейного программирования приведена к виду

$$\begin{cases} x_1 = \beta_1 - (\alpha_{1\,r+1}x_{r+1} + \alpha_{1\,r+2}x_{r+2} + \ldots + \alpha_{1n}x_n); \\ \vdots \\ x_r = \beta_r - (\alpha_{r\,r+1}x_{r+1} + \alpha_{r\,r+2}x_{r+2} + \ldots + \alpha_{rn}x_n); \end{cases}$$

$$Z_{\min} = \gamma_0 - (\gamma_{r+1}x_{r+1} + \gamma_{r+2}x_{r+2} + \ldots + \gamma_n x_n).$$

При этом $\beta_1 \geq 0, \ldots, \beta_r \geq 0$, тогда, положив свободные неизвестные x_{r+1}, x_{r+2}, x_n равными нулю, получаем опорное решение $x_1 = \beta_1$; $x_2 = \beta_2$; ...; $x_r = \beta_r$.

Рассмотрим метод нахождения опорного решения, основанный на введении искусственных переменных. Для этого запишем задачу линейного программирования в общем виде. Будем рассматривать задачу с числом неизвестных n и r ограничениями:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ \vdots \\ a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n = b_r. \end{cases} \tag{7.42}$$

Перепишем систему (7.42) в другом виде. Для этого введем искусственные переменные y_1, y_2, ..., y_r так, чтобы был выделен базис. Тогда система примет вид

$$\begin{cases} y_1 = b_1 - (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n); \\ \vdots \\ y_r = b_r - (a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n). \end{cases} \tag{7.43}$$

Системы (7.42) и (7.43) будут эквивалентны в том случае, если все y_i, для $i = \overline{1,r}$ будут равны 0. Кроме того, мы считаем, что все $b_i \geq 0$ для $i = \overline{1,r}$. В противном случае соответствующие ограничения из системы (7.42) умножим на – 1. Для того чтобы y_i были равны 0, мы должны преобразовать задачу таким образом, чтобы все искусственные переменные y_i перешли в свободные неизвестные.

В этом случае система (7.43) после преобразования примет вид:

$$\begin{cases} x_1 = \beta_1 - (\alpha_{1\,r+1}x_{r+1} + \alpha_{1\,r+2}x_{r+2} + \ldots + \alpha_{1n}x_n + c_{11}y_1 + \ldots + c_{1r}y_r); \\ \vdots \\ x_r = \beta_r - (\alpha_{rr+1}x_{r+1} + \alpha_{rr+2}x_{r+2} + \ldots + \alpha_{rn}x_n + c_{r1}y_1 + \ldots + c_{rr}y_r). \end{cases} \tag{7.44}$$

От системы (7.43) к системе (7.44) всегда можно перейти шагами симплекс-метода. При таком переходе в качестве линейной формы рассматривают функцию

$$F_{\min} = y_1 + y_2 + \text{Е} + y_r, \tag{7.45}$$

равную сумме искусственных переменных. Переход заканчивают, когда $F_{\min} = 0$ и все искусственные переменные y_i переведены в свободные неизвестные.

Анализ вариантов решений

1. Если $F_{\min} \neq 0$, а все y_i переведены в свободные переменные, то задача не имеет положительного решения.

2. Если $F_{\min} = 0$, а часть y_i осталась в базисе, то для перевода их в свободные переменные необходимо применять специальные приемы.

В симплекс-таблице, соответствующей системе (7.44), после того как $F_{min} = 0$, а все y_i – свободные, вычеркивают строку для F_{min} и столбцы для y_i и решают задачу для исходной линейной формы Z_{min}.

Рекомендуется вводить минимум искусственных переменных.

Пример 7.13. Решение задачи линейного программирования симплекс-методом. *Для нахождения опорного плана использовать метод искусственных переменных.*

Ограничения:

$$\begin{cases} 3x_1 - 5x_2 + x_3 + 2x_4 = 1; \\ 2x_1 - 2x_2 + x_4 - x_5 = -4; \\ x_1 - 3x_2 + 2x_4 - x_5 = -5. \end{cases}$$

Целевая функция:

$$Z_{min} = x_1 + 2x_2,\ x_i \geq 0,\ i = \overline{1, 5}.$$

В базис можно выделить переменную x_3. Введем две искусственные переменные – y_1 и y_2.

$$\begin{aligned} x_3 &= 1 - (3x_1 - 5x_2 + 2x_4); \\ y_1 &= 4 - (-2x_1 + 2x_2 - x_4 + x_5); \\ y_2 &= 5 - (-x_1 + 3x_2 - 2x_4 + x_5); \\ Z_{min} &= 0 - (-x_1 - 2x_2); \\ F_{min} &= y_1 + y_2 = 9 - (-3x_1 + 5x_2 - 3x_4 + 2x_5). \end{aligned}$$

Составим симплекс-таблицу (табл. 7.9):

Таблица 7.9

Свободные неизвестные / Базисные неизвестные	Свободный член	x_1	x_2	x_4	x_5	
x_3	1	3	–5	2	0	
y_1	4	–2	2	–1	$\boxed{1}$	←
y_2	5	–1	3	–2	1	
Z_{min}	0	–1	–2	0	0	
F_{min}	9	–3	5	–3	2	
					↑	

Наименьший положительный элемент в строке линейной формы $F_{\min} = 2$. Разрешающий элемент находится на пересечении столбца переменной x_5 и строки переменной y_1.

Заполним следующую симплекс-таблицу (табл. 7.10):

Таблица 7.10

Свободные неизвестные / Базисные неизвестные	Свободный член	x_1	x_2	x_4	y_1
x_3	1	3	−5	2	0
x_5	4	−2	2	−1	1
y_2	1	1	$\boxed{1}$	−1	−1
$Z_{\min}$	0	−1	−2	0	0
$F_{\min}$	1	1	1	−1	−2

Наименьший положительный элемент в строке линейной формы $F_{\min} = 1$. Минимальное симплекс-отношение соответствует строке переменной y_2.

Заполним новую симплекс-таблицу (табл. 7.11):

Таблица 7.11

Свободные неизвестные / Базисные неизвестные	Свободный член	x_1	y_2	x_4	y_1
x_3	6	8	5	−3	5
x_5	2	−4	−2	1	3
x_2	1	1	1	−1	−1
$Z_{\min}$	2	1	2	−2	−1
$F_{\min}$	0	0	−1	0	−1

Так как $F_{\min} = 0$, а y_1 и y_2 переведены в число свободных, переход к первому опорному решению завершен. Строку, соответствующую $F_{\min}$, и столбцы переменных y_1 и y_2 вычеркиваем в последней таблице и переписываем ее в новом виде (табл. 7.12):

Таблица 7.12

Базисные неизвестные \ Свободные неизвестные	Свободный член	x_1	x_4
x_3	6	$\boxed{8}$	–3
x_5	2	–4	1
x_2	1	1	–1
Z_{min}	2	1	–2

Решим задачу для исходной линейной формы Z_{min}. В табл. 7.12 находим разрешающий элемент. Он равен 8. Выполняя действия согласно алгоритму симплекс-метода, получим табл. 7.13:

Таблица 7.13

Базисные неизвестные \ Свободные неизвестные	Свободный член	x_3	x_4
x_1	6/8	1/8	–3/8
x_5	5	4/8	–12/8
x_2	2/8	–1/8	–5/8
Z_{min}	10/8	–1/8	–13/8

В последней строке (Z_{min}) положительных элементов нет, следовательно, оптимальное решение найдено.

Значение целевой функции Z_{min} равно 10/8. Оптимальный план $\overline{X} = (6/8;\ 2/8;\ 0;\ 0;\ 5)$.

Второй алгоритм отыскания опорного плана

Пусть задача линейного программирования записана в каноническом виде:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m; \end{cases} \tag{7.46}$$

$$Z_{\min} = c_1x_1 + c_2x_2 + ... + c_nx_n; \quad (7.47)$$

$$b_i \geq 0,\ i = \overline{1,m},\ \ x_j \geq 0, j = \overline{1,n}.$$

Построим первую таблицу Жордана–Гаусса для задач (7.46) и (7.47). Для единообразия вычислительной процедуры к исходной таблице приписываем строку целевой функции:

$$\left|\begin{array}{cccc|c|c} a_{11} & a_{12} & .... & a_{1n} & b_1 & \Sigma_1 \\ .... & .... & .... & .... & .... & .... \\ a_{m1} & a_{m2} & .... & a_{mn} & b_m & \Sigma_m \\ c_1 & c_2 & .... & c_n & 0 & \Sigma_n. \end{array}\right. \quad (7.48)$$

После приведения системы ограничений к единичному базису целевая функция, как и базисные переменные, будет выражена через свободные переменные. Аналогичным приемом мы пользовались, когда решали задачи графическим методом с числом переменных более двух.

Алгоритм метода

1. Запишем задачу в форме (7.48), при этом все элементы столбца свободных членов b_i должны быть неотрицательны $b_i \geq 0$, $i = \overline{1,m}$. Уравнения системы (7.46), в которых свободные члены отрицательны, предварительно нужно умножить на -1.

2. Таблицу (7.48) преобразуем шагами Жордана–Гаусса исключений. При этом на каждом шаге разрешающим может быть выбран любой столбец, содержащий хотя бы один положительный элемент. Строка целевой функции на выбор разрешающих столбцов влияние не оказывает.

3. Разрешающая строка определяется по наименьшему из отношений свободных членов к положительным элементам разрешающего столбца.

4. В процессе преобразований вычеркиваем строки, состоящие из одних нулей.

5. Если в процессе преобразований встречается строка, все элементы которой нули, а свободный член отличен от нуля, то задача не имеет решения. Если встретится строка, в которой, кроме свободного члена, других положительных элементов нет, то говорят, что задача не имеет положительных решений.

Пояснение. В п.1 алгоритма предполагается, что все элементы столбца свободных членов неотрицательны. Это требование необязательно. В случае когда в столбце свободных членов встречаются отрицательные числа, будем пользоваться теоремой.

Теорема. *Если разрешающий элемент выбирать по наименьшему положительному симплекс-отношению, то после шага Жордана–Гаусса свободный член в разрешающей строке становится положительным, а остальные члены сохраняют свой знак.*

Выбор разрешающего элемента производят иначе, а именно.

1. Просматривают строку, соответствующую какому-либо отрицательному свободному члену. Выбирают в ней какой-либо отрицательный элемент – соответствующий этому элементу столбец будет разрешающим.

2. Выбор разрешающего элемента производится по минимальному положительному симплекс-отношению. Если задача разрешима, то через конечное число шагов получают первое допустимое решение и можно применять симплекс-метод.

В некоторых случаях найденное таким образом первое допустимое решение является также и оптимальным решением.

Пример 7.14. Определение опорного плана.

Дана следующая система:

$$\begin{cases} x_1 + x_2 - x_3 - x_4 = -4; \\ x_1 - 2x_2 - x_3 - x_5 = -7; \\ 2x_1 - x_2 + x_4 + x_5 = 7; \end{cases} \tag{7.49}$$

$$Z_{\max} = 3x_1 - x_2 + 5; \tag{7.50}$$

$$x_i \geq 0,\ i = \overline{1,\ 5}.$$

Два свободных члена в системе ограничений отрицательны, поэтому перед тем, как записать задачу в виде (7.48), умножим соответствующие уравнения (7.49) на –1:

$$\begin{cases} -x_1 - x_2 + x_3 + x_4 = 4; \\ -x_1 + 2x_2 + x_3 + x_5 = 7; \\ 2x_1 - x_2 + x_4 + x_5 = 7; \end{cases}$$

$$Z_{\max} = 3x_1 - x_2 + 5.$$

Запишем данную задачу в виде выражения 7.48 и получим табл. 7.14.

Выберем произвольно столбец с положительным элементом. Затем по минимальному положительному отношению находим разрешающий элемент. В нашем примере он равен 1 и находится на пересечении столбца переменной x_4 и первой строки (элемент отмечен квадратом). Выполняя преобразования Жордана–Гаусса, получим табл. 7.15.

Таблица 7.14

x_1	x_2	x_3	x_4	x_5	Свободный член	Σ
−1	−1	1	$\boxed{1}$	0	4	4
−1	2	1	0	1	7	10
2	−1	0	1	1	7	10
3	−1	0	0	0	5	7

Таблица 7.15

x_1	x_2	x_3	x_4	x_5	Свободный член	Σ
−1	−1	1	1	0	4	4
−1	2	1	0	1	7	10
3	0	−1	0	$\boxed{1}$	3	6
3	−1	0	0	0	5	7

Выполняя в дальнейшем все действия алгоритма, получим ряд следующих аналогичных таблиц (табл. 7.16, 7.17):

Таблица 7.16

x_1	x_2	x_3	x_4	x_5	Свободный член	Σ
−1	−1	1	1	0	4	4
−4	2	$\boxed{2}$	0	0	4	4
3	0	−1	0	1	3	6
3	−1	0	0	0	5	7

Таблица 7.17

x_1	x_2	x_3	x_4	x_5	Свободный член	Σ
1	−2	0	1	0	2	2
−2	1	1	0	0	2	2
1	1	0	0	1	5	8
3	−1	0	0	0	5	7

На данном этапе расчетов в табл. 7.17 мы получили три нулевых столбца, что соответствует количеству ограничений в примере. Здесь необходимо закончить преобразования Жордана–Гаусса.

Запишем нашу задачу из табл. 7.17 так:

$$\begin{cases} x_1 - 2x_2 + x_4 = 2 \\ -2x_1 + x_2 + x_3 = 2 \\ x_1 + x_2 + x_5 = 5 \end{cases} \quad \text{или} \quad \begin{cases} x_4 = 2 - (x_1 - 2x_2); \\ x_3 = 2 - (-2x_1 + x_2); \\ x_5 = 5 - (x_1 + x_2); \end{cases}$$

$$Z_{max} = 3x_1 - x_2 + 5; \qquad Z_{max} = 5 - (-3x_1 + x_2).$$

Получено первое допустимое решение, выделим его. Для этого положим $x_1 = 0$; $x_2 = 0$, тогда $\overline{X} = (0; 0; 2; 2; 5)$; $Z_{max} = 5$.

Задача решена.

7.7. Экономическая интерпретация решения задачи линейного программирования

Любая экономико-математическая модель лишь упрощенно, грубо отображает реальный экономический процесс, и это упрощение существенно сказывается на получаемых результатах. Исследователя вряд ли устроила бы заключительная симплекс-таблица, из которой можно было бы получить только список переменных и их значения. На самом же деле результирующая симплекс-таблица «насыщена» весьма важными данными, лишь небольшую часть которых составляют оптимальные значения переменных. Из симплекс-таблицы можно получить информацию относительно:

- оптимального решения;
- статуса ресурсов;
- ценности каждого ресурса;
- чувствительности оптимального решения к изменению запасов ресурсов, вариациям коэффициентов целевой функции и интенсивности потребления ресурсов.

Сведения, относящиеся к первым трем пунктам, можно извлечь непосредственно из итоговой симплекс-таблицы. Получение информации, относящейся к четвертому пункту, требует дополнительных вычислений.

Для иллюстрации возможностей получения указанной выше информации из заключительной симплекс-таблицы воспользуемся опять задачей об ассортименте продукции (пример 7.2). Эта задача формулируется следующим образом:

максимизировать:	$Z = 3x_1 + 4x_2$	(доход);
при следующих ограничениях:	$2x_1 + 3x_2 \le 9$	(сырье А),
	$3x_1 + 2x_2 \le 13$	(сырье В),
	$x_1 - x_2 \le 1$	(спрос),
	$x_2 \le 2$	(спрос).

Оптимальная симплекс-таблица имеет вид (табл. 7.18):

Таблица 7.18

Базисные неизвестные \ Свободные неизвестные	Свободный член	y_1	y_3
x_1	2,4	0,2	0,6
y_2	3	−1	−1
y_4	0,6	−0,2	0,4
x_2	1,4	0,2	−0,4
Z_{max}	12,8	1,4	0,2

В таблице y_j, $j = \overline{1,4}$ – выравнивающие переменные.

Оптимальное решение. С точки зрения практического использования результатов решения задач линейного программирования классификация переменных на базисные и небазисные не имеет значения и при анализе оптимального решения может не учитываться. Переменные, отсутствующие в симплекс-таблице в столбце «базисные переменные», обязательно имеют нулевое значение. Значения остальных переменных приводятся в столбце «свободные члены».

При интерпретации результатов оптимизации в задаче об ассортименте продукции нас прежде всего интересуют объемы производства продукции $П_1$ и $П_2$, т. е. значения управляемых переменных x_1 и x_2. Используя данные, содержащиеся в симплекс-таблице для оптимального решения, основные результаты можно представить в следующем виде (табл. 7.19):

Таблица 7.19

Управляемые переменные	Оптимальные значения	Решение
x_1 x_2	2,4 1,4	Объем производства продукции $П_1$ должен быть равен 2,4 ед. в сутки Объем производства продукции $П_2$ должен быть равен 1,4 ед. в сутки
Z_{max}	12,8	Доход от реализации продукции будет равен 12,8 д. е. в сутки

Статус ресурсов. В подразд. 7.4 ресурсы относились либо к дефицитным, либо к недефицитным – в зависимости от того, полное или частичное их использование предусматривает оптимальное решение задачи. Сейчас цель состоит в том, чтобы получить соответствующую информацию непосредственно из оптимальной таблицы.

В модели, построенной для задачи об ассортименте продукции, фигурируют четыре ограничения со знаком «≤». Первые два ограничения (определяющие допустимый расход исходного сырья) представляют собой истинные ограничения на ресурсы. Третье и четвертое ограничения относятся к спросу. Эти требования можно рассматривать как ограничения на соответствующие ресурсы, так как увеличение спроса на продукцию эквивалентно расширению представительства предприятия на рынке сбыта. В отношении финансовых средств такая ситуация имеет те же последствия, что и увеличение запасов ресурсов, требующее распределения дополнительных вложений.

Из вышеизложенного следует, что статус ресурсов (дефицитный или недефицитный) для любой модели линейного программирования можно установить непосредственно из результирующей симплекс-таблицы, обращая внимание на значения выравнивающих переменных. Применительно к нашей задаче можно привести следующую сводную таблицу (табл. 7.20).

Положительное значение выравнивающей переменной указывает на неполное использование соответствующего ресурса, т. е. данный ресурс является недефицитным. Если же выравнивающая переменная равна 0, то это свидетельствует о полном потреблении соответствующего ресурса. Из сводной табл. 7.20 видно, что ресурсы 2 и 4 связаны с запасами сырья *В* и возможностями сбыта продукции $П_2$. Поэтому любое увеличение их запасов сверх установ-

Таблица 7.20

Ресурс	Выравнивающая переменная	Статус ресурса
Сырье A	$y_1 = 0$	Дефицитный
Сырье B	$y_2 = 3$	Недефицитный
Превышение объема производства продукции $П_1$ по отношению к объему производства продукции $П_2$	$y_3 = 0$	Дефицитный
Спрос на продукцию $П_2$	$y_4 = 0{,}6$	Недефицитный

ленного максимального значения приведет лишь к тому, что они станут еще более недефицитными. Оптимальное решение задачи при этом останется неизменным.

Ресурсы, увеличение запасов которых позволяет улучшить решение (увеличить доход), – это сырье A и возможности по сбыту продукции $П_1$, поскольку из оптимальной симплекс-таблицы (табл. 7.18) видно, что они дефицитные. В связи с этим логично поставить вопрос: какому из дефицитных ресурсов следует отдать предпочтение при вложении дополнительных средств на увеличение их запасов, с тем чтобы получить от них максимальную отдачу? Ответ на этот вопрос будет дан в следующем разделе этой главы, где рассматривается ценность различных ресурсов.

Ценность ресурса. Ценность ресурса характеризуется величиной улучшения оптимального значения Z, приходящегося на единицу прироста объема данного ресурса. Графическая интерпретация этого определения применительно к условиям задачи об ассортименте продукции была дана в подразд. 7.4 (вторая задача на чувствительность). Графический анализ показывает, что ценность ресурсов 1, 2, 3 и 4 равна:

$U_1 = 1{,}4$ д. е. на единицу прироста запасов ресурса сырья A;
$U_2 = 0$, $U_4 = 0$;
$U_3 = 0{,}2$ д. е. на единицу прироста превышения производства продукции $П_1$ по отношению к объему производства продукции $П_2$.

Эта информация представлена в оптимальной таблице (табл. 7.18). Обратим внимание на значения коэффициентов Z-уравнения, стоящих при переменных начального базиса y_1, y_2, y_3 и y_4. Значения указанных коэффициентов (1,4; 0; 0,2; 0) в точности соответствуют значениям U_1; U_2; U_3; U_4.

Хотя в подразд. 7.4 были даны необходимые разъяснения, связанные с определением ценности ресурсов, покажем, каким образом аналогичный результат можно получить непосредственно из симплекс-таблицы.

Рассмотрим Z-уравнение оптимальной симплекс-таблицы решения задачи об ассортименте продукции:

$$Z = 12{,}8 - (1{,}4 \cdot y_1 + 0 \cdot y_2 + 0{,}2 \cdot y_3 + 0 \cdot y_4).$$

Положительное приращение переменной y_1 относительно ее текущего нулевого значения приводит к пропорциональному уменьшению Z, причем коэффициент пропорциональности равен 1,4 д. е. Однако из первого ограничения модели следует

$$2x_1 + 3x_2 + y_1 = 9,$$

т. е. увеличение y_1 эквивалентно снижению запаса ресурса 1 (сырья A). Отсюда следует, что уменьшение запаса первого ресурса вызывает пропорциональное уменьшение целевой функции Z с коэффициентом пропорциональности, равным 1,4 д. е. Аналогичные рассуждения справедливы и для ресурса 3.

В отношении ресурсов 2 и 4 было установлено, что их ценность равна 0 ($U_2 = U_4 = 0$). Этого и следовало ожидать, так как ресурсы 2 и 4 оказались недефицитными. Такой результат получается всякий раз, когда соответствующие выравнивающие переменные имеют положительное значение.

Несмотря на то что ценность различных ресурсов, определяемая значениями переменных U_i, была представлена в стоимостном (д. е.) выражении, ее нельзя отождествлять с действительными ценами, по которым возможна закупка соответствующих ресурсов. На самом деле речь идет о некоторой мере, имеющей экономическую природу и количественно характеризующей ценность ресурса только относительно полученного оптимального значения Z. При изменении ограничений модели соответствующие экономические оценки будут меняться даже тогда, когда оптимизируемый процесс предполагает применение тех же ресурсов. Поэтому при характеристике ценности ресурсов экономисты предпочитают использовать такие термины, как теневая цена или двойственная оценка. Заметим, что теневая цена характеризует интенсивность улучшения оптимального решения Z. Однако при этом не фиксируется интервал значений увеличения запасов ресурсов, при которых интенсивность улучшения целевой функции остается постоянной. Для большинства практических ситуаций логично предположить наличие верхнего предела увеличения запасов, при превышении которого соответствующее ограничение становится избыточным, что в свою очередь приводит к новому базисному решению и соответствующим ему новым теневым ценам. Ниже определяется интервал значений запасов ресурса, при которых соответствующее ограничение не становится избыточным.

Максимальное изменение запаса ресурса. При решении вопроса о том, запас какого из ресурсов следует увеличивать в первую очередь, обычно используются двойственные оценки (теневые цены). Чтобы определить интервал значений изменения запаса ресурса, при которых двойственная оценка данного ресурса, фигурирующая в заключительной симплекс-таблице, остается неизменной, необходимо выполнить ряд дополнительных вычислений. Положим, что в задаче об ассортименте продукции запас первого ресурса (сырья *А*) изменился на Δ_1, т. е. запас сырья *А* составит $(9 + \Delta_1)$ единиц. Введем это изменение в начальную симплекс-таблицу и затем выполним всю последовательность вычислений.

Поскольку элементы правых частей ограничений никогда не используются в качестве разрешающих, то очевидно, что на каждой итерации вычислений Δ_1 будет оказывать влияние только на значения элементов столбца «свободные члены».

Результаты вычислений элементов столбца «свободные члены» сведены в табл. 7.21:

Таблица 7.21

Уравнение	Значения элементов столбца «свободные члены»	
	Начальная симплекс-таблица	Оптимальная симплекс-таблица
Z	0	$12{,}8 + 1{,}4 \cdot \Delta_1$
1	$9 + \Delta_1$	$2{,}4 + 0{,}2 \cdot \Delta_1$
2	13	$3 - 1 \cdot \Delta_1$
3	1	$0{,}6 - 0{,}2 \cdot \Delta_1$
4	2	$1{,}4 + 0{,}2 \cdot \Delta_1$

Все изменения элементов столбца «свободные члены» определяются непосредственно по данным, содержащимся в симплекс-таблицах. Каждый элемент столбца «свободные члены» представляет собой сумму двух величин:

1) постоянной;

2) члена, линейно зависящего от Δ_1.

Постоянные соответствуют числам, которые фигурируют в оптимальной симплекс-таблице до введения Δ_1 в столбце «свободные члены». Коэффициенты при Δ_1 во вторых слагаемых равны коэффициентам при y_1 в оптимальной симплекс-таблице.

Заметим, что при анализе изменений в правых частях второго, третьего и четвертого ограничений нужно пользоваться коэффициентами при переменных y_2, y_3, y_4 соответственно.

Так как введение Δ_1 сказывается лишь на правой части ограничений (на элементах столбца «свободные члены»), изменение запаса ресурса может повлиять только на допустимость решения. Поэтому Δ_1 не может принимать значений, при которых какая-либо из базисных переменных становится отрицательной. Из этого следует, что величина Δ_1 должна быть ограничена таким интервалом значений, при котором выполняется условие неотрицательности правых частей ограничений в результирующей симплекс-таблице, т. е.:

$$x_1 = 2{,}4 + 0{,}2 \cdot \Delta_1 \geq 0; \tag{7.51}$$

$$y_2 = 3 - 1 \cdot \Delta_1 \geq 0; \tag{7.52}$$

$$y_4 = 0{,}6 - 0{,}2 \cdot \Delta_1 \geq 0; \tag{7.53}$$

$$x_2 = 1{,}4 + 0{,}2 \cdot \Delta_1 \geq 0. \tag{7.54}$$

Для определения допустимого интервала изменения Δ_1 рассмотрим два случая.

Случай 1: $\Delta 1 > 0$.

Соотношения (7.51) и (7.54) всегда выполняются при $\Delta_1 \geq 0$. Соотношения (7.52) и (7.53) определяют следующие предельные значения Δ_1: $\Delta_1 \leq 3$; $\Delta_1 \leq 3$. Таким образом, все четыре соотношения выполняются при $\Delta_1 \leq 3$.

Случай 2: $\Delta_1 < 0$.

Соотношения (7.52) и (7.53) выполняются при $\Delta_1 < 0$. Соотношения (7.51) и (7.54) справедливы при $\Delta_1 \geq -12$; $\Delta_1 \geq -7$ соответственно.

Таким образом, оба соотношения справедливы при $\Delta_1 \geq -7$.

Объединяя результаты, полученные для обоих случаев, можно сделать вывод, что при $-7 \leq \Delta_1 \leq 3$ решение рассматриваемой системы всегда будет допустимым. Любое значение Δ_1, выходящее за предел указанного интервала (т.е. уменьшение запаса сырья *А* более чем на 7 единиц или увеличение более чем на 3 единицы), приведет к недопустимости решения и новой совокупности базисных переменных.

Анализ на чувствительность оптимального решения к вариации коэффициентов целевой функции. В подразд. 7.4 на основе графического представления модели было показано, что при определенных значениях изменения коэффициентов целевой функции оптимальные значения переменных остаются неизменными (хотя оптимальное значение *Z* при этом меняется). Возвращаясь к этому вопросу, покажем, каким образом интересующую нас информацию можно получить из данных, содержащихся в оптимальной симплекс-таблице.

Следует отметить, что уравнение целевой функции также не используется в качестве ведущего уравнения. Поэтому любые изменения коэффициентов целевой функции окажут влияние только на Z-уравнение результирующей симплекс-таблицы. Это означает, что такие изменения могут сделать полученное решение неоптимальным. Наша цель заключается в том, чтобы найти интервалы изменений коэффициентов целевой функции, при которых оптимальные значения переменных остаются неизменными.

Чтобы показать, как выполняются соответствующие вычисления, положим, что доход, получаемый с единицы продукции $П_1$, изменяется от 3 до $3 + \delta_1$, где δ_1 может быть как положительным, так и отрицательным числом. Целевая функция в этом случае принимает следующий вид:

$$Z_{max} = (3 + \delta_1)x_1 + 4x_2.$$

Если воспользоваться данными начальной симплекс-таблицы и выполнить все вычисления, необходимые для получения оптимальной симплекс-таблицы, то последнее Z_{max}-уравнение будет выглядеть следующим образом:

Свободные переменные	Свободные члены	y_1	y_2
Z_{max}	$12,8 + 2,4\ \delta_1$	$1,4 + 0,2\ \delta_1$	$0,2 + 0,6\ \delta_1$

Это уравнение (строка целевой функции) отличается от Z-уравнения до введения δ_1 только наличием членов, содержащих δ_1. Коэффициенты при δ_1 равны коэффициентам при соответствующих переменных в x_1-уравнении (x_1-строка) симплекс-таблицы для полученного ранее оптимального решения:

Базисные переменные \ Свободные переменные	Свободные члены	y_1	y_3
x_1	2,4	0,2	0,6

Мы рассматриваем x_1-уравнение, так как коэффициент именно при этой переменной в выражении для целевой функции в начальной симплекс-таблице изменился на δ_1.

Оптимальные значения переменных будут оставаться неизменными при значениях δ_1, удовлетворяющих условию неотрицательности (задача на отыскание максимума) всех коэффициентов при свободных переменных в Z-уравнении. Таким образом, должны выполняться следующие неравенства:

$$1{,}4 + 0{,}2\,\delta_1 \geq 0;$$
$$0{,}2 + 0{,}6\,\delta_1 \geq 0.$$

Из первого неравенства получаем, что $\delta_1 \geq -7$, а из второго следует, что $\delta_1 \geq -\frac{1}{3}$. Эти результаты определяют пределы изменения коэффициента $-\frac{1}{3} \leq \delta_1 < +\infty$.

Таким образом, при уменьшении коэффициента целевой функции при переменной x_1 до значения, равного $\left(3 + \left(-\frac{1}{3}\right)\right) = 2\frac{2}{3}$, или при его увеличении до $+\infty$ оптимальные значения переменных остаются неизменными. Этот вывод совпадает с результатом, полученным в подразд. 7.4.

Следует отметить, что оптимальное значение Z будет изменяться в соответствии с выражением $(12{,}8 + 2{,}4\,\delta_1)$, где $-\frac{1}{3} \leq \delta_1 < +\infty$.

Мы рассмотрели случай изменения коэффициента при базисной переменной x_1. В случае изменения коэффициента при свободной переменной в целевой функции происходит изменение коэффициента только при данной переменной в оптимальной симплекс-таблице. Рассмотрим в качестве иллюстрации случай, когда коэффициент при свободной переменной y_1 (первая выравнивающая переменная) изменяется от 0 до δ_2. Выполнение преобразований, необходимых для получения заключительной симплекс-таблицы, приводит к следующему результирующему Z-уравнению:

Свободные переменные	Свободные члены	y_1	y_2
Z_{max}	12,8	$1{,}4 - \delta_2$	0,2

Из приведенного фрагмента заключительной симплекс-таблицы видно, что единственное отличие от Z-уравнения до введения

δ_2 состоит в том, что коэффициент при y_3 уменьшился на δ_2. Таким образом, коэффициент при свободной переменной в результирующем Z-уравнении нужно уменьшить на ту же величину, на которую он увеличивался в исходном Z-уравнении.

7.8. Двойственные задачи линейного программирования

Взаимодвойственные задачи. Рассмотрим задачу об использовании ресурсов. Пусть предприятие № 1 производит n видов продукции и использует m видов сырья. Известна прибыль, получаемая с единицы продукции $c_j, j = \overline{1, n}$. Известны технологические коэффициенты a_{ij}, $i = \overline{1, m}$, $j = \overline{1, n}$. Требуется организовать производство так, чтобы предприятию была обеспечена максимальная прибыль. Сведем все исходные данные в табл. 7.22:

Таблица 7.22

Цены на ресурсы	Запасы сырья	Продукция			
		$П_1$	$П_2$	...	$П_n$
u_1	a_1	a_{11}	a_{12}	...	a_{1n}
u_2	a_2	a_{21}	a_{22}	...	a_{2n}
...	...	...	...	...	...
...	...	...	...	...	...
u_m	a_m	a_{m1}	a_{m2}	...	a_{mn}
	Прибыль с единицы продукции	c_1	c_2	...	c_n

Запишем в общем виде экономико-математическую модель задачи об использовании ресурсов. Для этого введем переменные x_j, $j = \overline{1, n}$ – количество продукции j-го вида. Тогда ограничения на сырье запишутся в виде

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le a_1; \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le a_2; \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le a_m. \end{aligned} \tag{7.55}$$

Целевая функция, определяющая максимум прибыли, имеет вид

$$Z_{\max} = c_1x_1 + c_2x_2 + \ldots + c_nx_n; \tag{7.56}$$

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

По этим же исходным данным сформулируем задачу по предприятию № 2.

Допустим, предприятие № 2 решило закупить все ресурсы, которыми располагает предприятие № 1. В этом случае предприятию № 2 необходимо установить оптимальные цены на эти ресурсы, исходя из *следующих условий*:

1) общая стоимость ресурсов для предприятия № 2 должна быть минимальной;

2) за каждый вид ресурса предприятию № 1 надо уплатить не менее той суммы, которую это предприятие может получать при переработке данного вида ресурса в готовую продукцию.

Обозначим цены, по которым предприятие № 2 покупает ресурсы у предприятия № 1, через u_i, $i = \overline{1, m}$. Запишем экономико-математическую модель для предприятия № 2 с учетом вышеуказанных условий 1) и 2).

Целевая функция, определяющая минимальную суммарную стоимость ресурсов, имеет вид

$$L_{\min} = a_1 u_1 + a_2 u_2 + \dots + a_m u_m. \tag{7.57}$$

В соответствии с условием 2) запишем систему ограничений:

$$\begin{aligned} &a_{11}u_1 + a_{21}u_2 + \dots + a_{m1}u_m \geq c_1; \\ &a_{12}u_1 + a_{22}u_2 + \dots + a_{m2}u_m \geq c_2; \\ &\vdots \\ &a_{1n}u_1 + a_{2n}u_2 + \dots + a_{mn}u_m \geq c_n. \end{aligned} \tag{7.58}$$

Сравним математические модели задач (7.55), (7.56) и (7.57), (7.58):

1) число неизвестных одной задачи равно числу ограничений другой задачи;

2) матрица коэффициентов системы ограничений получается одна из другой путем транспонирования;

3) неравенства в системах ограничений имеют противоположный смысл;

4) свободные члены системы ограничений одной из задач становятся коэффициентами целевой функции другой задачи, коэффициенты целевой функции превращаются в свободные члены ограничений;

5) целевые функции в задачах имеют противоположный смысл: в первой – max, во второй – min.

Задачи линейного программирования, обладающие пятью указанными формальными признаками, называются ***симметричными***. Одна из них называется основной, а другая – двойственной.

В линейном программировании кроме симметричных двойственных пар существуют несимметричные двойственные пары, которые имеют следующий вид:

основная задача

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1; \\ &\vdots \\ &a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m; \end{aligned} \tag{7.59}$$

$$Z_{\max} = c_1x_1 + c_2x_2 + ... + c_nx_n; \tag{7.60}$$

$$x_j \geq 0, j = \overline{1, n};$$

двойственная задача

$$L_{\max} = b_1y_1 + b_2y_2 + ... + b_my_m; \tag{7.61}$$

$$\begin{aligned} &a_{11}y_1 + a_{21}y_2 + ... + a_{m1}y_m \leq c_1; \\ &\vdots \\ &a_{1n}y_1 + a_{2n}y_2 + ... + a_{mn}y_m \leq c_n. \end{aligned} \tag{7.62}$$

Эти задачи отличаются от симметричной пары двумя особенностями:

1) ограничения задачи (7.59)–(7.60) выражены уравнениями вместо неравенств;

2) в задаче (7.61)–(7.62) отсутствуют условия неотрицательности переменных y_i, $i = \overline{1, m}$.

Общее правило построения двойственной пары. К пяти признакам, сформулированным ранее, необходимо добавить следующие:

1) в исходной задаче ограничения неравенства следует записывать со знаком $\geq$, если целевая функция стремится к min, и со знаком $\leq$, если целевая функция стремится к max;

2) каждому ограничению неравенства исходной задачи соответствует в двойственной задаче условие неотрицательности переменных $y_i \geq 0$;

3) каждому условию равенства соответствует переменная y_i без ограничения на знак, и наоборот: неотрицательным переменным x_k из основной задачи в двойственной задаче соответствуют ограничения неравенства, а неограниченным по знаку переменным соответствуют равенства.

Пример 7.15.

Дана следующая система:

$$\begin{cases} x_1 + 2x_2 + x_3 \geq 9; \\ 2x_1 + x_3 \geq 4; \end{cases}$$

$$x_1 \geq 0;\ x_2 \geq 0;\ x_3 \geq 0;$$

$$Z_{\min} = x_1 + x_2 + x_3.$$

Проверяем условие: целевая функция стремится к min, а знак неравенства должен быть $\geq$. Исходная задача соответствует данному условию, и можно сразу приступить к построению симметричной двойственной пары.

Так как в прямой задаче в системе ограничений два неравенства, то в двойственной будет две переменных u_1 и u_2, причем $u_1 \geq 0$, $u_2 \geq 0$.

Целевая функция

$$Z_{\max} = 9u_1 + 4u_2.$$

Учитывая, что целевая функция на max и в прямой задаче $x_1 \geq 0$; $x_2 \geq 0$; $x_3 \geq 0$, то система ограничений запишется в следующем виде:

$$\begin{cases} u_1 + 2u_2 \leq 1; \\ 2u_1 \ \leq 1; \\ u_1 + u_2 \leq 1. \end{cases}$$

Пример 7.16.

Имеем систему

$$\begin{cases} x_1 & -2x_2 & +x_3 & +3x_4 & -2x_5 & = & 6, \\ 2x_1 & +3x_2 & -2x_3 & -x_4 & +x_5 & \leq & 4, \\ x_1 & & +3x_3 & & -4x_5 & \geq & 8, \end{cases}$$

$x_1 \geq 0$; $x_3 \geq 0$; $x_5 \geq 0$; x_2 и x_5 не имеют ограничения знака;
$Z_{\min} = x_1 - 2x_2 + x_3 - x_4 + x_5.$

Так как целевая функция на min, то в исходной задаче ограничения неравенства должны иметь знак $\geq$. Умножим второе неравенство системы на -1:

$$-2x_1 - 3x_2 + 2x_3 + x_4 - x_5 \geq -4.$$

В прямой задаче число ограничений равно 3, значит, в двойственной будет три переменных: u_1; u_2; u_3. Так как u_2 и u_3 соответствуют неравенствам, то $u_2 \geq 0$; $u_3 \geq 0$. Параметр u_1 соответствует равенству, поэтому u_1 – переменная без ограничения знака.

Число ограничений в двойственной задаче равно 5, так как в прямой задаче пять переменных, в том числе первое, третье и пятое ограничения будут неравенствами, потому что они соответствуют неотрицательным переменным, а второе и четвертое ограничения будут уравнениями, так как соответствуют переменным без ограничения знака. Запишем двойственную задачу с учетом вышеизложенного:

целевая функция

$$L_{\max} = 6u_1 - 4u_2 + 8u_3;$$

ограничения

$$\begin{aligned}
&u_1 - 2u_2 + u_3 \leq 1;\\
&-2u_1 - 3u_2 = -2;\\
&u_1 + 2u_2 + 3u_3 \leq 1;\\
&3u_1 + u_2 = -1;\\
&-2u_1 - u_2 - 4u_3 \leq 1.
\end{aligned}$$

Решение симметричных двойственных задач. Сформулируем без доказательства теорему, справедливую для любых двойственных задач.

Теорема **(первая теорема двойственности).** *Если одна из двойственных задач имеет оптимальное решение, то и другая его имеет. Причем экстремальные значения целевых функций совпадают. Если же в одной задаче целевая функция не ограничена, то двойственная ей задача противоречива.*

Запишем в общем виде прямую и двойственную задачи линейного программирования:

1) прямая задача:

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \leq a_1;\\
&\vdots\\
&a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \leq a_m;\\
&Z_{\max} = c_1x_1 + c_2x_2 + \ldots + c_nx_n;\\
&x_i \geq 0;\ i = \overline{1, n};
\end{aligned} \tag{7.63}$$

2) двойственная:

$$\begin{aligned} &a_{11}u_1 + a_{21}u_2 + \ldots + a_{m1}u_m \ge c_1; \\ &\vdots \\ &a_{1n}u_1 + a_{2n}u_2 + \ldots + a_{mn}u_m \ge c_n; \\ &L_{\min} = a_1u_1 + a_2u_2 + \ldots + a_mu_m; \\ &u_j \ge 0,\ i = \overline{1,m}. \end{aligned} \tag{7.64}$$

Преобразуем задачи (7.63) и (7.64) к виду, допускающему применение симплекс-алгоритма. Для этого введем выравнивающие переменные y_j, $j = \overline{1,m}$ в прямую задачу и v_i, $i = \overline{1,n}$ – в двойственную задачу:

1) прямая:

$$\begin{aligned} &y_1 = a_1 - (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n); \\ &\vdots \\ &y_m = a_m - (a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n); \\ &Z_{\max} = 0 - (-c_1x_1 - c_2x_2 - \ldots - c_nx_n); \end{aligned} \tag{7.63′}$$

2) двойственная:

$$\begin{aligned} &v_1 = -c_1 - (-a_{11}u_1 - a_{21}u_2 - \ldots - a_{m1}u_m); \\ &\vdots \\ &v_n = -c_n - (-a_{1n}u_1 - a_{2n}u_2 - \ldots - a_{mn}u_m); \\ &L_{\min} = 0 - (-a_1u_1 - a_2u_2 - \ldots - a_mu_m). \end{aligned} \tag{7.64′}$$

Построим для задач (7.63′) и (7.64′) *общую симплекс-таблицу* (табл. 7.23), причем эта таблица имеет дополнительные столбец и строку для двойственной задачи:

Таблица 7.23

Двойственная задача	Б \ С	v_1 / x_1	v_2 / x_2	…	v_n / x_n	$L_{\min}$ / свободные члены
$-u_1$	y_1	a_{11}	a_{12}	…	a_{1n}	a_1
$-u_2$	y_2	a_{21}	a_{22}	…	a_{2n}	a_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$-u_m$	y_m	a_{m1}	a_{m2}	…	a_{mn}	a_m
	$Z_{\max}$	$-c_1$	$-c_2$	…	$-c_n$	0

Задачи, представленные в общей симплекс-таблице (табл. 7.23), решаются обычным симплекс-методом, алгоритм которого приведен выше.

Сформулируем признаки оптимальности двойственной пары задач:

• планы $X = (x_1, x_2, ..., x_n)$ и $U = (u_1, u_2, ..., u_m)$ оптимальны, если $Z_{max}(X) = L_{min}(U)$;

• решения $X = (x_1, x_2, ..., x_n)$ и $U = (u_1, u_2, ..., u_m)$ оптимальны, если все произведения сопряженных условий для этих решений равны 0.

Запишем следующие сопряженные условия:

I группа:

$$\begin{aligned}&(a_1 - a_{11}x_1 - a_{12}x_2 - \ldots - a_{1n}x_n)\, u_1 = 0;\\ &\vdots\\ &(a_m - a_{m1}x_1 - a_{m2}x_2 - \ldots - a_{mn}x_n)\, u_m = 0.\end{aligned}$$

II группа получается, если умножить на $x_1, x_2, ..., x_n$ выражения для базисных переменных двойственной задачи:

$$\begin{aligned}&(a_{11}u_1 + a_{21}u_2 + \ldots + a_{m1}u_m - c_1)\, x_1 = 0;\\ &\vdots\\ &(a_{1n}u_1 + a_{2n}u_2 + \ldots + a_{mn}u_m - c_n)\, x_n = 0.\end{aligned}$$

Пример 7.17. По исходной задаче построить двойственную и найти оптимальное решение обеих задач.

$$Z_{max} = -3x_1 + 5x_2 + x_3 + x_4;$$

$$\begin{aligned}&3x_1 + 8x_2 + x_3 + x_4 \le 50;\\ &5x_1 - 4x_2 - x_3 + x_4 \ge 14;\\ &x_j \ge 0,\ j = \overline{1, 4}.\end{aligned}$$

Согласно общему правилу составления двойственных задач запишем:

$$\begin{aligned}&3u_1 + 5u_2 \ge -3;\\ &8u_1 - 4u_2 \ge 5;\\ &u_1 - u_2 \quad \ge 1;\\ &u_1 + u_2 \quad \ge 1;\end{aligned}$$

$$\begin{aligned}&L_{min} = 50u_1 + 14u_2;\\ &u_1 \text{ и } u_2 \ge 0.\end{aligned}$$

Решим задачи в единой симплекс-таблице. Для этого представим их в виде, позволяющем применить симплекс-метод:

1) *прямая задача*:

вводим выравнивающие переменные x_5, x_6:

$$x_5 = 50 - (3x_1 + 8x_2 + x_3 + x_4);$$
$$x_6 = 14 - (5x_1 - 4x_2 - x_3 + x_4);$$

$$Z_{\max} = 0 - (3x_1 - 5x_2 - x_3 - x_4);$$

2) *двойственная задача*:

вводим в левую часть ограничений выравнивающие переменные v_1, v_2, v_3, v_4 со знаком «–»:

$$v_1 = 3 - (-3u_1 - 5u_2);$$
$$v_2 = -5 - (-8u_1 + 4u_2);$$
$$v_3 = -1 - (-u_1 + u_2);$$
$$v_4 = -1 - (u_1 - u_2);$$

$$L_{\min} = 0 - (-50u_1 - 14u_2).$$

Итак, составим таблицу, внешний контур которой образует двойственную задачу, внутренний – прямую задачу (табл. 7.24):

Таблица 7.24

	Б*	v_1	v_2	v_3	v_4	$L_{\min}$	
С**	С / Б	x_1	x_2	x_3	x_4	свободные члены	
$-u_1$	x_5	3	$\boxed{8}$	1	1	50	←
$-u_2$	x_6	5	–4	–1	1	14	
Свободные члены	$Z_{\max}$	3	–5	–1	–1	0	
			↑				

*Б – базисные переменные.

**С – свободные переменные.

Будем работать по симплекс-методу, но так как записанная в последней строчке функция стремится к максимуму, то в этой строчке мы будем искать наименьший отрицательный элемент.

Построим итоговую таблицу (табл. 7.25):

Таблица 7.25

Б* / С*	Б / С	v_1	u_1	v_3	v_4	L_{min}
С*	Б \ С	x_1	x_5	x_3	x_4	свободные члены
$-v_2$	x_2	3/8	1/8	$\boxed{1/8}$	1/8	50/8 ←
$-u_2$	x_6	52/8	4/8	−4/8	12/8	39
Свободные члены	Z_{max}	39/8	5/8	−3/8 ↑	−3/8	250/8

Б* / С*	Б / С	v_1	u_1	v_2	v_4	L_{min}
С*	Б \ С	x_1	x_5	x_2	x_4	свободные члены
$-v_3$	x_3	3	1	8	1	50
$-u_2$	x_6	8	1	4	$\boxed{2}$	64 ←
Свободные члены	Z_{max}	6	1	3	0 ↑	50

Б* / С*	Б / С	v_1	u_1	v_2	u_2	L_{min}
С*	Б \ С	x_1	x_5	x_2	x_6	свободные члены
$-v_3$	x_3	−1	1/2	6	−1/2	18
$-v_4$	x_4	4	1/2	2	1/2	32
Свободные члены	Z_{max}	6	1	3	0	50

Итак, мы получили итоговую таблицу (табл. 7.25), в последней строчке которой нет отрицательных элементов, следовательно, задачи решены:

прямая задача $X = (0; 0; 18; 32)$, $Z_{max} = 50$;

двойственная задача $U = (1; 0)$, $L_{min} = 50$.

Проверим полученное решение на оптимальность:
условие 1) выполнено: $Z(X) = L(U) = 50$.

Для выполнения условия 2) составим произведения сопряженных условий:

$x_1 v_1 = x_1 (3 - (-3u_1 - 5u_2)) = 0 (3 - (-3 \cdot 1 - 5 \cdot 0)) \equiv 0;$

$x_2 v_2 = x_2 (-5 - (-8u_1 + 4u_2)) = 0 (-5 - (-8 \cdot 1 + 4 \cdot 0)) \equiv 0;$

$x_3 v_3 = x_3 (-1 - (-u_1 + u_2)) = 18 (-1 - (-1 + 0)) \equiv 0;$

$x_4 v_4 = x_4 (-1 - (-u_1 - u_2)) = 18 (-1 - (-1 - 0)) \equiv 0;$

$u_1 x_5 = u_1 (50 - (3x_1 + 8x_2 + x_3 + x_4)) = 1 (50 - (3 \cdot 0 + 8 \cdot 0 + 18 + 32)) \equiv 0;$

$u_2 x_6 = u_2 (14 - (5x_1 - 4x_2 - x_3 + x_4)) = 0 (14 - (5 \cdot 0 - 4 \cdot 0 - 18 + 32)) \equiv 0.$

Итак, оба условия выполняются, значит, полученный план — оптимальный.

Рассмотренную задачу примера 7.17 можно решить иначе. Для этого необходимо:

1) решить прямую задачу обычным симплекс-методом. Решение получим следующее: $Z_{max} = 50$; $X = (0;\ 0;\ 18;\ 32)$;

2) составить произведение сопряженных условий и приравнять их 0:

$$\begin{aligned} x_1 v_1 &= x_1 (3 - (-3u_1 - 5u_2)); \\ x_2 v_2 &= x_2 (-5 - (-8u_1 + 4u_2)); \\ x_3 v_3 &= x_3 (-1 - (-u_1 + u_2)); \\ x_4 v_4 &= x_4 (-1 - (-u_1 - u_2)); \end{aligned}$$

$$\begin{aligned} u_1 x_5 &= u_1 (50 - (3x_1 + 8x_2 + x_3 + x_4)); \\ u_2 x_6 &= u_2 (14 - (5x_1 - 4x_2 - x_3 + x_4)). \end{aligned}$$

Подставим вектор X в записанные условия, тогда будем иметь:

$$\begin{aligned} 18 \cdot (-1 - (-u_1 + u_2)) &= 0; \\ 32 \cdot (-1 - (-u_1 - u_2)) &= 0. \end{aligned}$$

После преобразований получим:

$$\begin{aligned} 1 + (-u_1 + u_2) &= 0; \\ 1 + (-u_1 - u_2) &= 0. \end{aligned}$$

Откуда $u_1 = 1$; $u_2 = 0$, $L_{min} = 50$, так как $Z_{max} = L_{min}$, значит, задача решена верно.

7.9. Экономико-математический анализ полученных оптимальных решений

Оптимальное решение задачи линейного программирования существенно зависит от реальной экономической ситуации, складывающейся на предприятии. На решение задачи могут повлиять следующие экономические ситуации:

1) изменение запасов ресурсов;

2) внедрение нового технологического способа производства, позволяющего снизить расход сырья A и B;

3) произошедшие изменения в ценовой политике на предприятии;

4) предполагается выпуск нового вида продукции.

Результаты влияния данных экономических ситуаций на оптимальное решение можно получить в ходе проведения экономико-математического анализа модели на чувствительность.

Анализ на чувствительность оптимального решения базируется на следующих свойствах двойственных оценок.

1. Двойственные оценки характеризуют дефицитность ресурсов. Величина u_i в оптимальном решении двойственной задачи является оценкой i-го ресурса; чем больше значение оценки u_i, тем выше дефицитность ресурса. Для недефицитного ресурса $u_i = 0$.

2. Двойственные оценки показывают, как влияют изменения правой части ограничений (запасов ресурсов) на значение целевой функции. Практический интерес представляют границы (нижняя и верхняя) изменения ресурсов, в которых величины оценок остаются неизменными.

3. Двойственные оценки являются показателем эффективности производства отдельных видов продукции с позиции критерия оптимальности. С этой точки зрения в оптимальный план может быть включена лишь та продукция j-го вида, для которой выполняется условие

$$\sum_{i=1}^{m} a_{ij} u_i \le c_j,$$

где u_i – оптимальное значение двойственной оценки i-го ресурса;
a_{ij} – технологические коэффициенты;
c_j – доход, получаемый с единицы продукции j-го вида;
m – количество видов ресурсов.

4. Двойственные оценки позволяют провести сравнение суммарных условных затрат и результатов.

Это свойство следует из принципа двойственности, в котором устанавливается связь между значениями функции прямой и двойственной задач, т. е. $Z_{max} = Z_{min}$ (подразд. 7.8). Это означает, что оценка всех затрат производства должна равняться оценке произведенного продукта.

Используя данные свойства двойственных оценок, проведем анализ изменений исходной задачи, которые могут привести к недопустимости и неоптимальности решения.

Обратимся к конкретной задаче и проиллюстрируем применение анализа оптимального решения на чувствительность на примере задачи оптимизации ассортимента выпускаемой продукции (пример 7.2).

Составим математические модели прямой и двойственной задач.

1) *прямая задача*:

максимизировать доход

$$Z_{max} = 3x_1 + 4x_2 \tag{7.65}$$

при ограничениях

$$\begin{aligned} 2x_1 + 3x_2 &\le 9 \quad &&\text{(сырье A)};\\ 3x_1 + 2x_2 &\le 13 \quad &&\text{(сырье B)};\\ x_1 - x_2 &\le 1 \quad &&\text{(спрос)};\\ x_2 &\le 2 \quad &&\text{(спрос)}; \end{aligned} \tag{7.66}$$

$$x_1;\ x_2 \ge 0; \tag{7.67}$$

2) *двойственная задача*:

минимизировать

$$L_{min} = 9u_1 + 13u_2 + u_3 + 2u_4 \tag{7.68}$$

при ограничениях

$$2u_1 + 3u_2 + u_3 \ge 3; \tag{7.69}$$

$$\begin{aligned} &3u_1 + 2u_2 - u_3 + u_4 \ge 4;\\ &u_1;\ u_2;\ u_3;\ u_4 \ge 0. \end{aligned} \tag{7.70}$$

Установив соответствие между переменными обеих задач и решая задачи симплекс-методом, запишем итоговую таблицу с оптимальным решением (табл. 7.26).

В таблице v_1, v_2 – выравнивающие переменные двойственной задачи; x_3, x_4, x_5, x_6 – выравнивающие переменные прямой задачи; u_i – двойственная оценка i-го ресурса ($i = \overline{1, 4}$).

Таблица 7.26

Итоговая таблица

С* \ Б*	Б \ С	u_1 / x_3	u_3 / x_5	L_{min} / свободные члены
$-v_1$	x_1	0,2	0,6	2,4
$-u_2$	x_4	–1	–1	3
$-u_4$	x_6	–0,2	0,4	0,6
$-v_2$	x_2	0,2	–0,4	1,4
Свободные члены	Z_{max}	1,4	0,2	12,8

Изменение запасов ресурсов. Значение двойственной оценки того или иного ресурса показывает, насколько возросло бы значение целевой функции, если бы объем данного ресурса (запас) увеличился на 1 ед. На основе вышеизложенных свойств двойственных оценок можно записать следующее:

$$\Delta Z = u_i \,.\, \Delta b_i, \tag{7.71}$$

где u_i – двойственная оценка i-го ресурса;
Δb_i – приращение i-го ресурса;
ΔZ – изменение целевой функции.

В нашем примере увеличение сырья А на 1 ед. привело бы к росту Z_{max} на 1,4 ед. ($\Delta Z = u_1 \,.\, \Delta b_1 = 1{,}4 \cdot 1 = 1{,}4$).

Двойственная оценка для недефицитного ресурса равна нулю, так как ресурс используется не полностью и увеличение его запасов (Δb_i) не повлияет на оптимальное решение. В нашем примере $u_2 = u_4 = 0$, следовательно, ресурсы 2 и 4 недефицитные. Избыток ресурса 2 (сырья В) составляет 3 ед. (x_4 = 3 ед.), а ресурса 4 – 0,6 ед. (x_6 = 0,6 ед.).

Используя аналитическое выражение (7.71), мы можем выявить только направление деятельности по устранению «узких» мест, обеспечивающее наибольшее изменение целевой функции. Это изменение определяется величиной u_i и может быть установлено лишь тогда, когда при изменении величин b_i значения переменных u_i, соответствующих двойственной задаче, в оптимальном плане остаются неизменными. В связи с этим необходимо определить такие интервалы изменения каждого из свободных членов системы линейных ограничений, в которых оптимальный план двойственной задачи не меняется. Предельные оценки Δb_i нижнего и верхнего прироста запасов i-го ресурса можно определить по следующим формулам:

$$\Delta b_i^{(\text{н})} = \min_k \left\{ \frac{x_k}{d_{ki}} \right\} \text{ для } d_{ki} > 0; \tag{7.72}$$

$$\Delta b_i^{(\text{в})} = \left| \max_k \left\{ \frac{x_k}{d_{ki}} \right\} \right| \text{ для } d_{ki} < 0, \tag{7.73}$$

где x_k – значение k-й базисной переменной из оптимального решения;
d_{ki} – элементы обратной матрицы коэффициентов при базисных переменных.

Элементы обратной матрицы d_{ki} находятся в итоговой симплекс-таблице (табл. 7.26). Определим по формулам (7.72)–(7.73) возможные пределы изменения запасов ресурса 1, при которых двойственные оценки не изменяются.

$$\Delta b_1^{(\text{н})} = \min_k \left\{ \frac{x_k}{d_{k1}} \right\}, \text{ для } d_{ki} > 0;$$

$$\Delta b_1^{\text{н}} = \min \left\{ \frac{x_1}{d_{11}}; \frac{x_2}{d_{21}} \right\} = \min \left\{ \frac{2{,}4}{0{,}2}; \frac{1{,}4}{0{,}2} \right\} = 7;$$

$$\Delta b_1^{\text{в}} = \left| \max \left\{ \frac{x_4}{d_{41}}; \frac{x_6}{d_{61}} \right\} \right| = \left| \max \left\{ \frac{3}{-1}; \frac{0{,}6}{-0{,}2} \right\} \right| = 3.$$

Интервал устойчивости оценок по отношению и изменению ресурса 1 будет равен:

$$[b_1 - b_1^{н};\ b_1 + b_1^{в}] = [9 - 7;\ 9 + 3] = [2;\ 12].$$

Возможные пределы изменения запасов дефицитного ресурса 3, при которых двойственные оценки не изменяются, определяются следующим образом:

$$\Delta b_3^{\text{н}} = \min \left\{ \frac{x_1}{d_{13}}; \frac{x_6}{d_{63}} \right\} = \min \left\{ \frac{2{,}4}{0{,}6}; \frac{0{,}6}{0{,}4} \right\} = 1{,}5;$$

$$\Delta b_3^{\text{в}} = \left|\max\left\{\frac{x_4}{d_{43}}; \frac{x_2}{d_{23}}\right\}\right| = \left|\max\left\{\frac{3}{-1}; \frac{1,4}{-0,4}\right\}\right| = 3,5.$$

Интервал устойчивости оценок по отношению к изменению ресурса 3 (соотношение спроса на продукцию $П_1$ и $П_2$) будет равен:

$$[b_3 - \Delta b_3^{н}; b_3 + \Delta b_3^{в}] = [1 - 1,5; 1 + 3,5] = [-0,5; 4,5].$$

Недефицитный ресурс (2 и 4) используется в производстве не полностью, поэтому верхняя граница интервала устойчивости определяется однозначно исходными данными ($b_2 = 13$; $b_4 = 2$). Нижнюю границу устойчивости можно определить, используя данные табл. 7.24, учитывая, что ненулевые выравнивающие переменные, вошедшие в базис, характеризуют величину избытка недефицитного ресурса. В нашем примере избыток недефицитного ресурса будет равен:

$$\Delta b_2^{в} = x_4 = 3 \text{ и } \Delta b_4^{н} = x_6 = 0,6.$$

Тогда интервалы устойчивости оценок по отношению к изменению ресурсов 2 и 4 вычисляются следующим образом:

для ресурса 2:

$$[b_2 - \Delta b_2^{н}; b_2] = [13 - 3; 13] = [10; 13];$$

для ресурса 4:

$$[b_4 - \Delta b_4^{н}; b_4] = [2 - 0,6; 2] = [1,4; 2].$$

Так как изменения ресурсов находятся в пределах устойчивости оценок, то их раздельное влияние на величину доходов $\Delta Z_{i\,\max}$ определяется произведением оценки u_i и величины Δb_i:

$$\begin{aligned}
\Delta Z_{1\,\max} &= 1,4 \cdot [12 - 2] = 14;\\
\Delta Z_{2\,\max} &= 0 \cdot [13 - 10] = 0;\\
\Delta Z_{3\,\max} &= 0,2 \cdot [4,5 - (-0,5)] = 1;\\
\Delta Z_{4\,\max} &= 0 \cdot [2 - 1,4] = 0.
\end{aligned}$$

Суммарное возможное увеличение оптимального значения функции составит:

$$\Delta Z_{\max} = u_1 \cdot \Delta b_1^{в} + u_3 \cdot \Delta b_3^{в} = 1,4 \cdot 3 + 0,2 \cdot 3,5 \approx 4,9 \text{ д. е.}$$

Здесь можно определить также целесообразность дополнительного приобретения дефицитного ресурса, используя четвертое свойство двойственных оценок. Например, определить, выгодно ли приобретать дополнительно ресурс 1 (сырье А) в размере 2 ед. по цене $c_1 = 0{,}5$ д. е. за 1 ед. ресурса.

Приращение ресурса 1 на величину $\Delta b_1 = 2$ ед. находится в пределах устойчивости двойственных оценок. Следовательно, $\Delta Z_{\max} = \Delta b_1 \,.\, u_1 = 2 \,.\, 1{,}4 = 2{,}8$ ед., в то время как затраты на приобретение 2 ед. ресурса 1 вида составляют:

$$\Delta c = \Delta b_1 \cdot c_1 = 2 \cdot 0{,}5 = 1 \text{ д. е.}$$

Поскольку величина дополнительных доходов ($\Delta Z_{\max}$) больше дополнительных затрат, закупать ресурс 1 в размере 2 ед. по цене $c_1 = 0{,}5$ д. е. целесообразно.

Необходимо подчеркнуть, что изменение правых частей ограничений может повлиять только на элементы правой части симплекс-таблицы и, следовательно, на допустимость самого решения. Поэтому нужно при всяком изменении Δb_i в исходных условиях проводить расчеты новых значений элементов правой части симплекс-таблицы.

Внедрение нового технологического способа производства. Новый технологический способ производства предполагает либо выпуск нового вида продукции, либо изменение технологических коэффициентов, стоящих в левой части ограничений. Для определения эффективности внедряемого нового технологического способа с успехом могут быть использованы двойственные оценки.

Согласно третьему свойству двойственных оценок в оптимальный план включается новая продукция j-го вида, для которой выполняется условие

$$\sum_{i=1}^{m} a_{ij} u_i < c_j. \tag{7.74}$$

Оценим целесообразность введения в оптимальный план задачи (7.65 – 7.67) продукции третьего вида (x_3), для которой технологические коэффициенты $a_{13} = 3$ ед., $a_{23} = 1$ ед., а доход – $c_3 = 8$ д. е.:

$$(a_{13}u_1 + a_{23}u_2) - c_3 = (3 \cdot 1{,}4 + 1 \cdot 0) - 8 = -3{,}8.$$

Так как доходы превышают затраты, то введение в план третьего вида продукции выгодно.

Задачи

Построить математическую модель задачи линейного программирования (7.1 – 7.30).

7.1. Автотранспортному предприятию (АТП) необходимо освободить из-под груза складские помещения клиента. Вывоз груза следует осуществить в два рейса колоннами автомобилей. Условия перевозки требуют, чтобы в составе каждой колонны, предназначенной для вывоза груза в первый район, было 8 автомобилей ЗИЛ-131 и 8 автомобилей ЗИЛ-130; в колоннах второго рейса 8 автомобилей ЗИЛ-130 и 16 – МАЗ-500. Каждая из колонн может сделать за сутки одинаковое количество поездок. Парк подвижного состава АТП состоит из 32 автомобилей ЗИЛ-131 грузоподъемностью 3 т, 48 автомобилей ЗИЛ-130 грузоподъемностью 4 т, 48 автомобилей МАЗ-500 грузоподъемностью 7,5 т.

Определите количество колонн, которое нужно направить в каждый район, чтобы перевезти наибольшее количество груза.

7.2. Из пункта А в пункт В ежедневно отправляются пассажирские и скорые поезда. Данные об организации перевозок следующие:

Поезда	Количество вагонов в поезде				
	багажный	почтовый	плацкарт	купейный	мягкий
Скорый	1	1	5	6	3
Пассажирский	1	–	8	4	1
Число пассажиров	–	–	58	40	32
Парк вагонов	12	8	81	70	26

Сколько должно быть сформировано скорых и пассажирских поездов, чтобы перевезти наибольшее количество пассажиров?

7.3. Четыре овощехранилища каждый день обеспечивают картофелем три магазина. Магазины подали заявки соответственно на 17, 12 и 32 т. Овощехранилища имеют соответственно 20, 20, 15 и 25 т. Тарифы (в д.е. за 1 т) указаны в следующей таблице:

Овощехранилища	Магазины		
	1	2	3
1	2	7	4
2	3	2	1
3	5	6	2
4	3	4	7

Составьте план перевозок, минимизирующий суммарные транспортные расходы.

7.4. Имеются два склада готовой продукции: A_1 и A_2 с запасами однородного груза 200 и 300 т. Этот груз необходимо доставить трем потребителям: B_1, B_2 и B_3 в количестве 100, 150, 250 т соответственно. Стоимость перевозки 1 т груза из склада A_1 потребителям B_1, B_2 и B_3 равна 5, 3, 6 д.е., а из склада A_2 тем же потребителям – 3, 4, 2 д.е. соответственно.

Составьте план перевозок, минимизирующий суммарные транспортные расходы.

7.5. При откорме каждое животное должно получать не менее 9 ед. белков, 8 ед. углеводов и 11 ед. протеина. Для составления рациона используют два вида корма, представленных в следующей таблице:

Питательные вещества	Количество единиц питательных веществ на 1 кг	
	корма 1	корма 2
Белки	3	1
Углеводы	1	2
Протеин	1	6

Стоимость 1 кг корма первого вида – 4 д.е., второго – 6 д.е.

Составьте дневной рацион питательности, имеющий минимальную стоимость.

7.6. Хозяйство располагает следующими ресурсами: площадь – 100 ед., труд – 120 ед., тяга – 80 ед. Хозяйство производит четыре вида продукции $П_1$, $П_2$, $П_3$, $П_4$. Организация производства характеризуется следующей таблицей:

Продукция	Затраты на 1 ед. продукции			Доход от единицы продукции
	площадь	труд	тяга	
$П_1$	2	2	2	1
$П_2$	3	1	3	4
$П_3$	4	2	1	3
$П_4$	5	4	1	5

Составьте план выпуска продукции, обеспечивающий хозяйству максимальную прибыль.

7.7. Цех выпускает трансформаторы двух видов. Для изготовления трансформаторов обоих видов используются железо и проволока. Общий запас железа – 3 т, проволоки – 18 т. На один трансформатор первого вида расходуются 5 кг железа и 3 кг проволоки, а на один трансформатор второго вида расходуются 3 кг железа и 2 кг проволоки. За каждый реализованный трансформатор первого вида завод получает прибыль 3 д. е., второго – 4 д. е.

Составьте план выпуска трансформаторов, обеспечивающий заводу максимальную прибыль.

7.8. Совхоз отвел три земельных массива размером 5000, 8000, 9000 га на посевы ржи, пшеницы, кукурузы. Средняя урожайность в центнерах на 1 га по массивам указана в следующей таблице:

Посевы	Массивы		
	I	II	III
Рожь	12	14	15
Пшеница	14	14	22
Кукуруза	30	35	25

За 1 ц ржи совхоз получает 2 д. е., за 1 ц пшеницы – 2,8 д. е., за 1 ц кукурузы – 1,4 д. е.

Сколько гектаров и на *каких* массивах совхоз должен отвести на каждую культуру, чтобы получить максимальную выручку, если по плану он обязан сдать не менее 1900 т ржи, 158 000 т пшеницы и 30 000 т кукурузы?

7.9. Три типа самолетов следует распределить между четырьмя авиалиниями. Данные об организации процесса перевозок приведены в следующей таблице:

Тип самолета	Число самолетов, ед.	Месячный объем перевозок одним самолетом по авиалиниям, ед.				Эксплуатационные расходы на один самолет по авиалиниям, д. е.			
		I	II	III	IV	I	II	III	IV
1	50	15	10	20	50	15	20	25	40
2	20	20	25	10	10	70	28	15	45
3	30	35	50	30	45	40	70	50	65

Распределите самолеты по авиалиниям так, чтобы при минимальных суммарных эксплуатационных затратах перевезти по каждой из четырех авиалиний соответственно не менее 300, 200, 1000, 500 ед. груза.

7.10. Имеются четыре оперативные базы и три цели. В силу различия в типах самолетов и высоте полета вес бомб, доставляемых с любой базы к любой цели, определяется по следующей таблице:

База	Цель		
	1	2	3
1	8	6	5
2	6	6	6
3	10	8	4
4	8	6	4

Дневная интенсивность каждой базы составляет 150 самолето-вылетов в день. На каждую цель необходимо организовать 200 самолето-вылетов в день.

Определите план вылетов с каждой базы к каждой цели, дающий максимальный общий вес бомб, доставляемых к целям.

7.11. Из трех продуктов – I, II, III составляется смесь. В состав смеси должно входить не менее 6 ед. химического вещества A, 8 ед. – вещества B и не менее 12 ед. вещества C. Структура химических веществ приведена в следующей таблице:

Продукт	Содержание химического вещества в 1 ед. продукции			Стоимость 1 ед. продукции
	A	B	C	
I	2	1	3	2
II	1	2	4	3
III	3	1,5	2	2,5

Составьте наиболее дешевую смесь.

7.12. В институте проводится конкурс на лучшую стенгазету. Одному студенту дано следующее поручение:

• купить акварельной краски по цене 30 д. е. за коробку, цветные карандаши по цене 20 д. е. за коробку, линейки по цене 12 д. е., блокноты по цене 10 д. е.;

• красок нужно купить не менее трех коробок, блокнотов – столько, сколько коробок карандашей и красок вместе, линеек не более пяти. На покупки выделяется не менее 300 д. е.

В *каком количестве* студент должен купить указанные предметы, чтобы общее число предметов было наибольшим?

7.13. Заводы № 1, 2, 3 производят однородную продукцию в количестве соответственно 500, 400 и 510 единиц. Себестоимость производства единицы продукции на заводе № 1 составляет 25 д. е., на заводе № 2 – 20 д. е., на заводе № 3 – 23 д. е. Продукция отправляется в пункты A, B, C, потребности которых равны 310, 390 и 450 единицам. Стоимости перевозок 1 ед. продукции заданы матрицей

$$C = \begin{pmatrix} 7 & 5 & 1 \\ 2 & 3 & 2 \\ 3 & 5 & 4 \end{pmatrix}.$$

Составьте оптимальный план перевозок продукции при условии, что коммуникации между заводом № 2 и пунктом A не позволяют пропускать в рассматриваемый период более 250 единиц продукции.

7.14. Цех выпускает три вида деталей – A, B, C. Каждая деталь обрабатывается тремя станками. Организация производства в цехе характеризуется следующей таблицей:

Станок	Длительность обработки детали, мин.			Фонд времени, час
	A	B	C	
I	12	10	9	220
II	15	18	20	400
III	6	4	4	100
Отпускная цена за одну деталь	30	32	30	

Составьте план загрузки станков, обеспечивающий цеху получение максимальной прибыли.

7.15. Предприятие должно выпускать два вида продукции – A и B, используя при этом последовательно четыре станка. Данные о технологическом процессе указаны в следующей таблице:

Станок	Трудоемкость на 1 ед. продукции		Фонд времени, час.
	A	B	
1	3	3	15
2	2	6	18
3	4	0	16
4	1	2	8
Прибыль на 1 ед. продукции (д. е.)	2	3	

Составьте план выпуска продукции, обеспечивающий предприятию наибольшую прибыль.

7.16. На предприятии для производства запасных частей для автомобилей используются три вида ресурсов. Выпускаются три вида запасных частей. Организация производства на предприятии характеризуется следующей таблицей:

Ресурсы	Расход материалов на производство одной запасной части, кг			Запас ресурсов, кг
	1	2	3	
I	5	5	2	1200
II	4	–	3	300
III	–	2	4	800
Прибыль от реализации одной запасной части (д. е.)	5	8	6	

Составьте план производства запасных частей, обеспечивающий предприятию максимальную прибыль.

7.17. Имеются три специализированные мастерские по ремонту двигателей. Их производственные мощности равны соответственно 100, 700, 980 ремонтов в год. В пяти районах, обслуживаемых этими мастерскими, потребность в ремонте равна соответственно 90, 180, 150, 120, 80 двигателей в год. Затраты на перевозку одного двигателя из районов к мастерским следующие:

Районы	Мастерские		
	1	2	3
1	4,5	3,7	8,3
2	2,1	4,3	2,4
3	7,5	7,1	4,2
4	5,3	1,2	6,2
5	4,1	6,7	3,1

Спланируйте количество ремонтов каждой мастерской для каждого из районов, минимизирующее суммарные транспортные расходы.

7.18. Нефтеперерабатывающий завод получает четыре полуфабриката: 400 тыс. л алкилата, 250 тыс. л крекинг-бензина, 350 тыс. л бензина прямой перегонки и 100 тыс. л изопентона. В результате смешивания этих четырех компонентов в разных пропорциях образуется три сорта авиационного бензина: бензин А-2:3:5:2, бензин В-3:1:2:1, бензин С-2:2:1:3. Стоимость 1 тыс. л указанных сортов бензина характеризуется числами 120 д.е., 100 д. е., 150 д. е.

Составьте план выпуска разных сортов авиационного бензина из условия получения максимальной стоимости всей продукции.

7.19. Планируется нанесение удара по некоторому объекту тремя различными видами оружия: оружием А – в течение 3 мин., оружием Б – в течение 5 мин., оружием В – в течение 4 мин. Возможности средств обеспечения стрельбы таковы, что при применении оружия А в течение 3 мин., оружия Б в течение 2 мин., оружия В в течение 4 мин. общее количество залпов не должно превышать 15. При применении оружия А в течение 2 мин. и оружия В в течение 3 мин. общее количество залпов не должно превышать 8 ед. Кроме того, для преодоления противодействия противника необходимо, чтобы количество залпов оружием В за 1 мин. было больше, чем 5 ед.

Рассчитайте темп стрельбы (количество залпов в 1 мин.) всеми видами оружия, при котором общее количество залпов в ударе будет наибольшим.

7.20. Имеется 5 ракет и 5 целей. Вероятность поражения цели каждой из ракет задана в следующей таблице:

Ракеты	Цели				
	1	2	3	4	5
1	0,12	0,02	0,50	0,43	0,15
2	0,71	0,18	0,81	0,05	0,26
3	0,84	0,76	0,26	0,37	0,52
4	0,22	0,45	0,83	0,81	0,65
5	0,49	0,02	0,50	0,26	0,27

Распределите ракеты по целям так, чтобы математическое ожидание числа пораженных целей было максимальным.

7.21. Для участия в соревнованиях спортклуб должен выставить команду, состоящую из спортсменов I и II разрядов. Соревнования проводятся по бегу, прыжкам в высоту, прыжкам в длину. В беге должны участвовать 5 спортсменов, в прыжках в длину – 8 спортсменов, а в прыжках в высоту – не более 10. Количество очков, гарантируемых спортсмену каждого разряда по каждому виду, указано в следующей таблице:

Разряд	Бег	Прыжки в высоту	Прыжки в длину
I	4	5	5
II	2	3	3

Распределите спортсменов в команды так, чтобы сумма очков команды была наибольшей, если известно, что в команде I разряд имеют только 10 спортсменов.

7.22. Предприятию задана месячная программа на изготовление четырех типов изделий в количествах соответственно 5000, 2000, 3000 и 1800 шт. На предприятии имеется три группы станков с различной производительностью. Суммарное допустимое время для каждой группы станков составляет соответственно 800, 1000, 1500 час. Данные о технологическом процессе указаны в следующей таблице:

№ группы станков	Нормы времени на изготовление одного изделия, час.				Издержки на изготовление одного изделия, д. е.			
	I	II	III	IV	I	II	III	IV
1	0,5	0,15	0,4	0,6	0,12	0,2	0,3	0,25
2	0,4	0,12	0,2	0,5	0,16	0,14	0,35	0,2
3	0,42	0,14	0,35	0,45	0,17	0,25	0,4	0,3

Распределите изделия по станкам так, чтобы месячная программа была выполнена при наименьших издержках.

7.23. Звероферма выращивает черно-бурых лисиц и песцов. На звероферме имеется 10 000 клеток. В одной клетке могут быть либо две лисы, либо 1 песец. По плану на ферме должно быть не менее 3000 лис и 6000 песцов. В одни сутки необходимо выдавать каждой лисе корма – 4 ед., а каждому песцу – 5 ед. Ферма ежедневно может иметь не более 200 000 единиц корма. От реализации одной шкурки лисы ферма получает прибыль 10 д.е., а от реализации одной шкурки песца – 5 д. е.

Какое количество лисиц и песцов нужно держать на ферме, чтобы получить наибольшую прибыль?

7.24. *Найдите* оптимальное распределение трех видов механизмов, имеющихся в количествах 45, 20 и 35, между четырьмя участками работ, потребности которых соответственно равны 10, 20, 30, 40 при следующей матрице производительности:

$$W = \begin{pmatrix} 3 & 2 & 1 & 0 \\ 2 & 5 & 3 & 4 \\ 1 & 0 & 4 & 2 \end{pmatrix}.$$

7.25. Имеются два элеватора, в которых сосредоточено соответственно 4200 и 1200 т зерна. Зерно необходимо перевезти трем хлебозаводам в количестве 1000, 2000 и 1600 т каждому. Расстояние от элеватора до хлебозаводов указано в следующей таблице:

Элеваторы	Хлебозаводы		
	1	2	3
1	20	30	50
2	60	20	40

Затраты на перевозку 1 т продукта на 1 км составляют 25 д.е.

Спланируйте перевозки зерна из условия минимизации транспортных расходов.

7.26. Из двух сортов бензина образуются две смеси – А и В. Смесь А содержит бензина 60% 1-го сорта и 40% 2-го сорта; смесь В – 80% 1-го сорта и 20% 2-го сорта. Цена 1 кг смеси А – 10 д.е., а смеси В – 12 д.е.

Составьте план образования смесей, при котором будет получен максимальный доход, если в наличии имеется бензина 50 т 1-го сорта и 30 т 2-го сорта.

7.27. Имеются две почвенно-климатические зоны, площади которых соответственно равны 0,8 и 0,6 млн га. Данные об урожайности зерновых культур приведены в следующей таблице:

Зерновые культуры	Урожайность (ц/га)		Стоимость 1 ц, д. е.
	1-я зона	2-я зона	
Озимые	20	25	8
Яровые	25	20	7

Определите размеры посевных площадей озимых и яровых культур, необходимые для достижения максимального выхода продукции в стоимостном выражении.

7.28. На строительство четырех объектов кирпич поступает с трех заводов. Заводы имеют на складах соответственно 50, 100 и 50 тыс. шт. кирпича. Объекты требуют соответственно 50, 70, 40, 40 тыс. шт. кирпича. Тарифы (в д.е./тыс. шт) приведены в следующей таблице:

Заводы	Объекты			
	1	2	3	4
1	2	6	2	3
2	5	2	1	7
3	4	5	7	8

Составьте план перевозок, минимизирующий суммарные транспортные расходы.

7.29. Для полива различных участков сада, на которых растут сливы, яблони, груши, служат три колодца. Колодцы могут дать соответственно 180, 90 и 40 ведер воды. Участки сада требуют для полива соответственно 100, 120 и 90 ведер воды. Расстояния (в метрах) от колодцев до участков сада указаны в следующей таблице:

Колодцы	Участки		
	сливы	яблони	груши
1	10	5	12
2	23	28	33
3	43	40	39

Определите, как лучше организовать полив?

7.30. На заводе выпускают изделия четырех типов. От реализации 1 ед. каждого изделия завод получает прибыль соответственно 2, 1, 3, 5 д.е. На изготовление изделий расходуются ресурсы трех типов: энергия, материалы, труд. Данные о технологическом процессе приведены в следующей таблице:

Ресурсы	Затраты ресурсов на единицу изделия				Запасы ресурсов, ед.
	I	II	III	IV	
Энергия	2	3	1	2	30
Материалы	4	2	1	2	40
Труд	1	2	3	1	25

Спланируйте производство изделий так, чтобы прибыль от их реализации была наибольшей.

Решите задачи линейного программирования (7.31 – 7.60) графическим методом, проведите анализ на чувствительность.

Во всех задачах $x_1 \geq 0$, $x_2 \geq 0$.

7.31. $W = 2x_1 - 5x_2 \to \min;$

$$\begin{cases} 3x_1 + 4x_2 \leq 6; \\ 2x_1 + 3x_2 \leq 4. \end{cases}$$

7.32. $W = x_1 - 4x_2 \to \min;$

$$\begin{cases} 3x_1 + 5x_2 \geq 8; \\ -3x_1 + 10x_2 \leq 16. \end{cases}$$

7.33. $F = x_1 + x_2 \to \max;$

$$\begin{cases} x_1 + 3x_2 \le 30; \\ 2x_1 + x_2 \le 20. \end{cases}$$

7.34. $W = 2x_1 + 2x_2 \to \min;$

$$\begin{cases} x_1 + x_2 \ge 1; \\ -x_1 + x_2 \le 1. \end{cases}$$

7.35. $F = 2x_1 + 3x_2 \to \max;$

$$\begin{cases} x_1 \ge 4; \\ x_2 \ge 3; \\ x_1 + x_2 \le 8. \end{cases}$$

7.36. $W = x_1 - 3x_2 \to \min;$

$$\begin{cases} x_1 + x_2 \le 3; \\ -x_1 + 2x_2 \le 5. \end{cases}$$

7.37. $W = x_1 - 3x_2 \to \min;$

$$\begin{cases} -x_1 + 2x_2 \le 6; \\ x_1 + 2x_2 \le 5. \end{cases}$$

7.38. $W = 2x_1 + 5x_2 \to \max;$

$$\begin{cases} x_1 + x \le 500; \\ x_1 \le 400; \\ x_2 \le 300. \end{cases}$$

7.39. $W = x_1 + 4x_2 \to \max;$

$$\begin{cases} x_1 + x_2 \le 7; \\ x_1 \le 3; \\ x_2 \le 1. \end{cases}$$

7.40. $W = 2x_1 + x_2 \to \max;$

$$\begin{cases} 2x_1 + 6x_2 \le 15; \\ 4x_1 + 3x_2 \le 11. \end{cases}$$

7.41. $W = 2x_1 + 2x_2 \to \min;$

$$\begin{cases} x_1 + x_2 \ge 4; \\ -x_1 + 2x_2 \le 8. \end{cases}$$

7.42. $W = 3x_1 + 2x_2 \to \max;$

$$\begin{cases} x_1 \ge 1; \\ x_2 \ge 0{,}6; \\ 0{,}1x_1 + 0{,}4x_2 \le 2. \end{cases}$$

7.43. $W = x_1 + x_2 \to \max;$

$$\begin{cases} 3x_1 + x_2 \le 20; \\ 2x_1 + 3x_2 \le 30. \end{cases}$$

7.44. $W = 5x_1 + x_2 \to \max;$

$$\begin{cases} 3x_1 + 6x_2 \le 11; \\ x_1 \le 2{,}75; \\ 3x_2 \le 1{,}1. \end{cases}$$

7.45. $W = 2x_1 + 7x_2 \to \max;$

$$\begin{cases} x_1 \ge 3; \\ x_2 \ge 4; \\ 2x_1 + 2x_2 \le 9. \end{cases}$$

7.46. $W = x_1 - 2x_2 \to \min;$

$$\begin{cases} x_1 + 10x_2 \le 1; \\ -2x_1 + 24x_2 \le 1. \end{cases}$$

7.47. $W = x_1 + 3x_2 \to \max;$

$$\begin{cases} 4x_1 + 8x_2 \le 17; \\ x_1 \le 3; \\ x_2 \le 2. \end{cases}$$

7.48. $W = 2x_1 + 4x_2 \to \max;$

$$\begin{cases} 4x_1 + x_2 \le 15; \\ x_1 + 6x_2 \le 7. \end{cases}$$

7.49. $W = 2x_1 + 3x_2 \to \min;$

$$\begin{cases} 5x_1 + 2x_2 \ge 3; \\ -4x_1 + 6x_2 \le 9. \end{cases}$$

7.50. $W = 4x_1 + 6x_2 \to \max;$

$$\begin{cases} x_1 + 15x_2 \le 32; \\ x_1 \le 31; \\ x_2 \le 2. \end{cases}$$

7.51. $W = 4x_1 - x_2 \to \min;$

$$\begin{cases} 4x_1 + 6x_2 \le 9; \\ -5x_1 + 8x_2 \le 4. \end{cases}$$

7.52. $W = 2x_1 + x_2 \to \min;$

$$\begin{cases} 5x_1 + 3x_2 \ge 7; \\ -2x_1 + 9x_2 \le 21. \end{cases}$$

7.53. $W = x_1 + 5x_2 \to \max;$

$$\begin{cases} 2x_1 + x_2 \le 8; \\ x_1 + 3x_2 \le 7. \end{cases}$$

7.54. $W = 2x_1 + x_2 \to \min;$

$$\begin{cases} 5x_1 + 3x_2 \ge 7; \\ -2x_1 + 9x_2 \le 21. \end{cases}$$

7.55. $W = x_1 + 4x_2 \to \max;$

$$\begin{cases} x_1 \ge 4; \\ x_2 \ge 5; \\ 3x_1 + x_2 \le 16. \end{cases}$$

7.56. $W = 3x_1 - 2x_2 \to \min;$

$$\begin{cases} 2x_1 + 5x_2 \le 8; \\ -3x_1 + 8x_2 \le 4. \end{cases}$$

7.57. $W = x_1 + 3x_2 \to \max;$

$$\begin{cases} 16x_1 + 4x_2 \le 9; \\ x_1 \le \dfrac{1}{3}; \\ x_2 \le \dfrac{2}{3}. \end{cases}$$

7.58. $W = 5x_1 + 4x_2 \to \max;$

$$\begin{cases} 4x_1 + x_2 \le 9; \\ x_1 + 3x_2 \le 6. \end{cases}$$

7.59. $W = 3x_1 + x_2 \to \min;$

$$\begin{cases} 4x_1 + x_2 \ge 5; \\ -3x_1 + 10x_2 \le 50. \end{cases}$$

7.60. $W = 3x_1 + 0{,}5x_2 \to \max;$

$$\begin{cases} x_1 \le 2; \\ x_2 \ge 1{,}8; \\ 2x_1 + 5x_2 \ge 12. \end{cases}$$

Задачи линейного программирования (7.61 – 7.90) решите симплекс-методом и проведите анализ моделей на чувствительность, сформулируйте двойственную задачу к исходной и решите ее.

7.61. $\max L = x_1 - 2x_2 + 2x_3 + 3x_4;$

$$\begin{aligned} &x_1 + x_2 + 2x_3 + 2x_4 = 8; \\ &2x_1 + 2x_2 + x_3 + x_4 = 10; \\ &x_1 - 2x_2 + x_3 + 2x_4 = 1; \end{aligned}$$

$$x_j \geq 0; j = \overline{1, 4}.$$

7.62. $\max L = 2x_1 + x_2 + x_3 + 2x_4;$

$$\begin{aligned} &x_1 + 2x_2 + x_3 + 2x_4 = 16; \\ &2x_1 + x_2 + 2x_3 + x_4 = 14; \\ &2x_1 + 2x_2 - 2x_3 + x_4 = 4; \end{aligned}$$

$$x_j \geq 0; j = \overline{1, 4}.$$

7.63. $\min L = 3x_1 + 2x_2 + x_3 + x_4;$

$$\begin{aligned} &2x_1 + 2x_2 + 3x_4 = 9; \\ &x_2 + 2x_3 + x_4 = 4; \\ &x_1 + 2x_2 + 2x_3 + 2x_4 = 8; \end{aligned}$$

$$x_j \geq 0; j = \overline{1, 4}.$$

7.64. $\min L = 3x_1 + 2x_2 + x_3 + 2x_4;$

$$\begin{aligned} &2x_1 + 3x_2 + 3x_4 = 10; \\ &x_2 + 2x_3 + x_4 = 4; \\ &x_1 + 2x_2 + 2x_3 + 2x_4 = 8; \end{aligned}$$

$$x_j \geq 0; j = \overline{1, 4}.$$

7.65. $\max L = x_1 + 2x_2 + 3x_3 + x_4;$

$$\begin{aligned} &2x_1 + x_2 + 3x_3 + x_4 = 12; \\ &x_1 + 2x_2 + x_3 + 2x_4 = 8; \\ &3x_1 + 3x_2 + x_3 + 3x_4 = 15; \end{aligned}$$

$$x_j \geq 0; j = \overline{1, 4}.$$

7.66. $\max L = 2x_1 - x_2 + 3x_3 - 2x_4;$

$$\begin{aligned} &x_1 + x_2 + 2x_3 - x_4 = 3; \\ &2x_1 + x_2 - x_3 + 2x_4 = 4; \\ &x_1 + 2x_2 + x_3 + x_4 = 5; \end{aligned}$$

$$x_j \geq 0; j = \overline{1, 4}.$$

7.67. $\min L = 2x_1 + x_2 + 2x_3 + 2x_4;$

$$\begin{aligned} &2x_1 + x_2 + 2x_3 + x_4 = 8; \\ &x_1 + 2x_2 + x_3 + 2x_4 = 10; \\ &2x_1 + x_2 + 2x_3 + 2x_4 = 10; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

7.68. $\min L = x_1 + 2x_2 + x_3 + x_4;$

$$\begin{aligned} &x_1 + x_2 - x_3 + x_4 = 4; \\ &2x_1 + x_2 + 2x_3 - x_4 = 4; \\ &x_1 - x_2 + x_3 + x_4 = 2; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

7.69. $\min L = 4x_1 + 2x_2 + 2x_3 + x_4;$

$$\begin{aligned} &x_1 + x_2 + x_3 + 2x_4 = 8; \\ &2x_1 + x_2 + x_3 + 2x_4 = 10; \\ &x_1 + x_2 + 2x_3 - 2x_4 = 6; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

7.70. $\min L = x_1 + 2x_2 + 3x_3 + 4x_4;$

$$\begin{aligned} &x_1 + x_2 - 2x_3 + x_4 = 2; \\ &x_1 - 2x_2 + x_3 + 2x_4 = 4; \\ &x_1 + 2x_2 + 2x_3 + 2x_4 = 8; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

7.71. $\max L = x_1 + 2x_2 + 3x_3 - x_4;$

$$\begin{aligned} &x_1 + x_2 + x_3 + x_4 = 4; \\ &x_1 + 2x_2 + x_3 + 2x_4 = 6; \\ &x_1 + 2x_2 + 2x_3 + x_4 = 6; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

7.72. $\min L = x_1 - 2x_2 + 3x_3 + x_4;$

$$\begin{aligned} &x_1 + x_2 + 2x_3 + x_4 = 7; \\ &x_1 - 2x_2 + x_3 + 2x_4 = 1; \\ &3x_1 + x_2 + 3x_3 + 2x_4 = 13; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

7.73. $\max L = 3x_1 + x_2 + 2x_3 + x_4;$

$$\begin{aligned} &x_1 + 2x_2 + 2x_4 = 6; \\ &2x_2 + 2x_3 + x_4 = 7; \\ &x_1 + x_2 + x_3 + 2x_4 = 7; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.74. $\max L = 2x_1 + x_2 + 3x_3 + 2x_4;$

$$\begin{aligned} &x_2 + 2x_3 + 2x_4 = 8; \\ &2x_1 + 2x_2 + x_3 + x_4 = 9; \\ &2x_1 + 2x_3 + x_4 = 8; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.75. $\max L = 2x_1 + x_2 - x_3 + 2x_4;$

$$\begin{aligned} &x_1 + 2x_2 = 6; \\ &x_2 + x_3 + 2x_4 = 6; \\ &x_1 + 2x_2 + 2x_3 = 10; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.76. $\min L = x_1 + 2x_2 - x_3 + 3x_4;$

$$\begin{aligned} &x_1 + 2x_3 + 2x_4 = 5; \\ &x_1 + x_2 + 2x_3 = 4; \\ &2x_2 + x_3 = 4; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.77. $\min L = x_1 + x_2 - 2x_3 + 2x_4;$

$$\begin{aligned} &x_1 + x_2 + x_4 = 5; \\ &2x_1 + x_3 = 3; \\ &2x_1 + x_3 = 6; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.78. $\min L = x_1 - 2x_2 + 2x_3 + 3x_4;$

$$\begin{aligned} &x_1 + x_2 + 2x_4 = 4; \\ &x_2 + 2x_3 + x_4 = 6; \\ &x_1 - 2x_2 + x_3 + x_4 = 6; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.79. $\min L = -2x_2 + x_3;$

$$\begin{aligned}&-2x_1 + x_2 + x_4 = 0;\\&x_1 - x_2 + x_3 = 2;\\&2x_1 - x_2 + 4x_3 - x_4 = 12;\end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.80. $\min L = 4x_1 + 2x_2 + 2x_3 + x_4;$

$$\begin{aligned}&x_1 + x_2 + x_3 + 2x_4 = 8;\\&2x_1 + x_2 + x_3 + 2x_4 = 10;\\&x_1 + x_2 + x_3 - 2x_4 = 6;\end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.81. $\max L = 4x_1 + 3x_2 + 2x_3 + x_4;$

$$\begin{aligned}&x_1 - 3x_2 + x_3 + x_4 = 6;\\&x_1 - 2x_2 + x_3 + 2x_4 = 4;\\&x_1 + x_3 = 1;\end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.82. $\min L = x_1 + x_2 + x_3 + x_4;$

$$\begin{aligned}&3x_1 + 2x_2 + 5x_3 + x_4 = 22;\\&2x_1 + 5x_2 + 4x_3 + x_4 = 24;\\&3x_1 + 4x_2 + 5x_3 + x_4 = 26;\end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.83. $\max L = x_1 - x_2 + x_3 - x_4;$

$$\begin{aligned}&3x_1 + x_3 + 2x_4 = 2;\\&-x_1 + x_2 + x_3 + 5x_4 = 5;\end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 4}.$$

7.84. $\max L = 2x_1 - x_2 + 3x_3 - 5x_4;$

$$\begin{aligned}&x_1 + x_2 + x_3 = 6;\\&x_2 + x_3 + x_4 = 9;\\&x_3 + x_4 + x_5 = 12;\\&x_4 + x_5 + x_6 = 15;\end{aligned}$$

$$x_j \geq 0;\ j = \overline{1, 6}.$$

7.85. $\min L = 2x_1 - x_2 + 3x_3 - 2x_4 + x_5;$

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

$$x_j \geq 0;\ j = \overline{1,\ 5}.$$

7.86. $\max L = 2x_1 - x_2 + x_3 - 3x_4 + 4x_5;$

$$\begin{aligned} x_1 - x_2 + 3x_3 - 18x_4 + 2x_5 &= -4; \\ 2x_1 - x_2 + 4x_3 - 21x_4 + 4x_5 &= 22; \\ 3x_1 + 2x_2 + 8x_3 - 43x_4 + 11x_5 &= 38; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 5}.$$

7.87. $\min L = 3x_1 - 2x_2 + x_3 + 8x_5;$

$$\begin{aligned} x_1 - 2x_2 - 10x_4 + 3x_5 &= 12; \\ 2x_1 - 3x_2 - 15x_3 - 10x_4 + 6x_5 &= 28; \\ 2x_1 - x_2 + 8x_3 - 6x_4 + 10x_5 &= 40; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 5}.$$

7.88. $\max L = 3x_1 + 4x_2 + 2x_3 + 6x_4;$

$$\begin{aligned} 2x_1 + 3x_2 + 2x_3 + 3x_4 &= 10; \\ 2x_2 + 2x_3 + x_4 &= 7; \\ x_1 + 3x_2 + 2x_3 + 5x_4 &= 8; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 5}.$$

7.89. $\min L = 2x_1 + x_2 + x_3 + 2x_4;$

$$\begin{aligned} 2x_1 + x_2 + 3x_3 + x_4 &= 20; \\ x_1 + 2x_2 + x_3 + 2x_4 &= 30; \\ 3x_1 + 3x_2 + x_3 + 3x_4 &= 40; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

7.90. $\max L = x_1 + 2x_2 + 3x_3 + 4x_4;$

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 60; \\ x_1 + 2x_2 + x_3 + 4x_4 &= 9; \\ 2x_1 + 3x_2 + 4x_3 + 1x_4 &= 12; \end{aligned}$$

$$x_j \geq 0;\ j = \overline{1,\ 4}.$$

Глава 8
Транспортные задачи линейного программирования

8.1. Постановка задачи

Под термином «транспортные задачи» понимается широкий круг задач не только транспортного характера. Общим для них является, как правило, распределение ресурсов, находящихся у m производителей (поставщиков), по n потребителям этих ресурсов.

На автомобильном транспорте наиболее часто встречаются следующие задачи, относящиеся к транспортным:

- прикрепление потребителей ресурса к производителям;
- привязка пунктов отправления к пунктам назначения;
- взаимная привязка грузопотоков прямого и обратного направлений;
- отдельные задачи оптимальной загрузки промышленного оборудования;
- оптимальное распределение объемов выпуска промышленной продукции между заводами-изготовителями и др.

Рассмотрим экономико-математическую модель прикрепления пунктов отправления к пунктам назначения. Имеются m пунктов отправления груза и объемы отправления по каждому пункту a_1, a_2, ..., a_m. Известна потребность в грузах b_1, b_2, ..., b_n по каждому из n пунктов назначения. Задана матрица стоимостей доставки по каждому варианту c_{ij}, $i = \overline{1,m}, j = \overline{1,n}$. Необходимо рассчитать оп-

Таблица 8.1

Исходные данные

Поставщики	Потребители				Запасы (объемы отправления)
	B_1	B_2	...	B_n	
A_1	c_{11} x_{11}	c_{12} x_{12}	...	c_{1n} x_{1n}	a_1
A_2	c_{21} x_{21}	c_{22} x_{22}	...	c_{2n} x_{2n}	a_2
...	...	...	...	...	$\vdots$
A_m	c_{m1} x_{m1}	c_{m2} x_{m2}	...	c_{mn} x_{mn}	a_m
Потребность	b_1	b_2	...	b_n	

тимальный план перевозок, т. е. определить, сколько груза должно быть отправлено из каждого i-го пункта отправления (от поставщика) в каждый j-й пункт назначения (до потребителя) x_{ij} с минимальными транспортными издержками.

В общем виде исходные данные представлены в табл. 8.1.

Транспортная задача называется *закрытой*, если суммарный объем отправляемых грузов $\left(\sum_{i=1}^{m} a_i\right)$ равен суммарному объему потребности в этих грузах по пунктам назначения $\left(\sum_{j=1}^{n} b_j\right)$:

$$\sum_{i=1}^{m} a_j = \sum_{j=1}^{n} b_j. \tag{8.1}$$

Если такого равенства нет (потребности выше запасов или наоборот), задачу называют *открытой*, т. е.:

$$\sum_{i=1}^{m} a_j \neq \sum_{j=1}^{n} b_j. \tag{8.2}$$

Для написания модели необходимо все условия (ограничения) и целевую функцию представить в виде математических уравнений.

Все грузы из i пунктов должны быть отправлены, т. е.:

$$\sum_{j=1}^{n} x_{ij} = a_i, \quad i = \overline{1,m}. \tag{8.3}$$

Все j пункты (потребители) должны быть обеспечены грузами в плановом объеме:

$$\sum_{i=1}^{m} x_{ij} = b_j, \quad j = \overline{1,n}. \tag{8.4}$$

Суммарные объемы отправления должны равняться суммарным объемам назначения:

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j. \tag{8.5}$$

Должно выполняться условие неотрицательности переменных: $x_{ij} \geq 0$, $i = \overline{1,m}$; $j = \overline{1,n}$. Перевозки необходимо осуществить с минимальными транспортными издержками (функция цели):

$$Z_{\min} = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}. \tag{8.6}$$

В модели (8.3) – (8.6) вместо матрицы стоимостей перевозок (c_{ij}) могут задаваться матрицы расстояний. В таком случае в качестве целевой функции рассматривается минимум суммарной транспортной работы. Как видно из выражения (8.5), уравнение баланса является обязательным условием решения транспортной задачи. Поэтому, когда в исходных условиях дана открытая задача, то ее необходимо привести к закрытой форме. В случае если

- потребности по пунктам назначения превышают запасы пунктов отправления, то вводится фиктивный поставщик с недостающим объемом отправления;
- запасы поставщиков превышают потребности потребителей, то вводится фиктивный потребитель с необходимым объемом потребления.

Варианты, связывающие фиктивные пункты с реальными, имеют нулевые оценки. После введения фиктивных пунктов задача решается как закрытая.

Транспортным задачам присущи следующие особенности:

- распределению подлежат однородные ресурсы;
- условия задачи описываются только уравнениями;
- все переменные выражаются в одинаковых единицах измерения;
- во всех уравнениях коэффициенты при неизвестных равны единице;
- каждая неизвестная встречается только в двух уравнениях системы ограничений.

Транспортные задачи могут решаться симплекс-методом. Однако перечисленные особенности позволяют для транспортных задач применять более простые методы решения.

8.2. Алгоритм метода потенциалов

Наиболее распространенным методом решения транспортных задач является *метод потенциалов.*

Решение задачи методом потенциалов включает следующие *этапы*:

1) разработку начального плана (опорного решения);
2) расчет потенциалов;
3) проверку плана на оптимальность;
4) поиск максимального звена неоптимальности (если условие п. 3 не было достигнуто);
5) составление контура перераспределения ресурсов;

6) определение минимального элемента в контуре перераспределения и перераспределение ресурсов по контуру;

7) получение нового плана.

Описанная процедура повторяется несколько раз (итераций), пока не будет найдено оптимальное решение. Вычислительный алгоритм для каждой итерации не меняется.

Для транспортной задачи существует несколько методов отыскания начального плана (опорного решения):

- метод северо-западного угла;
- метод минимальной стоимости;
- метод двойного предпочтения и т. д.

Вычислительный алгоритм метода потенциалов рассмотрим на примере решения конкретной задачи прикрепления пунктов отправления $i = \overline{1,3}$ к пунктам назначения $j = \overline{1,4}$. В соответствии с принятыми в подразд. 8.1 обозначениями исходные данные задачи приведены в табл. 8.2.

Таблица 8.2

Исходные данные

Поставщики	Потребители				Запасы
	B_1	B_2	B_3	B_4	
A_1	1	2	3	4	60
A_2	4	3	2	0	80
A_3	0	2	2	1	100
Потребность	40	60	80	60	240

Начальный план можно составить одним из перечисленных выше методов. Воспользуемся наиболее простым методом – *методом северо-западного угла*. В соответствии с этим методом загрузка клеток (распределение объемов пунктов отправления по пунктам назначения) начинается с верхней левой клетки («северо-западная» часть таблицы) и продолжается вниз и вправо (по диагонали).

По указанному правилу загружаем первую клетку $(i - j) = (1 - 1)$ на основе следующего условия:

$$x_{11} = \min\{a_1;\ b_1\} = \min\{60;\ 40\} = 40.$$

Таким образом, первый пункт назначения загружен, а первый пункт отправления имеет остатки груза $\Delta a_1 = 60 - 40 = 20$, которые и распределяем на второй пункт назначения:

$$x_{12} = \min\{\Delta a_1;\ b_2\} = \min\{20;\ 60\} = 20;\ \Delta b_2 = 40.$$

Продолжая преобразования аналогичным образом, получаем:

$x_{22} = \min\{a_2; \Delta b_2\} = \min\{80; 40\} = 40$; $\Delta b_2 = 40$ и т. д.

Результаты начального плана и расчета потенциалов представлены в табл. 8.3.

Таблица 8.3

Начальный план перевозок

Поставщики	Потребители				Запасы	α_i
	B_1	B_2	B_3	B_4		
A_1	1 Р 40	2 20 З	3	4	60	0
A_2	4	3 40 Р	З 2 40	0	80	1
A_3	0 ⊕ З	2	2 Р 40	1 60	100	1
Потребность	40	60	80	60	240	
β_j	1	2	1	0		

В процессе решения после каждой итерации (в том числе и после получения допустимого решения) по загруженным клеткам проверяется выполнение следующего условия:

$$N = m + n - 1. \tag{8.7}$$

В нашем примере $m = 3$, $n = 4$, а число загруженных клеток равно 6, т. е. соответствует условию (8.7): $N = 3 + 4 - 1 = 6$. Если условие (8.7) не выполняется, план называется *вырожденным*. В этом случае в любые свободные клетки надо поставить столько нулей, чтобы с их учетом выполнялось условие (8.7). Клетка, в которой стоит ноль, считается занятой. Значение целевой функции по результатам расчета допустимого плана

$$Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} = 1 \cdot 40 + 2 \cdot 20 + 3 \cdot 40 + 2 \cdot 40 + 2 \cdot 40 + 1 \cdot 60 = 420 \text{ д.е.}$$

Расчет потенциалов выполняют по загруженным клеткам, для которых должно выполняться следующее равенство:

$$\alpha_i + \beta_j = c_{ij}, \tag{8.8}$$

где α_i – потенциал i-й строки;

β_j – потенциал j-го столбца.

Вычисляя потенциалы по выражению (8.8), принимаем для первой строки $\alpha_1 = 0$. Используя загруженные клетки $(i - j) = (1 - 1)$, $(1 - 2)$, получаем:

$$\alpha_1 + \beta_1 = c_{11} = 0 + \beta_1 = 1;\ \beta_1 = 1;$$
$$\alpha_1 + \beta_2 = c_{12} = 0 + \beta_2 = 2;\ \beta_2 = 2.$$

Далее по загруженным клеткам (2 – 2), (2 – 3) определяем другие потенциалы:

$$\alpha_2 + \beta_2 = 3;\ \alpha_2 + 2 = 3;\ \alpha_2 = 1;$$
$$\alpha_2 + \beta_3 = 2;\ 1 + \beta_3 = 2;\ \beta_3 = 1.$$

Результаты расчета потенциалов представлены в табл 8.3.

Проверяем план на оптимальность по незагруженным клеткам, используя следующее неравенство:

$$\alpha_i + \beta_j \le c_{ij}. \tag{8.9}$$

Если для незагруженных клеток условие (8.9) выполняется, то план – оптимальный. По табл. 8.3 осуществляем проверку начального плана на оптимальность:

$(i - j) = (1–3)$, $0 + 1 \le 3$;
$(i - j) = (1–4)$, $0 + 0 \le 4$;
$(i - j) = (2–1)$, $1 + 1 \le 4$;
$(i - j) = (2–4)$, $1 + 0 > 0$, $\Delta c_{24} = 1$;
$(i - j) = (3–1)$, $1 + 1 > 0$, $\Delta c_{31} = 2$;
$(i - j) = (3–2)$, $1 + 2 > 2$, $\Delta c_{32} = 1$.

Итак, по трем клеткам условие (8.9) не выполняется, следовательно, начальный план требует улучшения. Характеристики Δc_{ij} показывают размер экономии транспортных издержек на 1 ед. перевозимого груза. В нашем примере наибольшую экономию можно получить по клетке $(i - j) = (3 - 1)$, где $\Delta c_{31} = 2 > \{\Delta c_{24};\ \Delta c_{32}\}$. Следовательно, клетку (3–1) необходимо загрузить за счет перераспределения ресурсов из других загруженных клеток. В табл. 8.3 клетку (3–1) помечаем знаком «+», так как здесь в начальном плане находится вершина максимальной неоптимальности.

Контур перераспределения ресурсов составляют по следующим правилам:

- этот контур представляет замкнутый многоугольник с вершинами в загруженных клетках, за исключением клетки с вершиной максимальной неоптимальности «+», и звеньями, лежащими вдоль строк и столбцов матрицы;
- ломаная линия должна быть связанной в том смысле, что из любой ее вершины можно попасть в любую другую вершину по звеньям ломаной цепи (по строке или по столбцу);
- в каждой вершине контура встречаются только два звена, одно из них располагается по строке, другое – по столбцу;
- число вершин контура четное, все они в процессе перераспределения делятся на загружаемые и разгружаемые;
- в каждой строке (столбце) имеются две вершины: одна – загружаемая, другая – разгружаемая.

В этой клетке намечаем одну из вершин контура и далее по вышеизложенным правилам строим контур, вершины которого будут находиться в клетках (3–1) – (1–1) – (1–2) – (2–2) – (2–3) – – (3–3). Вершины контура последовательно подразделяем на загружаемые (З) и разгружаемые (Р), начиная с вершины максимальной неоптимальности «+» (табл. 8.3).

Перераспределение ресурсов по контуру осуществляется с целью получения оптимального плана. В процессе перераспределения ресурсов по контуру в соответствии с условием неотрицательности переменных x_{ij} ни одно из этих значений не должно превратиться в отрицательное число. Поэтому анализируют только клетки, помеченные знаком Р, из которых выбирают клетку с минимальным объемом перевозок. В нашем примере $X_{\min} = \min\{40; 40; 40\} = 40$. Следовательно, клетки (1–1), (2–2), (3–3) полностью разгружаются. В клетке (1–2) загрузка увеличится на 40 и достигнет 60, в клетке (2–3) загрузка составит 40 + 40 = 80, и клетка (3–1) загрузится на 40 (табл. 8.4).

Проверяем условие $N = m + n - 1$. В нашем примере $m = 3$, $n = 4$, а число загруженных клеток равно 4, т. е. условие не выполняется и $6 \neq 4$. В процессе перераспределения ресурсов произошла полная разгрузка трех клеток, а мы должны освободить только одну клетку. В этом случае следует в две клетки проставить нули (нулевой ресурс) и считать условно их загруженными. Например, в клетки (1–1) и (3–3) проставим нулевой ресурс (рис. 8.4). Получение нового плана (итерации) осуществляется в том же порядке, который был рассмотрен выше, т. е.

- по загруженным клеткам (в соответствии с новой загрузкой) вычисляются потенциалы α_i и β_j;

Таблица 8.4

Первый план перевозок

Поставщики	Потребители				Запасы	α_i
	B_1	B_2	B_3	B_4		
A_1	1 0	2 60	3	4	60	0
A_2	4	3	2 80 Р	⊕ 0 З	80	–1
A_3	0 40	2	2 0 З	1 Р 60	100	–1
Потребность	40	60	80	60	240	
β_j	1	2	3	2		

- по незагруженным клеткам производится проверка плана на оптимальность;
- находится вершина максимальной неоптимальности и строится новый контур перераспределения и т. д., до тех пор, пока не будет найдено оптимальное решение, удовлетворяющее неравенству (8.9).

По результатам первой итерации имеем

$$Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} = 2 \cdot 60 + 2 \cdot 80 + 1 \cdot 60 + 0 \cdot 40 = 340.$$

Ниже приведены расчеты по второй итерации и оптимальный план.

Поиск потенциалов следующий:

$$\begin{array}{lll}
\alpha_1 + \beta_1 = 1; & 0 + \beta_1 = 1; & \beta_1 = 1; \\
\alpha_1 + \beta_2 = 2; & 0 + \beta_2 = 2; & \beta_2 = 2; \\
\alpha_2 + \beta_3 = 2; & \alpha_2 + 3 = 2; & \alpha_2 = -1; \\
\alpha_3 + \beta_1 = 0; & \alpha_3 + 1 = 0; & \alpha_3 = -1; \\
\alpha_3 + \beta_3 = 2; & -1 + \beta_3 = 2; & \beta_3 = 3; \\
\alpha_3 + \beta_4 = 1; & -1 + \beta_4 = 1; & \beta_4 = 2.
\end{array}$$

Проведем проверку на оптимальность:

$$(i - j) = (1\text{–}3);\ 0 + 3 \le 3;$$
$$(i - j) = (1\text{–}4);\ 0 + 2 < 4;$$
$$(i - j) = (2\text{–}1);\ 1 - 1 < 4;$$

$$(i-j) = (2\text{–}2);\ 2-1 < 3;$$
$$(i-j) = (3\text{–}2);\ 2-1 < 2;$$
$$(i-j) = (2\text{–}4);\ 2-1 > 0.$$

Клетку (2–4) необходимо загрузить.

В соответствии с перераспределением ресурсов по контуру получаем табл. 8.5, для которой вновь рассчитываем потенциалы α_i и β_j, и последовательность вычислений повторяется.

Таблица 8.5

Оптимальный план перевозок

Поставщики	Потребители				Запасы	α_i
	B_1	B_2	B_3	B_4		
A_1	1 0	2 60	3	4	60	0
A_2	4	3	2 20	0 60	80	–1
A_3	0 40	2	2 60	1	100	–1
Потребность	40	60	80	60	240	
β_j	1	2	3	1		

Для распределения, полученного в табл. 8.5, условие $\alpha_i + \beta_j \le c_{ij}$ выполняется, следовательно, план – оптимальный.

Транспортные издержки по оптимальному плану следующие:

$$Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} = 1 \cdot 0 + 2 \cdot 60 + 2 \cdot 20 + 0 \cdot 60 + 0 \cdot 40 + 2 \cdot 60 = 280 \text{ д.е.}$$

Таким образом, построением начального плана с последующим расчетом двух итераций получено оптимальное решение по прикреплению пунктов отправления грузов к пунктам назначения.

8.3. Усложненные задачи транспортного типа

Нами рассмотрена классическая транспортная задача, на которой показано, как используется метод потенциалов для нахождения оптимального плана. В экономике предприятия такие задачи встречаются крайне редко. Обычно при составлении экономикоматематической модели в задачи транспортного типа приходится вводить целый ряд дополнительных ограничений, а затем пользоваться методом потенциалов.

Ряд экономических задач легко сводимы к транспортной задаче. Рассмотрим наиболее часто встречающиеся *ситуации в экономике предприятия.*

1. Отдельные поставки от определенных поставщиков некоторым потребителям должны быть исключены (из-за отсутствия необходимых условий хранения, чрезмерной перегрузки коммуникаций и т. д.). Это ограничение требует, чтобы в матрице перевозок, содержащей оптимальный план, определенные клетки оставались свободными. Последнее достигается искусственным завышением затрат на перевозки c_{ij} в клетках, перевозки через которые следует запретить. При этом производят завышение величины c_{ij} до таких значений, которые будут заведомо больше всех, с которыми их придется сравнивать в процессе решения задачи.

2. На предприятии необходимо определить минимальные суммарные затраты на производство и транспортировку продукции. С подобной задачей сталкиваются при решении вопросов, связанных с оптимальным размещением производственных объектов. Здесь может оказаться экономически более выгодным доставлять сырье из более отдаленных пунктов, но зато при меньшей его себестоимости. В таких задачах за критерий оптимальности принимают сумму затрат на производство и транспортировку продукции.

3. Ряд транспортных маршрутов, по которым необходимо доставить грузы, имеют ограничения по пропускной способности. Если, например, по маршруту A_iB_j можно провести не более q единиц груза, то B_j столбец матрицы разбивается на два столбца – B'_j и B''_j. В первом столбце спрос принимается равным разности между действительным спросом b_j и ограничением q: $b'_j = b_j - q$, во втором – равным ограничению q, т. е. $b''_j = q$. Затраты c_{ij} в обоих столбцах одинаковы и равны данным, но в первом столбце B'_j, в клетке, соответствующей ограничению i, вместо истинного тарифа c_{ij} ставится искусственно завышенный тариф M (клетка блокируется). Затем задача решается обычным способом.

4. Поставки по определенным маршрутам обязательны и должны войти в оптимальный план независимо от того, выгодно это или нет. В этом случае уменьшают запас груза у поставщиков и спрос потребителей и решают задачу относительно тех поставок, которые необязательны. Полученное решение корректируют с учетом обязательных поставок.

5. Экономическая задача не является транспортной, но в математическом отношении подобна транспортной, так как описывается аналогичной моделью, например распределение производства

изделий между предприятиями, оптимальное закрепление механизмов по определенным видам работы.

6. Необходимо максимизировать целевую функцию задачи транспортного типа. В этой ситуации при составлении опорного плана в первую очередь стараются заполнить клетки с наиболее высокими значениями показателей c_{ij}. Выбор клетки, подлежащей заполнению при переходе от одного допустимого плана к другому, должен производиться не по минимальной отрицательной разнице $[c_{ij} - (\alpha_i + \beta_j)]$, а по максимальной положительной разнице $[c_{ij} - (\alpha_i + \beta_j)]$. Оптимальным будет план, которому в последней таблице сопутствуют свободные клетки с неположительными элементами: все разности $[c_{ij} - (\alpha_i + \beta_j)] \leq 0$.

7. Необходимо в одно время распределить груз различного рода по потребителям. Задачи данного типа называются *многопродуктовыми транспортными задачами*. В этих задачах поставщики m родов грузов разбиваются на m условных поставщиков, а потребители n родов грузов разбиваются на n условных потребителей. С учетом этой разбивки составляют полную транспортную таблицу. При этом заметим, что некоторые маршруты A_iB_j должны быть блокированы (закрыты), поскольку в данной постановке задачи грузы разного рода не могут заменять друг друга. Этим маршрутам A_iB_j должна соответствовать очень высокая стоимость перевозки. Многопродуктовую задачу не всегда обязательно описывать одной моделью. Например, если поставки грузов различного рода независимы, то задачу можно представить в виде комплекса транспортных задач по каждому роду груза. Однако если между грузами различного рода существует связь (например, одни из грузов можно заменить другими), то в общем случае исходную модель (задачу) не удается разбить на комплекс простых транспортных задач.

Рассмотрим примеры задач транспортного типа.

Пример 8.1. Одно фермерское хозяйство (A_1) имеет продовольственное зерно двух видов: 3 тыс. т – III класса и 4 тыс. т – IV класса. Второе фермерское хозяйство (A_2) также имеет зерно двух классов: 5 тыс. т – III класса и 2 тыс. т – IV класса. Зерно должно быть вывезено на два элеватора: на первый элеватор (B_1) необходимо поставить 2 тыс. т пшеницы III класса, 3 тыс. т пшеницы IV класса и остальные 2 тыс. т пшеницы любого класса.

Аналогично второй элеватор (B_2) должен получить 8,25 тыс. т, из них пшеницы – 1 тыс. т III класса и 1,5 тыс. т IV класса.

Стоимость перевозки в д. е. 1 т зерна составляет: из пункта A_1 в пункты B_1 и B_2 – 1 и 1,5 соответственно; из пункта A_2 в пункты B_1 и B_2 – 2 и 1 д. е. соответственно.

Составить оптимальный план перевозок.

Решение

Каждого поставщика условно разбиваем на две части согласно двум видам зерна (A_1^3 и A_1^4; A_2^3 и A_2^4), аналогично потребителей разбиваем на три части (пшеница III класса, IV класса и любой класс): B_1^3, B_1^4 и B_1^0, а также B_2^3, B_2^4 и B_2^0. Потребности превышают запасы, поэтому вводим фиктивного поставщика A_3. Часть клеток в таблице запираем большими числами M; например, в клетке (1; 2) стоит большое число. Это значит, что поставщик A_1^3 не может удовлетворить потребителя B_1^4 пшеницей IV класса за счет имеющейся пшеницы III класса.

С учетом сделанных замечаний составим первую таблицу (табл. 8.6).

Таблица 8.6

Исходные данные

Поставщики		Потребители						Запас, тыс. т
		B_1			B_2			
		B_1^3	B_1^4	B_1^0	B_2^3	B_2^4	B_2^0	
A_1	A_1^3	1	M	1	1,5	M	1,5	3
	A_1^4	M	1	1	M	1,5	1,5	4
A_2	A_2^3	2	M	2	1	M	1	5
	A_2^4	M	2	2	M	1	1	2
A_3		0	0	0	0	0	0	1,25
Спрос, тыс. т		2	3	2	1	1,5	5,75	15,25

Перевозки от фиктивного поставщика не производятся, поэтому $c_{51} = c_{52} = c_{53} = c_{54} = c_{55} = c_{56} = 0$. Величина M намного больше c_{ij}. Применяя метод потенциалов, в итоге получим таблицу с оптимальным решением (табл. 8.7).

Анализ решения.

Первый поставщик поставит на первый элеватор (B_1) пшеницу III класса ($x_{22} = 2$); пшеницу IV класса ($x_{22} = 3$), а также пшеницу любого класса (III или IV) ($x_{13} = 1$; $x_{23} = 1$).

Второй поставщик (A_2) поставит на второй элеватор (B_2) пшеницу III класса ($x_{31} = 1$), пшеницу IV класса ($x_{45} = 1,5$) и частично

Таблица 8.7

Оптимальное решение

Поставщики		Потребители						Запас, тыс. т
		B_1			B_2			
		B_1^3	B_1^4	B_1^0	B_2^3	B_2^4	B_2^0	
A_1	A_1^3	1 2	M	1 1	1,5	M	1,5	3
	A_1^4	M	1 3	1 1	M	1,5	1,5	4
A_2	A_2^3	2	M	2	1 1	M	1 4	5
	A_2^4	M	2	2	M	1 1,5	1 0,5	2
A_3		0	0	0	0	0	0 1,25	1,25
Спрос, тыс. т		2	3	2	1	1,5	5,75	15,25

любую пшеницу ($x_{36} = 4$; $x_{46} = 0,5$). Потребность элеватора в любой пшенице не удовлетворена на 1,25 тыс. т ($x_{56} = 1,25$). Минимальные затраты на перевозку составили: $Z_{\min} = 14$ д. е.

Пример 8.2. Модель производства с запасами.

Фирма переводит свой головной завод на производство определенного вида изделий, которые будут выпускаться в течение четырех месяцев. Величины спроса в течение этих четырех месяцев составляют 100, 200, 180 и 300 изделий соответственно. В каждый месяц спрос можно удовлетворить за счет:

- запасов изделий, произведенных в прошлом месяце, сохраняющихся для реализации в будущем;
- производства изделий в течение текущего месяца;
- избытка производства изделий в более поздние месяцы в счет невыполненных заказов.

Затраты на одно изделие в каждом месяце составляют 4 д. е. Изделие, произведенное для более поздней реализации, влечет за собой дополнительные издержки на хранение в 0,5 д. е. в месяц. С другой стороны, каждое изделие, выпускаемое в счет невыполненных заказов, облагается штрафом в размере 2 д. е. в месяц.

Объем производства изделий меняется от месяца к месяцу в зависимости от выпуска других изделий. В рассматриваемые четыре месяца предполагается выпуск 50, 180, 280 и 270 изделий соответственно.

Требуется составить план, имеющий минимальную стоимость производства и хранения изделий.

Решение

Задачу можно сформулировать как транспортную. Эквивалентность между элементами производственной и транспортной систем устанавливается следующим образом (табл. 8.8):

Таблица 8.8

Транспортная система	Производственная система
1. Исходный пункт i 2. Пункт назначения j 3. Предложение в пункте i 4. Спрос в пункте j 5. Стоимость перевозки из i и j	1. Период производства i 2. Период потребления j 3. Объем производства за период i 4. Реализация за период j 5. Стоимость производства и хранения за период i и j

Перед нами структура транспортной модели. Для рассматриваемой задачи стоимость «перевозки» изделия из периода i в период j выражается как:

$$c_{ij} = \begin{cases} \bullet \text{ стоимость производства в } i\text{-й период, } i = j; \\ \bullet \text{ стоимость производства в } i\text{-й период плюс стои-} \\ \quad \text{мость задержки от } i \text{ до } j,\ i < j; \\ \bullet \text{ стоимость производства в } i\text{-й период плюс штраф} \\ \quad \text{за нарушение срока, } i > j. \end{cases}$$

Из определения c_{ij} следует, что затраты в период i при реализации продукции в тот же период i ($i = j$) оцениваются только стоимостью производства. Если в период i производится продукция, которая будет потребляться позже ($i < j$), то имеют место дополнительные издержки, связанные с хранением. Аналогично производство в i-й период в счет невыполненных заказов $i > j$ влечет за собой дополнительные расходы в виде штрафа. Например,

$$c_{11} = 4 \text{ д. е.};$$
$$c_{24} = 4 + (0{,}5 + 0{,}5) = 5 \text{ д. е.};$$
$$c_{41} = 4 + (2 + 2 + 2) = 10 \text{ д. е.}$$

Исходная транспортная таблица выглядит следующим образом (табл. 8.9).

Таблица 8.9

Исходные данные

Период	Период				Объем производства
	1	2	3	4	
1	4	4,5	5	5,5	50
2	6	4	4,5	5	180
3	8	6	4	4,5	280
4	10	8	6	4	270
Спрос	100	200	180	300	

Задача решается обычным методом потенциалов на минимум затрат по производству и хранению продукции.

Пример 8.3. Имеются три сорта бумаги в количестве 10, 8 и 5 т, которую можно использовать на издание четырех книг тиражом 8000, 6000, 10 000, 15 000 экземпляров. Расход бумаги на одну книгу составляет: 0,6; 0,4; 0,8; 0,5 кг, а себестоимость тиража книги при использовании i-го сорта бумаги задается следующей матрицей (д. е.):

$$c_{ij} = \begin{pmatrix} 24 & 16 & 32 & 25 \\ 18 & 24 & 24 & 20 \\ 30 & 24 & 16 & 20 \end{pmatrix}.$$

Определить оптимальное распределение бумажных резервов.

Решение

Задача по своему экономическому смыслу не является транспортной, в то же время можно построить математическую модель, аналогичную транспортной задаче.

Потребности в бумаге легко определить, зная тираж и расход на одну книгу (т):

$$8000 \cdot 0{,}6 = 4{,}8;$$
$$15\ 000 \cdot 0{,}4 = 6;$$
$$6000 \cdot 0{,}8 = 4{,}8;$$
$$10\ 000 \cdot 0{,}5 = 5.$$

Общие запасы бумаги составляют 23 т, а общие потребности — 20,5 т, поэтому необходимо в таблицу ввести фиктивный тираж B_5^* с нулевыми затратами. В связи с тем что мы составляем модель относительно бумаги, а матрица c_{ij} характеризует себестоимость печатания книги, необходимо исходную матрицу преобразовать относительно единицы бумаги (каждый столбец матрицы c_{ij} разделим на количество бумаги, приходящейся на одну книгу).

Согласно изложенному составим первую таблицу (табл. 8.10).

Таблица 8.10

Исходные данные

Поставщики	Потребители					Запасы, т
	B_1	B_2	B_3	B_4	B_5^*	
A_1	40	20	80	50	0	10
A_2	30	30	60	40	0	8
A_3	50	30	40	40	0	5
Потребность, т	4,8	4,8	6	5	2,4	23

Используя метод потенциалов, получим оптимальное решение (табл. 8.11).

Таблица 8.11

Оптимальное решение

Поставщики	Потребители					Запасы, т
	B_1	B_2	B_3	B_4	B_5^*	
A_1	40	20 4,8	80	50 2,8	0 2,4	10
A_2	30 4,8	30	60 1	40 2,2	0	8
A_3	50	30	40 5	40	0	5
Потребность, т	4,8	4,8	6	5	2,4	23

Анализ решения.

Бумаги 1-го сорта в количестве 4,8 т затрачено на издание второй книги; 2,8 т – на издание четвертой книги; 2,4 т – не использовано. Бумаги 2-го сорта затрачено: на первую книгу – 4,8 т; на издание третьей книги – 1,0 т; на издание четвертой книги – 2,2 т; бумага 3-го сорта использована на издание третьей книги в количестве 5 т.

8.4. Метод Фогеля

Как известно, метод потенциалов позволяет за конечное число шагов найти оптимальный план, следовательно, желательно, чтобы первый опорный план был ближе к оптимальному. Способ получения опорного плана, предложенный американским ученым У. Фогелем, позволяет найти практически оптимальный план. Найденный план или совпадение с оптимальным, или незначительно от него отличается.

Способ Фогеля рассмотрим в табл. 8.12. Процесс начинается с определения разностей между двумя наименьшими элементами каждой строки и каждого столбца таблицы. Так, в столбце B_5 минимальный элемент равен 3, следующий за ним по величине элемент – 8, разность между ними равна 5. Эта и другие разности по строкам и столбцам записаны в таблицу. Затем из всех разностей выбирается наибольшая. В приведенном примере это $\Delta = 6$ в строке A_6. Минимальный элемент в соответствующей строке равен 5 и совпадает с клеткой (A_6B_4), сюда же записывается поставка (см. табл. 8.12).

Все запасы A_6 исчерпаны, поэтому эту строку в дальнейшем не рассматриваем (ее можно вычеркнуть).

Смысл способа Фогеля легко понять. Найденные разности показывают, насколько больше будут затраты, если в соответствующем столбце (или строке) поставка будет записана не в клетку, где находится минимальный в этом столбце (строке) элемент, а в клетку, где находится элемент, следующий за ним по величине.

Теперь, когда из рассмотрения исключены элементы столбца B_1, изменяются разности по строкам, но по столбцам они не изменяются (см. 4-ю строку и столбец). В столбце B_2 осталась разность $\Delta = 6$, она и является наибольшей.

Находим в столбце B_2 минимальный элемент, равный 3, и в клетку A_1B_2 делаем поставку. Вновь находим разности (см. 5-ю строку и столбец).

Опять имеем две максимальные разности в строке A_3 и в столбце B_3 ($\Delta = 4$). В строке A_3 минимальный элемент 13 в A_3B_3, но он

Таблица 8.12

Пример нахождения опорного плана методом Фогеля

	B_1	B_2	B_3	B_4	B_5	B_6	B_7	Запасы	Разности по строчкам 1	2	3	4	5
A_1	1 17	3 12	5	7	9	11	4 1	30	2	2	2	1	1
A_2	14	12	10	10	8 26	6	3 3	29	3	3	3	3	3
A_3	7	9	13 19	17 8	21	24 1	29	28	2	2	2	4	4
A_4	35	32	29	26	23	20 27	20	27	0	0	0	0	0
A_5	7	10	13	16	12	8 11	5 15	26	2	2	2	3	3
A_6	25	20	15	5 25	11	17	23	25	6	–	–	–	–
A_7	10	8	9	16	12	17	10 24	24	1	1	1	1	1
A_8	2	5	8	8	3 23	15	21	23	1	1	–	–	–
Потребности	17	12	19	33	49	39	43	212					
Разности по столбцам 1	1	2	3	2	5	2	1						
2	1	2	3	1	5	2	1						
3	6	6	4	3	1	2	1						
4	–	6	4	3	1	2	1						
5	–	–	4	3	1	2	1						

не является минимальным в своем столбце. Точно также не является минимальным в своей строке минимальный элемент столбца B_3.

Если учитывать, что мы ищем не абсолютный оптимум, а лишь один из базисов, можно произвольно выбрать или строку A_3, или столбец B_3. Но можно вычислить в соответствующих столбце и строке вторые разности – это разности между минимальными элементами и элементами, не ближайшими к ним по величине, а следующими за ними.

Вторая разность для столбца B_3 равна (9 – 5 = 4), а для строки A_3 (21 – 13 = 8), поэтому поставку надо сделать в клетку A_3B_3.

Проводя аналогичные рассуждения, можно получить допустимый план (см. табл.), общие затраты на перевозку составляют $Z = 1818$.

8.5. Транспортная задача в сетевой постановке

Если условия транспортной задачи заданы в виде схемы, на которой условно изображены поставщики, потребители и связывающие их дороги, указаны величины запасов груза и потребности в нем, а также числа C_{ij}, являющиеся показателями принятого в задаче критерия оптимальности (тарифы, расстояния и т. п.), то говорят, что транспортная задача поставлена в сетевой форме (рис. 8.1).

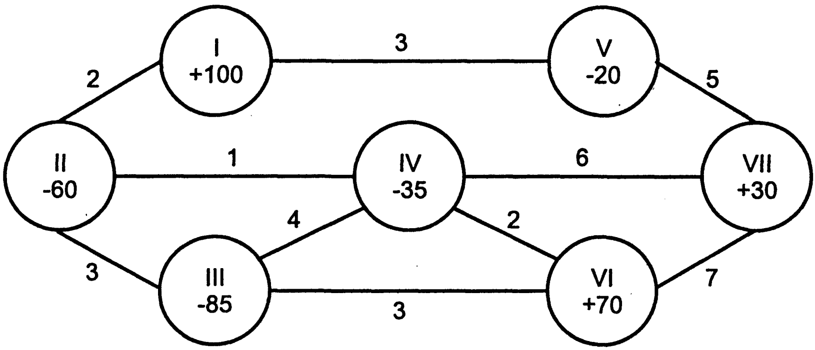


Рис. 8.1. Транспортная задача в сетевой форме

Пункты расположения поставщиков и потребителей будем изображать кружками и называть вершинами (узлами) сети, запасы груза будем записывать в кружках положительными, а потребности – отрицательными числами. Дороги, связывающие пункты расположения и потребления груза, будем изображать линиями и называть ребрами (дугами, звеньями) сети. На сети могут быть изображены вершины, в которых нет ни поставщиков, ни потребителей. Наличие таких вершин не повлияет на способ решения, если считать, что запасы (потребности) груза в них равны нулю.

Различия между транспортными задачами в матричной и сетевой формах весьма незначительны, так как методы их решения основаны на одних и тех же идеях (метод потенциалов).

Решение задачи на сети начинается с построения начального опорного плана. Последовательность решения задачи рассмотрена на конкретном примере (см. рис. 8.1). Поставку груза из вершины в вершину будем обозначать стрелками с указанием величины поставок.

Опорный план должен удовлетворять следующим требованиям:

1) все запасы должны быть распределены, а потребности удовлетворены;

2) к каждой вершине должна подходить или выходить из нее хотя бы одна стрелка;

3) общее количество стрелок должно быть на единицу меньше числа вершин ($n + m - 1$);

4) стрелки не должны образовывать замкнутый контур.

План распределения груза, отвечающий этим требованиям, представлен на рис. 8.2.

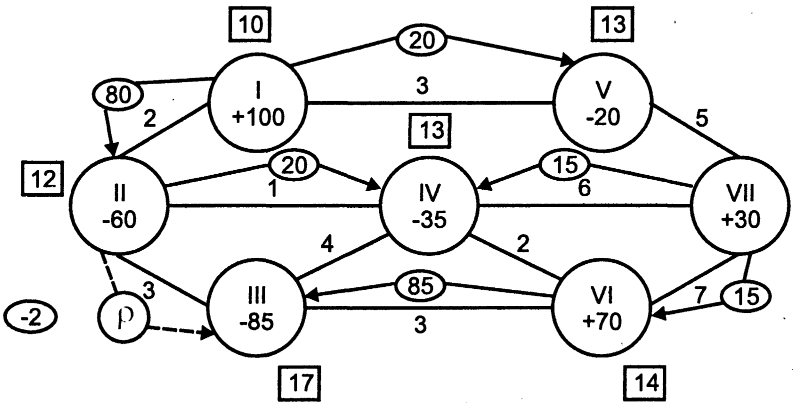


Рис. 8.2. Пример опорного план распределения груза при решении задачи на сети

Далее следует проверить план на оптимальность. Для этого вычисляем потенциалы. Одной из вершин (например, вершине 1) присвоим некоторое значение потенциала (например, равное 10). Для наглядности потенциалы будем заключать в рамки. После этого, двигаясь по стрелкам, определяем потенциалы остальных вершин, руководствуясь правилом: если стрелка выходит из вершины, то к потенциалу этой вершины прибавляем показатель C_{ij} критерия оптимальности, если же направление стрелки противоположно, C_{ij} вычитаем (см. рис. 8.2).

После вычисления потенциалов находят характеристики ребер без стрелок по правилу: из большего потенциала вычитается меньший, а разность вычитается из показателя C_{ij}, отвечающего данному ребру.

Если все ребра без стрелок имеют неотрицательные характеристики, то составленный план является оптимальным.

Вычислим характеристики ребер без стрелок

$$S_{5,7} = 5 - (13 - 7) = -1;\ S_{4,6} = 1;\ S_{4,3} = 0;\ S_{2,3} = -2.$$

Два ребра имеют отрицательные характеристики, в этом случае выбирается ребро с наименьшей отрицательной характеристикой и к нему подрисовывается новая стрелка, при этом образуется замкнутый контур из стрелок. Новая стрелка направляется от вершины с меньшим потенциалом к вершине с большим потенциалом.

В нашем примере новая стрелка направлена от вершины II к вершине III (штриховая линия (см. рис. 8.2).

Для определения величины поставки (ρ) для ребра S_{23} рассматриваются все стрелки образовавшегося замкнутого контура, имеющие направление, противоположное новой стрелке (участок S_{23}), и среди них находится стрелка с наименьшей поставкой λ (в нашем примере λ = 15 на ребре $S_{6,7}$). Выбранная таким образом величина прибавляется ко всем поставкам в стрелках, имеющих то же направление, что и новая стрелка, и вычитается из поставок в стрелках, имеющих противоположное направление. Поставки в стрелках, не входящих в контур, сохраняются неизменными. Стрелка, на

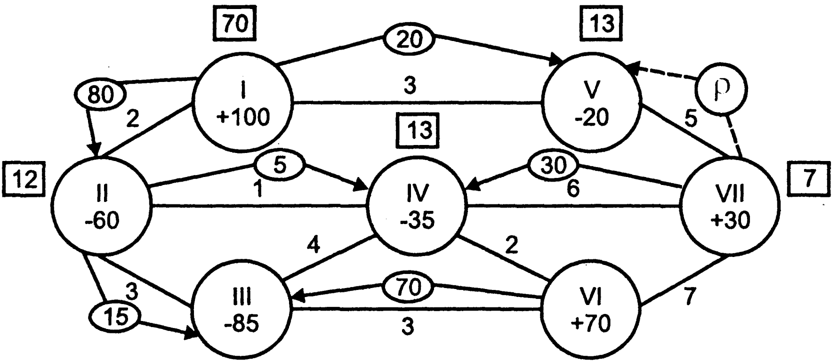


Рис. 8.3. Новый опорный план распределения груза

которой выбрано число λ, ликвидируется и общее число стрелок остается прежним. Новый опорный план представлен на рис. 8.3.

Полученный план исследуется на оптимальность, подобно предыдущему. Сделав еще шаг, получим оптимальный план (рис. 8.4), когда все характеристики S_{ij} на участках без стрелок неотрицательны.

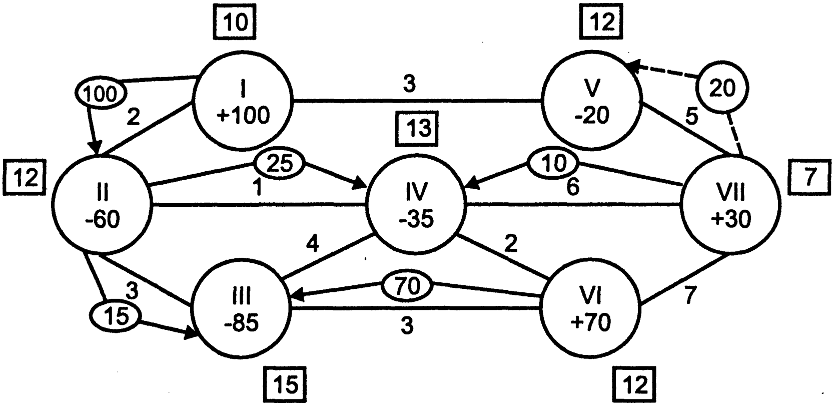


Рис. 8.4. Оптимальный план распределения груза

Определим значение целевой функции:

$$Z_{\min} = 100 \cdot 2 + 25 \cdot 1 + 15 \cdot 3 + 70 \cdot 3 + 10 \cdot 6 + 20 \cdot 5 = 640.$$

Вырождение плана транспортной задачи в сетевой постановке проявляется в том, что при полном использовании запасов и полном удовлетворении потребностей количество стрелок оказывается меньше чем $m + n - 1$, где n – поставщики, а m – потребители.

Для преодоления вырождения вводится нужное количество стрелок с нулевыми поставками, направление стрелок выбирается произвольно, однако они не должны образовывать замкнутый контур.

В случае открытой модели вводят фиктивного потребителя (поставщика) со спросом, равным небалансу. Фиктивный потребитель (поставщик) соединяется дугами непосредственно со всеми поставщиками (потребителями), при этом показатели C_{ij} ребер, соединяющих фиктивного потребителя (поставщика) с реальными поставщиками (потребителями), следует брать одинаковыми и сравнительно большими. Это делается для того, чтобы исключить возможность использования фиктивной вершины в качестве промежуточного пункта.

8.6. Доставка груза в кратчайший срок

В практической деятельности могут возникнуть ситуации, когда нас в первую очередь интересуют не затраты на перевозку груза (их минимизация), а время доставки этих грузов потребителям. Например, при подготовке крупных военных операций, когда необходимо в кратчайший срок сосредоточить ресурсы в намеченных пунктах или при стихийных бедствиях (землетрясение, ураганы и т. п.), возникает задача обеспечения пострадавших районов различными ресурсами в кратчайший срок.

Для решения подобных задач рассмотренный ранее метод потенциалов непригоден. Эти задачи решаются с помощью *специального алгоритма*.

1. Любым способом строим один из опорных планов.

2. Определяем наибольший элемент t' из всех t_{ij}, соответствующих занятым клеткам, и все клетки с элементами $t_{ij} \geq t'$ (это могут быть лишь свободные клетки) вычеркиваются.

3. Начиная с клетки с наибольшим временем доставки t', строим разгрузочный цикл так, чтобы клетки с нечетными номерами (считая первой разгружаемую клетку с элементом t') были занятыми. Одна из вершин разгрузочного цикла будет свободной. В общем случае построение разгрузочного цикла неоднозначно.

4. Сделав в свободную вершину цикла поставку ρ, проводим компенсации по вершинам цикла, определяем величину ρ (так же, как в методе потенциалов), строим новый план.

5. Переходим ко второму пункту алгоритма, естественно, не учитывая ранее вычеркнутые клетки.

6. Алгоритмом пользуемся до тех пор, пока построение разгрузочного цикла становится невозможным.

7. Последний полученный план является оптимальным, наибольшее время, соответствующее занятой клетке в этом плане, определяет наименьшее время по доставке грузов всем потребителям.

Пример 8.4.

Определить оптимальный план перевозок из условия доставки груза в кратчайший срок. Известны ресурсы a_i: 30; 35; 40, потребности $в_k$: 20; 34; 16; 10; 25 и матрица

$$T = \left\| t_{ik} \right\| = \begin{pmatrix} 2 & 6 & 3 & 4 & 8 \\ 1 & 5 & 6 & 9 & 7 \\ 3 & 4 & 1 & 6 & 10 \end{pmatrix},$$

где t_{ik} – время, затрачиваемое на перевозку груза из i-го пункта отправления в k-й пункт назначения.

Решение

Запишем в виде таблицы исходную информацию и внесем в нее один из опорных планов (пункт 1 алгоритма).

Таблица 8.13

Исходная информация и опорный план

Постав-щики	Потребители					Запасы, т
	B_1	B_2	B_3	B_4	B_5	
A_1	2 20	6 10	3	4	8	30
A_2	1	5 24	6 $11 - \rho$	9	7 ρ	35
A_3	3	4	1 $5 + \rho$	6 10	10 $25 - \rho$	40
Потреб-ность, т	20	34	16	10	25	105

Согласно пункту 2 алгоритма $t' = t_{35} = 10$ строим цикл перерасчета (3-й пункт). В свободную вершину цикла (2,5) делаем поставку ρ и проводим компенсации по вершинам цикла, затем определяем величину ρ:

$$\rho = 11.$$

Строим новый план и переходим к пункту 2 алгоритма (табл. 8.14).

Таблица 8.14

Промежуточный план распределения

Постав-щики	Потребители					Запасы, т
	B_1	B_2	B_3	B_4	B_5	
A_1	2 $20 - \rho$	6 $10 + \rho$	3	4	8	30
A_2	1	5 $24 - \rho$	6	9	7 $11 + \rho$	35
A_3	3 ρ	4	1 16	6 10	1 $14 - \rho$ 0	40
Потреб-ность, т	20	34	16	10	25	105

В новой таблице $t' = t_{35} = 10$ (клетка 3,5 оказалась не полностью разгруженной). Из анализа нового цикла пересчета заключаем: $\rho = 14$; строим новый план (табл. 8.15).

Таблица 8.15

Оптимальный план

Поставщики	Потребители					Запасы, т
	B_1	B_2	B_3	B_4	B_5	
A_1	2 6	6 24	3	4	$\cancel{8}$	30
A_2	1	5 10	6	$\cancel{9}$	7 25	35
A_3	3 14	4	1 16	6 10	$\cancel{10}$	40
Потребность, т	20	34	16	10	25	105

$t' = t_{25} \geq 7$, вычеркиваем все свободные клетки, для которых $t_{ij} \geq 7$. Для дальнейшего улучшения плана необходимо разгрузить клетку (2,5), но построение разгрузочного цикла для этой клетки невозможно. На основании этого заключаем, что решение, полученное в последней таблице, является оптимальным, обеспечивающим перевозки за время $t_{\min} = 7$.

Задачи

8.1. В пунктах A и B находятся соответственно 150 и 90 т горючего. Пунктам 1, 2, 3 требуются соответственно 60, 70, 110 т горючего. Стоимость перевозки 1 т горючего из пункта A в пункты 1, 2, 3 равна 60, 10, 40 тыс. руб. за 1 т соответственно, а из пункта B в пункты 1, 2, 3 – 120, 20, 80 тыс. руб. за 1 т соответственно.

Составьте план перевозок горючего, минимизирующий общую сумму транспортных расходов.

8.2. Три завода выпускают грузовые автомобили, которые отправляются четырем потребителям. Первый завод поставляет 90 платформ грузовиков, второй – 30 платформ, третий – 40 платформ. *Требуется поставить* платформы следующим потребителям: первому – 70 шт., второму – 30, третьему – 20, четвертому – 40 шт. Стоимость перевозки одной платформы от поставщика до потребителя указана в следующей таблице (д. е.):

Поставщики	Потребители			
	1	2	3	4
I	18	20	14	10
II	10	20	40	30
III	16	22	10	20

Составьте оптимальный план доставки грузовых автомобилей.

8.3. Строительство магистральной дороги включает задачу заполнения имеющихся на трассе выбоин до уровня основной дороги и срезания в некоторых местах дороги выступов. Срезанным грунтом заполняются выбоины. Перевозка грунта осуществляется грузовиками одинаковой грузоподъемности. Расстояние в километрах от срезов до выбоин и объем работ указаны в следующей таблице:

Поставщики	Потребители			Наличие грунта, т
	I	II	III	
A	1	2	3	110
B	2	1	3	130
C	1	2	4	20
Требуемое количество грунта, т	100	140	60	

Составьте план перевозок, минимизирующий общий пробег грузовиков.

8.4. Груз, хранящийся на трех складах и требующий для перевозки 60, 80, 106 автомашин соответственно, необходимо перевезти в четыре магазина. Первому магазину требуется 44 машины груза, второму – 70, третьему – 50 и четвертому – 82 машины. Стоимость пробега одной автомашины за 1 км составляет 10 д. е. Расстояния от складов до магазинов указаны в следующей таблице:

Склады	Магазины			
	1	2	3	4
1	13	17	6	8
2	2	7	10	41
3	12	18	2	22

Составьте оптимальный по стоимости план перевозки груза от складов до магазинов.

8.5. На складах А, В, С находится сортовое зерно 100, 150, 250 т, которое нужно доставить в четыре пункта. Пункту 1 необходимо поставить 50 т, пункту 2 – 100, пункту 3 – 200, пункту 4 – 150 т сортового зерна. Стоимость доставки 1 т зерна со склада А в указанные пункты соответственно равна (д. е.) 80, 30, 50, 20; со склада В – 40, 10, 60, 70; со склада С – 10, 90, 40, 30.

Составьте оптимальный план перевозки зерна из условия минимума стоимости перевозки.

8.6. Завод имеет три цеха – А, В, С и четыре склада – 1; 2; 3; 4. Цех А производит 30 тыс. шт. изделий, цех В – 40; цех С – 20 тыс. шт. изделий. Пропускная способность складов за то же время характеризуется следующими показателями: склад 1 – 20 тыс. шт. изделий; склад 2 – 30; склад 3 – 30 и склад 4 – 10 тыс. шт. изделий. Стоимость перевозки 1 тыс. шт. изделий из цеха А на склады 1, 2, 3, 4 – соответственно (д. е.): 20, 30, 40, 40, из цеха В – соответственно 30, 20, 50, 10, а из цеха С – соответственно 40, 30, 20, 60.

Составьте такой план перевозки изделий, при котором расходы на перевозку 90 тыс. шт. изделий были бы наименьшими.

8.7. На строительном полигоне имеется пять кирпичных заводов, объем производства которых в сутки равен 600; 600; 500; 650; 700 т. Заводы удовлетворяют потребности семи строительных объектов соответственно в количестве 350; 450; 300; 450; 300; 200; 450 т. Оставшийся кирпич отправляют по железной дороге в другие районы. Кирпич на строительные объекты доставляется автомобильным транспортом. Расстояние в километрах от заводов до объектов указано в следующей таблице:

Заводы	Объекты						
	B_1	B_2	B_3	B_4	B_5	B_6	B_7
A_1	14	5	10	8	16	10	25
A_2	13	4	11	9	20	12	23
A_3	18	8	14	18	23	13	21
A_4	14	7	13	19	15	16	23
A_5	11	15	14	25	19	15	20

Определите, с каких заводов и на какие объекты должен доставляться кирпич, а также какие заводы и в каком количестве должны отправлять кирпич в другие районы, чтобы транспортные издержки по доставке кирпича автотранспортом были минимальными. Стоимость перевозки 1 т кирпича автотранспортом удовлетворяет условию $c = a + d(\ell - 1)$, где $a = 25$ д. е., $d = 5$ д. е., ℓ – пробег, км.

8.8. Имеются две станции технического обслуживания (СТО), выполняющие ремонтные работы для трех автопредприятий. Производственные мощности СТО, стоимость ремонта в различных СТО, затраты на транспортировку от автопредприятий на СТО и обратно и прогнозируемое количество ремонтов в планируемом периоде на каждом автопредприятии приведены в следующей таблице:

СТО	Стоимость ремонта ед., д. е.	Затраты на транспортировку, тыс. руб.			Производственная мощность, шт.
		АТП-1	АТП-2	АТП-3	
1	520	60	70	20	10
2	710	40	50	30	8
Потребное количество, д. е.		6	7	5	18

Требуется определить, какое количество автомашин из каждого автопредприятия необходимо отремонтировать на каждой СТО, чтобы суммарные расходы на ремонт и транспортировку были минимальными.

8.9. Найдите оптимальный план распределения заявок на ремонт для условий, приведенных в следующей таблице:

СТО	Затраты на ТО и ремонт одного автомобиля, д.е.	Затраты на транспортировку, тыс. руб.				Производственная мощность, шт.
		АТП-1	АТП-2	АТП-3	АТП-4	
1	720	20	40	30	10	80
2	650	30	20	25	45	20
3	690	35	50	20	30	40
Прогнозируемое количество ТО, ед.		30	10	40	20	

8.10. Имеются два хранилища с однородным продуктом, в которых сосредоточено 200 и 120 т продукта соответственно. Продукты необходимо перевезти трем потребителям соответственно в количестве 80, 100 и 120 т. Расстояния от хранилищ до потребителей (в км) следующие:

Хранилище	Потребители		
	1	2	3
1	20	30	50
2	60	20	40

Затраты на перевозку 1 т продукта на 1 км постоянны и равны 5 д. е.

Определите план перевозок продукта от хранилищ до потребителей из условия минимизации транспортных расходов.

8.11. Промышленный концерн имеет два завода и пять складов в различных регионах страны. Каждый месяц первый завод производит 40, а второй – 70 ед. продукции. Вся продукция, производимая заводами, должна быть направлена на склады. Вместимость первого склада равна 20 ед. продукции; второго – 30; третьего – 15; четвертого – 27; пятого – 28 ед. Издержки транспортировки продукции от завода до склада следующие (ед.):

Заводы	Склады				
	1	2	3	4	5
1	520	480	650	500	720
2	450	525	630	560	750

Распределите план перевозок из условия минимизации ежемесячных расходов на транспортировку.

8.12. Три нефтеперерабатывающих завода с суточной производительностью 10; 8 и 6 млн галлонов бензина снабжают три бензохранилища, спрос которых составляет 6; 11 и 7 млн галлонов. Бензин транспортируется в бензохранилища по трубопроводу. Стоимость перекачки бензина на 1 км составляет 5 д. е. на 100 галлонов. Завод 1 не связан с хранилищем 3. Расстояние от заводов до бензохранилищ следующее (км):

Номер завода	Бензохранилища		
	1	2	3
1	100	150	–
2	420	180	60
3	200	280	120

Сформулируйте соответствующую транспортную задачу и решите на минимум транспортных затрат.

8.13. Пусть в задаче 8.12 производительность нефтеперерабатывающего завода 1 снизилась до 8 млн галлонов. Кроме того, обязательно полное удовлетворение спроса бензохранилища 2. Недопоставки в хранилища 1 и 3 штрафуются на сумму 8 д. е. за каждый галлон.

Сформулируйте соответствующую транспортную задачу и решите на минимум издержек.

8.14. Автомобили перевозятся на трайлерах из трех центров распределения пяти продавцам. Стоимость перевозки в расчете на 1 км пути, пройденного трайлером, равна 60 д. е. Один трайлер может перевозить до 15 автомобилей. Стоимость перевозок не зависит от того, насколько полно загружается трайлер. В приведенной ниже таблице указаны расстояния (км) между центрами распределения и продавцами, а также величины, характеризующие ежемесячный спрос и объемы поставок, исчисляемые количеством автомобилей:

Центр распределения	Продавцы					Объем поставок, шт.
	1	2	3	4	5	
1	80	120	180	150	50	300
2	60	70	50	65	90	350
3	30	80	120	140	90	120
Спрос на автомобили, шт.	110	250	140	150	120	770

Определите минимальные затраты на доставку автомобилей.

8.15. Решите задачу распределения станков четырех различных типов по шести типам работ. Пусть имеются 30; 45; 25 и 20 станков соответствующих типов. Шесть типов работ характеризуются 30; 20; 10; 40; 10 и 10 операциями соответственно. На станке 3 не может выполняться работа 6. Исходя из коэффициентов стоимости операции, представленных в следующей таблице, *постройте* модель и *выполните* оптимальное распределение станков по работам:

Тип станков	Тип работ					
	1	2	3	4	5	6
1	10	1	3	7	14	8
2	4	8	12	2	10	7
3	12	3	14	6	2	–
4	11	12	9	5	1	3

8.16. В данной транспортной задаче суммарный спрос превосходит суммарный объем производства. Пусть штрафы за недопоставку единицы продукции в пункты назначения 1, 2 и 3 равны соответственно 5, 3 и 2.

Исходные данные следующие:

Заводы	Потребители			Объем производства, шт.
	1	2	3	
A_1	3	2	4	50
A_2	5	4	5	75
A_3	1	6	7	30
Потребность, шт.	60	40	70	

Найдите оптимальное решение.

8.17. Пусть в задаче 8.16 не введены штрафы, а спрос пункта назначения 1 должен быть полностью удовлетворен.

Сформулируйте новую задачу и *найдите* оптимальное решение.

8.18. В таблице представлена несбалансированная транспортная задача, в которой назначается плата за хранение каждой единицы невывезенного из исходного пункта i груза. Пусть коэффициенты стоимости хранения груза в исходных пунктах 1; 2 и 3 соответственно равны 5; 6 и 2.

Пункты хранения (склады)	Потребители			Запасы продукции, т
	1	2	3	
1	1	0	4	300
2	3	1	2	400
3	1	2	1	250
Спрос, т	280	320	200	

Найдите оптимальное решение, если весь объем груза исходного пункта 2 должен быть вывезен для того, чтобы освободить место для новой продукции.

Для задач 8.19 – 8.38 дано следующее условие.

Имеются три пункта поставки однородного груза – A_1; A_2; A_3 и пять пунктов потребления этого груза – B_1; B_2; B_3; B_4; B_5. В пунктах A_1; A_2; A_3 находится груз a_1; a_2; a_3 соответственно. Груз необходимо доставить в пункты B_1; B_2; B_3; B_4; B_5 в количестве b_1; b_2; b_3;

b_4; b_5 соответственно. Расстояния между пунктами заданы следующей матрицей (км):

$$D=\begin{pmatrix} d_{11} & d_{12} & \dots & d_{15} \\ d_{21} & d_{22} & \dots & d_{25} \\ d_{31} & d_{32} & \dots & d_{35} \end{pmatrix}.$$

Требуется найти оптимальный план закрепления потребителей за поставщиками однородного груза при условии минимизации общего пробега автомобилей, используя параметры, представленные ниже.

8.19. $A^T = (a_1; a_2; a_3) = (200; 175; 225)$;
$B^T = (b_1; b_2; b_3; b_4; b_5) = (100; 130; 80; 190; 100)$;

$$D=\begin{pmatrix} 5 & 7 & 4 & 2 & 5 \\ 7 & 1 & 3 & 1 & 10 \\ 2 & 3 & 6 & 8 & 7 \end{pmatrix}.$$

8.20. $A^T = (a_1; a_2; a_3) = (200; 450; 250)$;
$B^T = (b_1; b_2; b_3; b_4; b_5) = (100; 125; 325; 250; 100)$;

$$D=\begin{pmatrix} 5 & 8 & 7 & 10 & 3 \\ 4 & 2 & 2 & 5 & 6 \\ 7 & 3 & 5 & 9 & 2 \end{pmatrix}.$$

8.21. $A^T = (a_1; a_2; a_3) = (250; 200; 200)$;
$B^T = (b_1; b_2; b_3; b_4; b_5) = (120; 130; 100; 160; 110)$;

$$D=\begin{pmatrix} 27 & 36 & 35 & 31 & 29 \\ 22 & 23 & 26 & 32 & 35 \\ 35 & 42 & 38 & 32 & 39 \end{pmatrix}.$$

8.22. $A^T = (a_1; a_2; a_3) = (350; 330; 270)$;
$B^T = (b_1; b_2; b_3; b_4; b_5) = (210; 170; 220; 150; 200)$;

$$D=\begin{pmatrix} 3 & 12 & 9 & 1 & 7 \\ 2 & 4 & 11 & 2 & 10 \\ 7 & 14 & 12 & 5 & 8 \end{pmatrix}.$$

8.23. $A^T = (a_1; a_2; a_3) = (300; 250; 200);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (210; 170; 220; 150; 200);$

$$D = \begin{pmatrix} 4 & 8 & 13 & 2 & 7 \\ 9 & 4 & 11 & 9 & 17 \\ 3 & 16 & 10 & 1 & 4 \end{pmatrix}.$$

8.24. $A^T = (a_1; a_2; a_3) = (350; 200; 300);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (170; 140; 200; 195; 145);$

$$D = \begin{pmatrix} 22 & 14 & 16 & 28 & 30 \\ 19 & 17 & 26 & 36 & 36 \\ 37 & 30 & 31 & 39 & 41 \end{pmatrix}.$$

8.25. $A^T = (a_1; a_2; a_3) = (200; 250; 200);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (190; 100; 120; 110; 130);$

$$D = \begin{pmatrix} 28 & 27 & 18 & 27 & 24 \\ 18 & 26 & 27 & 32 & 21 \\ 27 & 33 & 23 & 31 & 34 \end{pmatrix}.$$

8.26. $A^T = (a_1; a_2; a_3) = (230; 250; 170);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (140; 90; 160; 110; 150);$

$$D = \begin{pmatrix} 40 & 19 & 25 & 26 & 35 \\ 42 & 25 & 27 & 15 & 38 \\ 46 & 27 & 36 & 40 & 45 \end{pmatrix}.$$

8.27. $A^T = (a_1; a_2; a_3) = (200; 300; 250);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (210; 150; 120; 135; 135);$

$$D = \begin{pmatrix} 20 & 10 & 12 & 13 & 16 \\ 25 & 19 & 20 & 14 & 10 \\ 17 & 18 & 15 & 10 & 17 \end{pmatrix}.$$

8.28. $A^T = (a_1; a_2; a_3) = (200; 350; 300);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (270; 130; 190; 150; 110);$

$$D = \begin{pmatrix} 24 & 50 & 45 & 27 & 15 \\ 20 & 32 & 40 & 35 & 30 \\ 22 & 16 & 18 & 28 & 20 \end{pmatrix}.$$

8.29. $A^T = (a_1; a_2; a_3) = (150; 150; 200);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (100; 70; 130; 110; 90);$

$$D = \begin{pmatrix} 15 & 3 & 6 & 10 & 30 \\ 12 & 8 & 12 & 16 & 25 \\ 14 & 11 & 9 & 8 & 15 \end{pmatrix}.$$

8.30. $A^T = (a_1; a_2; a_3) = (330; 270; 350);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (220; 170; 210; 150; 200);$

$$D = \begin{pmatrix} 10 & 12 & 11 & 20 & 40 \\ 14 & 8 & 9 & 11 & 15 \\ 8 & 6 & 12 & 14 & 20 \end{pmatrix}.$$

8.31. $A^T = (a_1; a_2; a_3) = (150; 200; 100);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (90; 150; 75; 60; 75);$

$$D = \begin{pmatrix} 15 & 23 & 26 & 19 & 18 \\ 17 & 13 & 14 & 25 & 10 \\ 12 & 21 & 24 & 12 & 9 \end{pmatrix}.$$

8.32. $A^T = (a_1; a_2; a_3) = (300; 350; 200);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (145; 195; 200; 140; 170);$

$$D = \begin{pmatrix} 18 & 30 & 35 & 25 & 40 \\ 12 & 14 & 22 & 20 & 35 \\ 10 & 28 & 23 & 19 & 30 \end{pmatrix}.$$

8.33. $A^T = (a_1; a_2; a_3) = (300; 300; 250);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (150; 140; 115; 225; 220);$

$$D = \begin{pmatrix} 18 & 20 & 23 & 15 & 24 \\ 25 & 15 & 16 & 19 & 29 \\ 6 & 11 & 10 & 8 & 9 \end{pmatrix}.$$

8.34. $A^T = (a_1; a_2; a_3) = (300; 230; 320);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (190; 150; 130; 180; 200);$

$$D = \begin{pmatrix} 25 & 20 & 22 & 31 & 32 \\ 11 & 18 & 20 & 15 & 16 \\ 10 & 9 & 16 & 20 & 25 \end{pmatrix}.$$

8.35. $A^T = (a_1; a_2; a_3) = (300; 250; 300);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (130; 130; 150; 190; 250);$

$$D = \begin{pmatrix} 17 & 21 & 24 & 32 & 24 \\ 23 & 10 & 15 & 20 & 26 \\ 20 & 25 & 22 & 24 & 25 \end{pmatrix}.$$

8.36. $A^T = (a_1; a_2; a_3) = (200; 300; 250);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (120; 140; 160; 180; 150);$

$$D = \begin{pmatrix} 16 & 21 & 24 & 22 & 20 \\ 25 & 30 & 35 & 20 & 27 \\ 34 & 26 & 25 & 28 & 21 \end{pmatrix}.$$

8.37. $A^T = (a_1; a_2; a_3) = (270; 450; 330);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (190; 210; 200; 230; 220);$

$$D = \begin{pmatrix} 37 & 30 & 15 & 20 & 35 \\ 16 & 20 & 12 & 17 & 21 \\ 10 & 26 & 20 & 25 & 29 \end{pmatrix}.$$

8.38. $A^T = (a_1; a_2; a_3) = (210; 450; 290);$
$B^T = (b_1; b_2; b_3; b_4; b_5) = (200; 220; 170; 210; 150);$

$$D = \begin{pmatrix} 19 & 25 & 30 & 32 & 20 \\ 40 & 21 & 12 & 21 & 41 \\ 15 & 14 & 28 & 27 & 22 \end{pmatrix}.$$

8.39. Три завода производят продукцию в объемах: 200, 300, 500 ед. Эта продукция необходима четырем потребителям в количестве: 210, 320, 150 и 200 ед. Готовая продукция поступает потребителям через склады D_1 и D_2, емкость которых соответственно равна 500 и 300 ед. Транспортные расходы на доставку единицы продукции на склады и со складов потребителям заданы матрицами C_1 и C_2.

$$C_1 = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ 2 & 5 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 3 & 1 & 5 & 2 \\ 4 & 3 & 6 & 1 \end{pmatrix}.$$

Необходимо составить план доставки продукции от заводов к потребителям с учетом наименьших затрат.

Найти методом Фогеля план перевозок в задачах **8.40–8.43**:

8.40

Поставщики	Потребители				Запас
	B_1	B_2	B_3	B_4	
A_1	4	7	3	9	115
A_2	2	1	8	5	70
A_3	7	9	6	1	68
Потребности	95	38	50	70	

8.41

Поставщики	Потребители							Запас
	B_1	B_2	B_3	B_4	B_5	B_6	B_7	
A_1	10	8	2	5	4	8	2	50
A_2	8	7	1	5	2	4	9	45
A_3	9	11	13	7	1	3	8	45
A_4	10	10	2	6	4	5	9	10
A_5	9	8	2	4	3	4	8	5
Потребности	30	20	15	30	20	30	20	

8.42

Поставщики	Потребители								Запас
	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	
A_1	21	19	17	18	15	16	27	18	650
A_2	16	14	7	20	18	19	15	20	600
A_3	15	13	11	18	19	22	23	14	200
A_4	14	12	12	17	21	23	14	14	100
A_5	10	11	10	20	16	21	12	12	200
Потребности	200	300	400	250	150	100	150	300	

8.43

Поставщики	Потребители							Запас
	B_1	B_2	B_3	B_4	B_5	B_6	B_7	
A_1	19	24	26	28	30	22	18	450
A_2	31	29	27	25	21	21	19	450
A_3	25	15	17	22	24	18	13	300
A_4	28	25	21	20	23	27	29	400
A_5	13	20	27	18	14	30	24	350
Потребности	100	350	50	150	200	300	350	

Необходимо решить транспортные задачи **8.44÷8.47**, заданные в сетевой форме (рис. 8.5–8.8):

8.44.

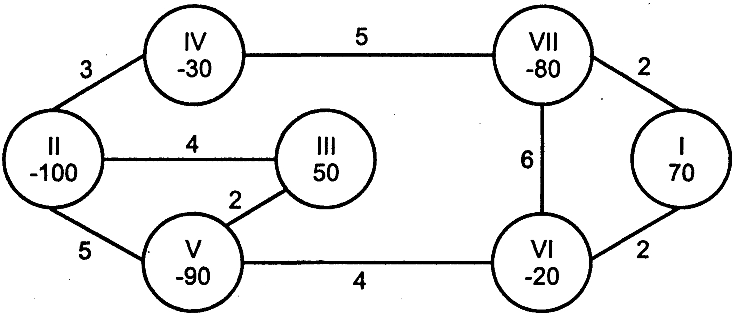


Рис. 8.5.

8.45.

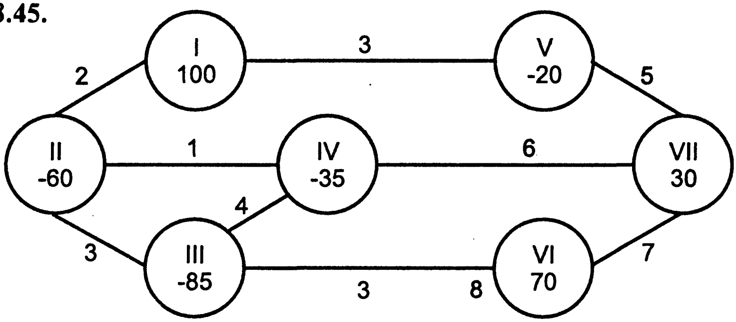


Рис. 8.6.

8.46.

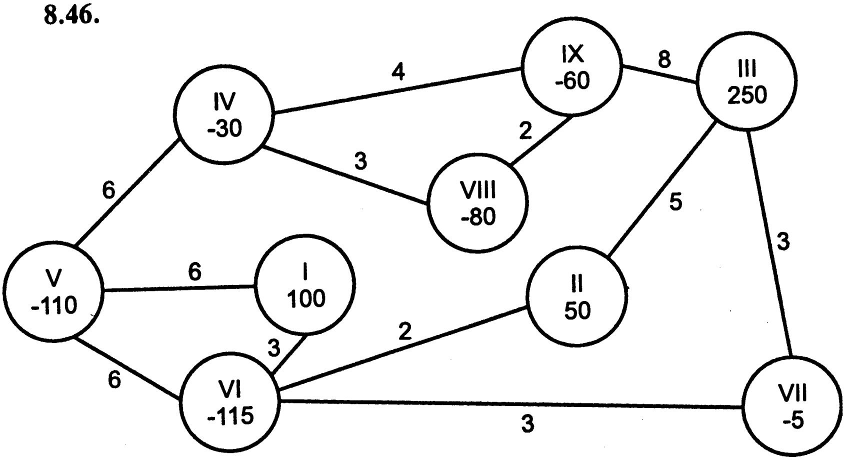


Рис. 8.7.

8.47.

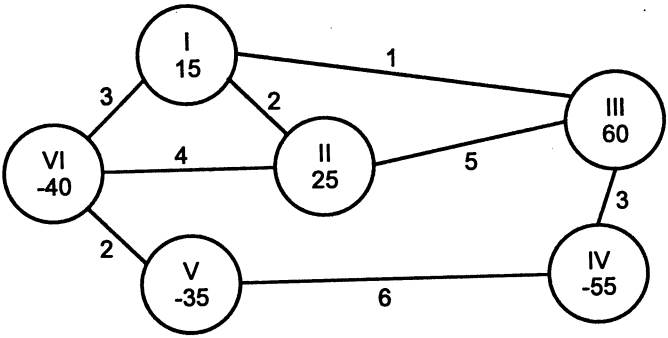


Рис. 8.8.

Глава 9
Теория игр и принятия решений

9.1. Основные понятия

Рассмотренные задачи линейного программирования формулировались и решались в предположении наличия *полной информации.* Их можно отнести к совокупности задач принятия решений в условиях определенности. Например, в транспортной задаче издержки c_{ij}, связанные с доставкой груза от i-го поставщика к j-му потребителю, считались фиксированной величиной. Если x_{ij} – оптимальное значение переменной, определяющей объем перевозок груза от i-го поставщика к j-му потребителю, то общий вклад в издержки от транспортировки грузов равен произведению c_{ij}, которое также является фиксированной величиной при заданном значении x_{ij}. В реальных экономических условиях приходится решать отдельные задачи при ограниченности, неточности исходной информации о самом объекте и внешней среде, в которой он функционирует и развивается.

При принятии управленческих решений о функционировании и развитии экономического объекта необходимо учитывать важную характеристику внешней среды – неопределенность.

Под *неопределенностью* следует понимать отсутствие, неполноту, недостаточность информации об объекте, процессе, явлении или неуверенность в достоверности информации. В условиях рыночной экономики существует множество источников возникновения неопределенности для различных экономических объектов. Например, к основным источникам возникновения неопределенности на транспорте можно отнести следующие:

1) существенную зависимость транспортного процесса от погодных условий. Например, погодные условия могут вызвать непредвиденные последствия в перевозках сельскохозяйственной продукции;

2) наличие, кроме транспортного предприятия, других участников транспортного процесса – поставщиков грузов, потребителей грузов, ГАИ и др.; результат их влияния на транспортный процесс носит неопределенный и неоднозначный характер;

3) наличие в работе автотранспорта элементов вероятности и случайности (надежность подвижного состава, неравномерность спроса на транспортные услуги во времени и др.);

4) недостаточность, неполнота информации об объекте, процессе, явлении, по отношению к которому принимается решение, ограниченность в сборе и обработке информации, постоянная ее изменчивость;

5) наличие в общественной жизни страны противоборствующих тенденций, столкновение противоречивых интересов;

6) невозможность однозначной оценки объекта при сложившихся в данных условиях уровне и методах научного познания;

7) относительная ограниченность сознательной деятельности лица, принимающего решение, существующие различия в социально-психологических установках, идеалах, намерениях, оценках, стереотипах поведения.

Неопределенность обусловливает появление ситуаций, не имеющих однозначного исхода (решения). Среди различных видов ситуаций, с которыми в процессе производства сталкиваются предприятия, особое место занимают ситуации риска.

Под *ситуацией риска* следует понимать сочетание, совокупность различных обстоятельств и условий, создающих обстановку того или иного вида деятельности. Ей сопутствуют три условия:

- наличие неопределенности;
- необходимость выбора альтернативы (отказ от выбора таковых является разновидностью альтернативы);
- возможность оценить вероятность осуществления выбираемых альтернатив.

Таким образом, если существует возможность количественно и качественно определить степень вероятности того или иного варианта, то это и будет ситуация риска.

Для того чтобы снять ситуацию риска, руководители предприятий вынуждены принимать решения и стремиться реализовать их. Этот процесс находит свое выражение в понятии «риск». Несмотря на то что риск объективно присутствует во всех сферах общественной жизни и в большинстве видов управленческой деятельности, обнаруживается, что понятие «риск» до сих пор не получило универсальной трактовки.

Следует упомянуть об экономическом риске применительно к процессам принятия решений в условиях неопределенности и риска, иными словами, в условиях дефицита информации или неуверенности в достоверности информации. В этом случае риск предстает в виде совокупности вероятных экономических, политических, нравственных и других положительных и неблагоприятных последствий, которые могут наступить при реализации выбранных решений. Определим риск как целенаправленные действия, в ходе которых имеется возможность количественно и качественно оценить вероятность достижения желаемого результата, неудачи и отклонения от цели (положительного или отрицательного свойства).

Процесс установления рыночных отношений в нашей стране порождает различные виды рисковых ситуаций, более того, в работе предприятий риск становится необходимым и обязательным его компонентом.

Чтобы проиллюстрировать различие между ситуациями, когда приходится принимать решения в условиях риска или в условиях неопределенности, рассмотрим задачу оптимального выбора ассортимента выпускаемой продукции.

В условиях риска доход c_j от реализации единицы продукции j не является фиксированной величиной. Напротив, это случайная величина, точное числовое значение которой не известно, но описывается с помощью функции распределения $f(c_j)$. Часть дохода $c_j x_j$, определяемая продукцией j, также случайная величина, если даже значение переменной x_j, определяющей уровень выпуска продукции j, задано.

В условиях неопределенности функция распределения $f_j(c)$ неизвестна. В действительности неопределенность не означает полного отсутствия информации о задаче. Например, известно, что c_j может принимать пять значений, но неизвестны вероятности этих значений. Эта ситуация рассматривается как принятие решений в условиях неопределенности.

Таким образом, с точки зрения полноты исходных данных определенность и неопределенность представляют два крайних случая, а риск определяет промежуточную ситуацию, в которой приходится принимать решение.

Степень неинформированности данных определяет, каким образом задача формализуется и решается.

При решении задач в условиях неопределенности внешней среды наиболее часто возникают две ситуации. При первой ситуации сама система препятствует принятию решений, например задача составления графика выпуска на работу подвижного состава, занимающегося перевозкой сельхозпродукции, в зависимости от того, будет дождь или нет. В этой задаче природа будет восприниматься как «доброжелательный» противник.

Во второй ситуации возможно наличие конкуренции, когда два (или более) участника находятся в конфликте и каждый стремится как можно больше выиграть у другого (других). Эта ситуация отличается от обычных процессов принятия решений в условиях неопределенности тем, что лицу, принимающему решение, противостоит мыслящий противник. Теория, в которой рассматриваются задачи принятия решений в условиях неопределенности при наличии противника («доброжелательного» или мыслящего), известна как *теория игр*.

9.2. Принятие решений в условиях полной определенности

Математические модели исследуемых явлений или процессов могут быть заданы в виде таблиц, элементами которых являются значения частных критериев эффективности функционирования системы, вычисленные для каждой из сравниваемых стратегий при строго заданных внешних условиях. Для рассматриваемых условий принятие решений может производиться:

- по одному критерию;
- по нескольким критериям.

Пример 9.1. Одной из фирм требуется выбрать оптимальную стратегию по обеспечению нового производства оборудованием. С помощью экспериментальных наблюдений были определены значения частных критериев функционирования соответствующего оборудования (a_{ij}), выпускаемого тремя заводами-изготовителями. Рассмотрим данные для выбора оптимальной стратегии в условиях полной определенности (табл. 9.1).

Таблица 9.1

Варианты оборудования (стратегии, решения)	Частные критерии эффективности оборудования*			
	производительность, д. е.	стоимость, д. е.	энергоемкость, у. е.	надежность, у. е.
Оборудование завода 1, x_1	$a_{11} = 5$	$a_{12} = 7$	$a_{13} = 5$	$a_{14} = 6$
Оборудование завода 2, x_2	$a_{21} = 3$	$a_{22} = 4$	$a_{23} = 7$	$a_{24} = 3$
Оборудование завода 3, x_3	$a_{31} = 4$	$a_{32} = 6$	$a_{33} = 2$	$a_{34} = 4$

* Значения частных критериев даны в условных единицах.

На основе экспертных оценок были также определены веса частных критериев λ_j, $j = \overline{1,4}$:

$$\lambda_1 = 0{,}4; \quad \lambda_2 = 0{,}2; \quad \lambda_3 = 0{,}1; \quad \lambda_4 = 0{,}3.$$

Очевидно, выбор оптимальной стратегии (варианта оборудования) по одному критерию в данной задаче не вызывает затруднений. Например, если оценивать оборудование по надежности, то лучшим является оборудование завода 1 (стратегия x_1).

Выбор оптимального решения по комплексу нескольких критериев (в нашем примере – по четырем критериям) является *задачей многокритериальной.*

Один из подходов к решению многокритериальных задач управления связан с процедурой образования обобщенной функции F_i (a_{i1}; a_{i2}; a_{i3}; ...; a_{in}), монотонно зависящей от критериев a_{i1}; a_{i2}; a_{i3}; ...; a_{in}. Данная процедура называется *процедурой (методом) свертывания критериев.* Существует несколько методов свертывания, например:

- метод аддитивной оптимизации;
- метод многоцелевой оптимизации и др.

Рассмотрим подробнее *метод аддитивной оптимизации.*

Пусть

$$F_i(a_{ij}) = \sum_{j=1}^{n} \lambda_j \cdot a_{ij}. \tag{9.1}$$

Здесь выражение (9.1) определяет аддитивный критерий оптимальности. Величины λ_i являются весовыми коэффициентами, которые определяют в количественной форме степень предпочтения j-го критерия по сравнению с другими критериями. Другими словами, коэффициенты λ_j определяют важность j-го критерия оптимальности. При этом более важному критерию приписывается больший вес, а общая важность всех критериев равна единице, т. е.

$$\sum_{j=1}^{n} \lambda_j = 1, \quad \lambda \geq 0, \quad j = \overline{1,n}. \tag{9.2}$$

Обобщенная функция цели (9.1) может быть использована для *свертывания частных критериев оптимальности,* если:

- частные (локальные) критерии количественно соизмеримы по важности, т. е. каждому из них можно поставить в соответствие некоторое число λ_j, которое численно характеризует его важность по отношению к другим критериям;
- частные критерии являются однородными (имеют одинаковую размерность; в нашем примере критерии «стоимость оборудования» и «производительность оборудования» в условных денежных единицах будут однородными).

В этом случае для решения задачи многокритериальной оптимизации оказывается справедливым применение аддитивного критерия оптимальности.

Допустим, в примере 9.1 необходимо выбрать оптимальный вариант оборудования по двум однородным локальным критериям:

- производительность (д. е.);
- стоимость оборудования (д. е.).

На основе экспертных оценок были определены весовые коэффициенты этих двух частных критериев: $\lambda_1 = 0{,}667$, $\lambda_2 = 0{,}333$. Вычислим аддитивный критерий оптимальности для трех вариантов:

$$F_1(a_{1j}) = \lambda_1\, a_{11} + \lambda_2\, a_{12} = 0{,}667 \cdot 5 + 0{,}333 \cdot 7 = 5{,}666;$$
$$F_2(a_{2j}) = \lambda_1\, a_{21} + \lambda_2\, a_{22} = 0{,}667 \cdot 3 + 0{,}333 \cdot 4 = 3{,}333;$$
$$F_3(a_{3j}) = \lambda_1\, a_{31} + \lambda_2\, a_{32} = 0{,}667 \cdot 4 + 0{,}333 \cdot 6 = 4{,}666.$$

Очевидно, первый вариант оборудования по двум частным стоимостным критериям будет оптимальным, так как $F_{max} = F_1(a_{1j}) = 5{,}666$. В примере 9.1 четыре локальных критерия не однородны, т.е. имеют различные единицы измерения. В этом случае требуется нормализация критериев. Под *нормализацией критериев* понимается такая последовательность процедур, с помощью которой все критерии приводятся к единому, безразмерному масштабу измерения. К настоящему времени разработано большое количество схем нормализации. Рассмотрим некоторые из них.

Определим максимум и минимум каждого локального критерия, т. е.

$$a_j^+ = \max a_{ij},\ i = \overline{1,m}; \tag{9.3}$$

$$a_j^- = \min a_{ij},\ i = \overline{1,m}. \tag{9.4}$$

Выделим группу критериев a_j, $j = \overline{1,\ell}$, которые максимизируются при решении задачи, и группу критериев a_j, $j = \overline{\ell + 1,n}$, которые минимизируются при решении задачи.

Тогда в соответствии с принципом максимальной эффективности нормализованные критерии определяются из следующих соотношений:

$$\hat{a}_{ij} = \frac{a_{ij}}{a_j^+},\quad j = \overline{1,\ell}; \tag{9.5}$$

$$\hat{a}_{ij} = 1 - \frac{a_{ij}}{a_j^+},\quad j = \overline{\ell+1,\ n} \tag{9.6}$$

или

$$\hat{a}_{ij} = \frac{a_{ij} - a_j^-}{a_j^+ - a_j^-},\quad j = \overline{1,\ell}; \tag{9.7}$$

$$\hat{a}_{ij} = \frac{a_j^+ - a_{ij}}{a_j^+ - a_j^-},\qquad j = \overline{\ell+1,\ n}. \tag{9.8}$$

Оптимальным будет тот вариант (стратегия), который обеспечивает максимальное значение функции цели:

$$F_i = \sum_{j=1}^{n} \lambda_j \cdot \hat{a}_{ij}, \quad i = \overline{1,m}. \tag{9.9}$$

В соответствии с принципом минимальной потери нормализованные критерии определяются из соотношений

$$\hat{a}_{ij} = 1 - \frac{a_{ij}}{a_j^+}, \quad j = \overline{1,\ell}; \tag{9.10}$$

$$\hat{a}_{ij} = \frac{a_{ij}}{a_j^+}, \quad j = \overline{\ell+1, n} \tag{9.11}$$

или

$$\hat{a}_{ij} = \frac{a_j^+ - a_{ij}}{a_j^+ - a_j^-}, \quad j = \overline{1,\ell}; \tag{9.12}$$

$$\hat{a}_{ij} = \frac{a_{ij} - a_j^-}{a_j^+ - a_j^-}, \quad j = \overline{\ell+1, n}. \tag{9.13}$$

При этом оптимальным будет тот вариант (стратегия), который обеспечивает минимальное значение функции цели (9.9).

Пример 9.2.

Используя данные примера 9.1, определите оптимальную стратегию выбора оборудования из трех возможных ($m = 3$) с учетом четырех локальных критериев ($n = 4$).

Решение

1. Определим max и min каждого локального критерия:

$$a_1^+ = 5;\ a_2^+ = 7;\ a_3^+ = 7;\ a_4^+ = 6.$$

2. При решении задачи максимизируются первый (производительность) и четвертый (надежность) критерии, а минимизируются второй (стоимость оборудования) и третий (энергоемкость) критерии.

3. Исходя из принципа максимизации эффективности, нормализуем критерии:

$$\hat{a}_{i1} = \frac{a_{i1}}{a_1^+}:$$

$$\hat{a}_{11} = \frac{a_{11}}{a_1^+} = \frac{5}{5} = 1;$$

$$\hat{a}_{21} = \frac{a_{21}}{a_1^+} = \frac{3}{5} = 0{,}6;$$

$$\hat{a}_{31} = \frac{a_{31}}{a_1^+} = \frac{4}{5} = 0{,}8.$$

$$\hat{a}_{i4} = \frac{a_{i4}}{a_4^+}:$$

$$\hat{a}_{14} = \frac{a_{14}}{a_4^+} = \frac{6}{6} = 1;$$

$$\hat{a}_{24} = \frac{a_{24}}{a_4^+} = \frac{3}{6} = 0{,}5;$$

$$\hat{a}_{34} = \frac{a_{34}}{a_4^+} = \frac{4}{6} = \frac{2}{3}.$$

$$\hat{a}_{i2} = 1 - \frac{a_{i2}}{a_2^+}:$$

$$\hat{a}_{12} = 1 - \frac{a_{12}}{a_2^+} = 1 - \frac{7}{7} = 0;$$

$$\hat{a}_{22} = 1 - \frac{a_{22}}{a_2^+} = 1 - \frac{4}{7} = \frac{3}{7};$$

$$\hat{a}_{32} = 1 - \frac{a_{32}}{a_2^+} = 1 - \frac{6}{7} = \frac{1}{7}.$$

$$\hat{a}_{i3} = 1 - \frac{a_{i3}}{a_3^+};$$

$$\hat{a}_{13} = 1 - \frac{a_{13}}{a_3^+} = 1 - \frac{5}{7} = \frac{2}{7};$$

$$\hat{a}_{23} = 1 - \frac{a_{23}}{a_3^+} = 1 - \frac{7}{7} = 0;$$

$$\hat{a}_{33} = 1 - \frac{a_{33}}{a_3^+} = 1 - \frac{2}{7} = \frac{5}{7}.$$

4. Определим обобщенную функцию цели по каждому варианту:

$$F_1 = \lambda_1 \hat{a}_{11} + \lambda_2 \hat{a}_{12} + \lambda_3 \hat{a}_{13} + \lambda_4 \hat{a}_{14} =$$

$$= 0,4 \cdot 1 + 0,2 \cdot 0 + 0,1 \cdot \frac{2}{7} + 0,3 \cdot 1 \approx 0,729;$$

$$F_2 = \lambda_1 \hat{a}_{21} + \lambda_2 \hat{a}_{22} + \lambda_3 \hat{a}_{23} + \lambda_4 \hat{a}_{24} =$$

$$= 0,4 \cdot 0,6 + 0,2 \cdot \frac{3}{7} + 0,1 \cdot 0 + 0,3 \cdot 0,5 \approx 0,476;$$

$$F_3 = \lambda_1 \hat{a}_{31} + \lambda_2 \hat{a}_{32} + \lambda_3 \hat{a}_{33} + \lambda_4 \hat{a}_{34} =$$

$$= 0,4 \cdot 0,8 + 0,2 \cdot \frac{1}{7} + 0,1 \cdot \frac{5}{7} + 0,3 \cdot \frac{2}{3} \approx 0,603.$$

Оптимальным является первый вариант оборудования, так как $F_{max} = F_1 = 0,729$.

Рассмотренный подход к решению многокритериальных задач зачастую применяется при решении экономических задач, связанных с оценкой качества промышленной продукции и оценкой уровня технического совершенства технических устройств и систем по нескольким показателям.

Другим возможным методом решения многокритериальных задач является *метод последовательных уступок*. Вначале критерии ранжируются и нумеруются в порядке убывания важности. Абсолютное значение коэффициентов важности λ_j на этом этапе не играет никакой роли. Оптимизируется первый по важности критерий a_1 и определяется его экстремальное значение a_1^*. Затем назначается величина допустимого отклонения критерия от оптимального значения (уступка) Δa_1 и ищется экстремальное значение второго по важности критерия a_2, при условии, что отклонение первого от оптимального значения не превзойдет величины уступки. Затем назначается уступка для второго критерия, и задача оптимизируется по третьему критерию и т. д. Таким образом, многокритериальная задача оптимизации заменяется последовательностью однокритериальных задач. Решение каждой предыдущей задачи используется при решении последующих для формирования дополнительных условий, состоящих в ограничении на величину уступки.

9.3. Принятие решений в условиях риска

Основными критериями оценки принимаемых решений в условиях риска являются:

- ожидаемое значение результата;
- ожидаемое значение результата в сочетании с минимизацией его дисперсии;
- известный предельный уровень результата;
- наиболее вероятное событие (исход) в будущем.

Критерий ожидаемого значения используется в случаях, когда требуется определить экстремальное значение (max или min) результативного показателя (прибыль, расходы, экономические потери и т. д.). Применение этого критерия рассмотрим на конкретном примере, связанном с постановкой задачи проведения ремонтно-профилактических воздействий автомобилей. Оптимальное количество ремонтных воздействий, определенное минимизацией суммарных затрат на заданной наработке $L_к$ с учетом рисков пропуска отказов и выполнения лишних ТО, приравнивается к количеству ТО на указанном пробеге. Модель данной задачи является моделью вероятностного спроса на ремонты с мгновенным восстановлением. Здесь минимизируются суммарные издержки за пробег $L_к$, которые определяются затратами на плановый ремонт $S_р$, профилактику $S_{ТО}$ и незапланированный аварийный ремонт $S_ш$, рассматриваемый как штраф за пропуск отказа:

$$S = S_р + S_{ТО} + S_ш \to \min. \tag{9.14}$$

Составляющие суммарных затрат формулы (9.14) зависят от количества ремонтно-профилактических операций за наработку $L_к$, определяемых по формуле

$$n = \frac{L_к}{L_{от}}, \tag{9.15}$$

где $L_{от}$ – наработка до отказа.

Наработка до отказа – величина случайная, определяемая плотностью распределения $f(L_{от})$, $L_{от} < L_к$. В силу случайности $L_{от}$ величина n также будет случайной с плотностью распределения

$$f(n)=\left|-\frac{L_к}{n^2}\right|\cdot f\left(\frac{L_к}{n}\right). \qquad (9.16)$$

Используя $f(n)$ как весовую функцию и выражая составляющие суммарных затрат через соответствующие стоимости из (9.14), получим

$$S = C_р n_р + \int\limits_0^{n_р} C_{ТО}(n_р - n)f(n)dn + \int\limits_{n_р}^{\infty} C_ш(n - n_р)f(n)dn \to \min, \qquad (9.17)$$

где $C_р$ – средняя стоимость предупредительного (планового) ремонта;
$C_{то}$ – средняя стоимость профилактики (или убыток от недоиспользования ресурса замененных при ТО деталей);
$C_ш$ – ущерб (штраф) от пропуска отказа (или стоимость устранения аварийного отказа). Очевидно, $C_ш > C_{то}$.

Интеграл (9.16) в пределах $[0, n_р]$ соответствует риску выполнения лишних ТО (избыточность затрат на ТО), а интеграл в пределах $[n_р, \infty]$ – риску пропуска аварийных отказов (избыточность затрат на ТР по потребности). Из уравнения (9.17) находим оптимальное количество ремонтов $n_р$ на пробеге $L_к$ (обычно $L_к$ – пробег до КР). Далее, заменяя необходимые ремонты обслуживаниями, при которых выполняется комплекс операций по предупреждению отказов, включая предупредительные замены деталей, получим

$$L_{ТО} = L_к / n_р. \qquad (9.18)$$

Пример 9.3. Определить оптимальную периодичность ТО (у. е.) при $L_к = 200$ тыс. км, $C_ш = 69$, $C_р = 24$, $C_{то} = 15$, если наработки до отказа имеют нормальное распределение с параметрами $\bar{L}_{от} =$ $= 20$ тыс. км и $\sigma_L = 5$ тыс. км.

$$f(L_{от}) = \frac{1}{\sigma_L\sqrt{2\pi}}\exp\left[-0{,}5\cdot\left(\frac{L_{от}-\bar{L}_{от}}{\sigma_L}\right)^2\right]. \qquad (9.19)$$

Решение

Выполнив преобразование распределения (9.19) по формуле (9.15), получим ($n \geq 1$):

$$f(n)=\frac{L_{\text{к}}}{n^2\sigma_L\sqrt{2\pi}}\exp\left[-0{,}5\left(\frac{L_{\text{к}}/n-\bar{L}_{\text{от}}}{\sigma_L}\right)^2\right]. \qquad (9.20)$$

После подстановки выражения (9.20) в (9.17) получим задачу оптимизации, для решения которой воспользуемся математическим пакетом EUREKA.

Решая задачу, получим оптимальную периодичность $L_{\text{ТО}}$ = 15,3 тыс. км при $n_{\text{р}}$ = 13,08, которая обеспечивает минимальные суммарные издержки S.

Критерий ожидаемого значения позволяет получить достоверные оценки в случае, когда одно и то же решение приходится принимать достаточно большое число раз, так как замена математического ожидания выборочными данными правомерна лишь при большом объеме выборки.

Если необходимость в принятии решения встречается редко, то выборочное значение может значительно отличаться от математического ожидания, а применение критерия ожидаемых значений может приводить к ошибочным результатам. В таких случаях рекомендуется применять *критерий ожидаемого значения в сочетании с минимизацией его дисперсии*, что приближает выборочное значение к математическому ожиданию. Критерий принимает следующий вид:

$$\begin{cases} M(X)+K\cdot D(X)\to\min \\ M(X)-K\cdot D(X)\to\max, \end{cases} \qquad (9.21)$$

где X – случайная величина (например, суммарные издержки);
$D(X)$ – дисперсия этой величины;
K – заданная постоянная.

Постоянную K иногда интерпретируют как уровень несклонности к риску. Считается, что K определяет «степень важности» дисперсии $D(X)$ по отношению к $M(X)$. Например, предприниматель, особенно остро реагирующий на большие отрицательные отклонения прибыли вниз от $M(X)$, может выбрать K много больше единицы. Это придает больший вес дисперсии и приводит к решению, уменьшающему большие потери прибыли.

Критерий предельного уровня не позволяет получить оптимальное решение, найти максимум прибыли и минимум расходов. Этот критерий дает возможность определить приемлемый (допустимый) способ действий. Например, транспортная фирма распродает автомобили, бывшие в эксплуатации. По каждой модели автомобиля определенного возраста определяется лимитная цена, т. е. минимально допустимая цена продажи автомобиля. Продажа автомобилей по цене ниже лимитной приведет к убыточной работе транспортной фирмы. Это и есть предельный уровень, позволяющий транспортной фирме согласиться на первое же превышающее этот уровень предложение цены. Такой критерий не определяет оптимальное решение, поскольку одно из последующих предложений может оказаться более выгодным, чем принятое.

Одно из преимуществ критерия предельного уровня заключается в том, что для него нет необходимости задавать в явном виде плотность распределения случайных величин. В нашем примере случайная величина – рыночная цена автомобиля. Транспортная фирма располагает информацией о распределении рыночных цен на подобные автомобили в неявном виде. Иначе при полном отсутствии информации о распределении рыночных цен фирма установила бы предельные цены на автомобили очень высокими или, наоборот, очень низкими.

Критерий наиболее вероятного события (исхода) основан на преобразовании случайной ситуации в детерминированную путем замены случайной величины единственным значением, имеющим наибольшую вероятность реализации.

9.4. Принятие решений в условиях неопределенности

Неопределенность является характеристикой внешней среды (природы), в которой принимается управленческое решение о развитии (или функционировании) экономического объекта. Здесь будем рассматривать неопределенность «природы», вызванную отсутствием, недостатком информации о действительных условиях (факторах), при которых развивается объект управления. Внешняя среда («природа») может находиться в одном из множества возможных состояний. Это множество может быть конечным и бесконечным. Будем считать, что множество состояний конечно или по крайней мере количество состояний можно пронумеровать.

Пусть S_i – состояние «природы», при этом $i = \overline{1, n}$, где n – число возможных состояний. Все возможные состояния известны, не известно только, какое состояние будет иметь место в условиях, когда планируется реализация принимаемого управленческого решения. Будем считать, что множество управленческих решений (планов) R_j также конечно и равно m. Реализация R_j плана в условиях, когда «природа» находится в S_i состоянии, приводит к определенному результату, который можно оценить, введя количественную меру. В качестве этой меры могут служить выигрыши от принимаемого решения (плана); потери от принимаемого решения, а также полезность, риск и другие количественные критерии.

Данные, необходимые для принятия решения в условиях неопределенности, обычно задаются в форме матрицы, строки которой соответствуют возможным действиям (управленческим решения) R_j, а столбцы – возможным состояниям «природы» S_i.

Допустим, каждому R_j-му действию и каждому возможному S_i-му состоянию «природы» соответствует результат (исход), определяющий результат (выигрыш, полезность) при выборе j-го действия и реализации i-го состояния, – V_{ji}.

$$
\begin{array}{c|cccccc}
 & S_1 & S_2 & \dots & S_i & \dots & S_n \\
\hline
R_1 & V_{11} & V_{12} & \dots & V_{1i} & \dots & V_{1n} \\
R_2 & V_{21} & V_{22} & \dots & V_{2i} & \dots & V_{2n} \\
\vdots & \vdots & \vdots & & \vdots & & \vdots \\
R_j & V_{j1} & V_{j2} & \dots & V_{ji} & \dots & V_{jn} \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
R_m & V_{mi} & V_{m2} & \dots & V_{mi} & \dots & V_{mn}.
\end{array}
\qquad (9.22)
$$

Следовательно, математическая модель задачи принятия решений определяется множеством состояний $\{S_i\}$, множеством планов (стратегий) $\{R_j\}$ и матрицей возможных результатов $\|V_{ji}\|$. В качестве результатов в отдельных задачах рассматривается матрица рисков $\|r_{ji}\|$.

Риск – мера несоответствия между разными возможными результатами принятия определенных стратегий (действий).

Элементы матрицы рисков $\|r_{ji}\|$ связаны с элементами матрицы полезностей (выигрышей) следующим соотношением:

$$r_{ji} = V_i - V_{ji}, \qquad (9.23)$$

где $V_i = \max\limits_j V_{ji}$ – максимальный элемент в столбце i матрицы полезностей.

Если матрица возможных результатов $\|V_{ji}\|$ представляет собой матрицу потерь (затрат), то элементы матрицы рисков $\|r_{ji}\|$ следует определять по формуле

$$r_{ji} = V_{ji} - V_i, \qquad (9.24)$$

где $V_i = \min\limits_j V_{ji}$ – минимальный элемент в столбце i матрицы потерь (результатов).

Таким образом, риск – это разность между результатом, который можно получить, если знать действительное состояние «природы», и результатом, который будет получен при j-й стратегии.

Матрица рисков дает более наглядную картину неопределенной ситуации, чем матрица выигрышей (полезностей).

Непосредственный анализ матриц выигрышей $\|V_{ji}\|$ или рисков $\|r_{ji}\|$ не позволяет в общем случае принять решение по выбору оптимальной стратегии (плана), за исключением тривиального случая, когда выигрыши при одной стратегии выше, чем при любой другой для каждого состояния «природы» (элементы матрицы выигрышей в некоторой строке больше, чем в любой из других). Другими словами, имеется в наличии «доминирующая» стратегия.

Для принятия решения в условиях неопределенности используется ряд критериев. Рассмотрим некоторые из них. Это критерий Лапласа, критерий Вальда, критерий Сэвиджа, критерий Гурвица.

Критерий Лапласа. Этот критерий опирается на «принцип недостаточного основания» Лапласа, согласно которому все состояния «природы» S_i, $i = \overline{1,n}$ полагаются равновероятными. В соответствии с этим принципом каждому состоянию S_i ставится вероятность q_i, определяемая по формуле

$$q_i = \frac{1}{n}. \tag{9.25}$$

При этом исходной может рассматриваться задача принятия решения в условиях риска, когда выбирается действие R_j, дающее наибольший ожидаемый выигрыш. Для принятия решения для каждого действия R_j вычисляют среднее арифметическое значение выигрыша:

$$M_j(R) = \frac{1}{n}\sum_{i=1}^{n} V_{ji}. \tag{9.26}$$

Среди $M_j(R)$ выбирают максимальное значение, которое будет соответствовать оптимальной стратегии R_j.

Другими словами, находится действие R_j^*, соответствующее

$$\max_{R_j}\left\{\frac{1}{n}\sum_{i=1}^{n} V_{ji}\right\}. \tag{9.27}$$

Если в исходной задаче матрица возможных результатов представлена матрицей рисков $\|r_{ji}\|$, то критерий Лапласа принимает следующий вид:

$$\min_{R_j} \left\{ \frac{1}{n} \sum_{i=1}^{n} r_{ji} \right\}. \tag{9.28}$$

Пример 9.4. Одно из транспортных предприятий должно определить уровень своих провозных возможностей так, чтобы удовлетворить спрос клиентов на транспортные услуги на планируемый период. Спрос на транспортные услуги не известен, но ожидается (прогнозируется), что он может принять одно из четырех значений: 10, 15, 20 или 25 тыс. т. Для каждого уровня спроса существует наилучший уровень провозных возможностей транспортного предприятия (с точки зрения возможных затрат). Отклонения от этих уровней приводят к дополнительным затратам либо из-за превышения провозных возможностей над спросом (из-за простоя подвижного состава), либо из-за неполного удовлетворения спроса на транспортные услуги. Ниже приводится табл. 9.3, определяющая возможные прогнозируемые затраты на развитие провозных возможностей:

Таблица 9.3

Варианты провозных возможностей транспортного предприятия	Варианты спроса на транспортные услуги, д. е.			
	1	2	3	4
1	6	12	20	24
2	9	7	9	28
3	23	18	15	19
4	27	24	21	15

Необходимо выбрать оптимальную стратегию.

Решение

Согласно условию задачи имеются четыре варианта спроса на транспортные услуги, что равнозначно наличию четырех состояний «природы»: S_1, S_2, S_3, S_4. Известны также четыре стратегии развития провозных возможностей транспортного предприятия: R_1, R_2, R_3, R_4. Затраты на развитие провозных возможностей при каждой паре S_i и R_i заданы следующей матрицей (таблицей):

$$V = \begin{array}{c|cccc} & S_1 & S_2 & S_3 & S_4 \\ \hline R_1 & 6 & 12 & 20 & 24 \\ R_2 & 9 & 7 & 9 & 28 \\ R_3 & 23 & 18 & 15 & 19 \\ R_4 & 27 & 24 & 21 & 15. \end{array}$$

Принцип Лапласа предполагает, что S_1, S_2, S_3, S_4 равновероятны. Следовательно, $P\{S = S_i\} = \frac{1}{n} = \frac{1}{4} = 0{,}25$, i = 1, 2, 3, 4 и ожидаемые затраты при различных действиях R_1, R_2, R_3, R_4 составляют:

$$W\{R_1\} = 0{,}25\ (6 + 12 + 20 + 24) = 15{,}5;$$

$$W\{R_2\} = 0{,}25\ (9 + 7 + 9 + 28) = 13{,}25;$$

$$W\{R_3\} = 0{,}25\ (23 + 18 + 15 + 19) = 18{,}75;$$

$$W\{R_4\} = 0{,}25\ (27 + 24 + 21 + 15) = 21{,}75.$$

Таким образом, наилучшей стратегией развития провозных возможностей в соответствии с критерием Лапласа будет R_2.

Критерий Вальда (минимаксный или максиминный критерий). Применение данного критерия не требует знания вероятностей состояний S_i. Этот критерий опирается на принцип наибольшей осторожности, поскольку он основывается на выборе наилучшей из наихудших стратегий R_j.

Если в исходной матрице (по условию задачи) результат V_{ji} представляет потери лица, принимающего решение, то при выборе оптимальной стратегии используется *минимаксный критерий*. Для определения оптимальной стратегии R_j необходимо в каждой строке матрицы результатов найти наибольший элемент $\max_i\{V_{ji}\}$, а затем выбирается действие R_j (строка j), которому будет соответствовать наименьший элемент из этих наибольших элементов, т. е. действие, определяющее результат, равный

$$W = \min_j \max_i \left\{V_{ji}\right\}. \qquad (9.29)$$

Если в исходной матрице по условию задачи результат V_{ji} представляет выигрыш (полезность) лица, принимающего решение, то при выборе оптимальной стратегии используется *максиминный критерий.*

Для определения оптимальной стратегии R_j в каждой строке матрицы результатов находят наименьший элемент $\min_j\left\{V_{ji}\right\}$, а затем выбирается действие R_j (строка j), которому будут соответствовать наибольшие элементы из этих наименьших элементов, т. е. действие, определяющее результат, равный

$$W = \min_j \max_i \left\{V_{ji}\right\}. \qquad (9.30)$$

Пример 9.5. Рассмотрим пример 9.4. Так как V_{ji} в этом примере представляет потери (затраты), применим минимаксный критерий. Необходимые результаты вычисления приведены в табл. 9.4:

Таблица 9.4

Состояния S_i / Стратегия R_j	Затраты, д. е. (V_{ji})				max $\{V_{ji}\}$	$W =$ $= \min\max\ \{V_{ji}\}$
	S_1	S_2	S_3	S_4		
R_1	6	12	20	24	24	–
R_2	9	7	9	28	28	–
R_3	23	18	15	19	23	23
R_4	27	24	21	15	27	–

Таким образом, наилучшей стратегией развития провозных возможностей в соответствии с минимаксным критерием «лучшим из худших» будет третья, т. е. R_3.

Минимаксный критерий Вальда иногда приводит к нелогичным выводам из-за своей чрезмерной «пессимистичности». «Пессимистичность» этого критерия исправляет критерий Сэвиджа.

Критерий Сэвиджа использует матрицу рисков $\|r_{ji}\|$. Элементы данной матрицы можно определить по формулам (9.23), (9.24), которые перепишем в следующем виде:

$$r_{ji} = \begin{cases} \max\limits_{j}\{V_{ji}\} - V_{ji}, & \text{если } V - \text{выигрыш} \\ V_{ji} - \min\limits_{j}\{V_{ji}\}, & \text{если } V - \text{потери.} \end{cases} \tag{9.31}$$

Это означает, что r_{ji} есть разность между наилучшим значением в столбце i и значениями V_{ji} при том же i. Отметим, что независимо от того, является ли V_{ji} доходом (выигрышем) или потерями (затратами), r_{ji} в обоих случаях определяет величину потерь лица, принимающего решение. Следовательно, можно применять к r_{ji} только минимаксный критерий. Критерий Сэвиджа рекомендует в условиях неопределенности выбирать ту стратегию R_j, при которой величина риска принимает наименьшее значение в самой неблагоприятной ситуации (когда риск максимален).

Пример 9.6. Рассмотрим пример 9.4. Заданная матрица определяет потери (затраты). По формуле (9.31) вычислим элементы матрицы рисков $\|r_{ji}\|$:

$$
\left\| r_{ji} \right\| = \begin{array}{c|cccc}
 & S_1 & S_2 & S_3 & S_4 \\
\hline
R_1 & 0 & 5 & 11 & 9 \\
R_2 & 3 & 0 & 0 & 13 \\
R_3 & 17 & 11 & 6 & 4 \\
R_4 & 21 & 17 & 12 & 0
\end{array}.
$$

Полученные результаты вычислений с использованием критерия минимального риска Сэвиджа оформим в табл. 9.5.

Таблица 9.5

Состояние S_i / Стратегия R_j	Величина риска, д. е. (V_{ji})				$\max \{r_{ji}\}$	$W = \min \max \{r_{ji}\}$
	S_1	S_2	S_3	S_4		
R_1	0	5	11	9	11	11
R_2	3	0	0	13	13	–
R_3	17	11	6	4	17	–
R_4	21	17	12	0	21	–

Введение величины риска r_{ji} привело к выбору первой стратегии R_1, обеспечивающей наименьшие потери (затраты) в самой неблагоприятной ситуации (когда риск максимален).

Применение критерия Сэвиджа позволяет любыми путями избежать большого риска при выборе стратегии, а значит, избежать большего проигрыша (потерь).

Критерий Гурвица. Основан на следующих двух предположениях: «природа» может находиться в самом невыгодном состоянии с вероятностью $(1 - \alpha)$ и в самом выгодном состоянии с вероятностью α, где α – коэффициент доверия. Если результат V_{ji} – прибыль, полезность, доход и т. п., то критерий Гурвица записывается так:

$$
W = \max_j \left[\alpha \max_i V_{ji} + (1 - \alpha) \min_i V_{ji} \right]. \tag{9.32}
$$

Когда V_{ji} представляет затраты (потери), то выбирают действие, дающее

$$
W_{\min} = \min_j \left[\alpha \min_i V_{ji} + (1 - \alpha) \max_i V_{ji} \right]. \tag{9.33}
$$

Если $\alpha = 0$, получим пессимистический критерий Вальда.

Если $\alpha = 1$, то приходим к решающему правилу вида $\max_j \min_i V_{ji}$ или к так называемой стратегии «здорового оптимиста», т. е. критерий слишком оптимистичный.

Критерий Гурвица устанавливает баланс между случаями крайнего пессимизма и крайнего оптимизма путем взвешивания обоих способов поведения соответствующими весами $(1 - \alpha)$ и α, где $0 \le \alpha \le 1$. Значение α от 0 до 1 может определяться в зависимости от склонности лица, принимающего решение, к пессимизму или к оптимизму. При отсутствии ярко выраженной склонности $\alpha = 0{,}5$ представляется наиболее разумной.

Пример 9.7. Критерий Гурвица используем в примере 9.4. Положим $\alpha = 0{,}5$. Результаты необходимых вычислений приведены в табл. 9.6.

Таблица 9.6

W_j	$\max\limits_i V_{ji}$	$\min\limits_i V_{ji}$	$\alpha \cdot \min\limits_i V_{ji} + (1-\alpha) \cdot \max\limits_i V_{ji}$	$\min\limits_j W_j$
W_1	6	24	15	15
W_2	7	28	17,5	–
W_3	15	23	19	–
W_4	15	27	21	–

Оптимальное решение заключается в выборе W.

Таким образом, в примере предстоит сделать выбор:

по критерию Лапласа – выбор стратегии R_2;

по критерию Вальда – выбор стратегии R_3;

по критерию Сэвиджа – выбор стратегии R_1;

по критерию Гурвица – выбор стратегии R_1 при $\alpha = 0{,}5$, а если лицо, принимающее решение, – пессимист ($\alpha = 0$), то выбор стратегии R_3.

Какое из возможных решений предпочтительнее?

Это определяется выбором соответствующего критерия (Лапласа, Вальда, Сэвиджа или Гурвица).

Выбор критерия принятия решений в условиях неопределенности является наиболее сложным и ответственным этапом в исследовании операций. При этом не существует каких-либо общих советов или рекомендаций. Выбор критерия должно производить лицо, принимающее решение (ЛПР), с учетом конкретной специфики решаемой задачи и в соответствии со своими целями, а также опираясь на прошлый опыт и собственную интуицию.

В частности, если даже минимальный риск недопустим, то следует применять критерий Вальда. Если, наоборот, определенный риск вполне приемлем и ЛПР намерено вложить в некоторое предприятие столько средств, чтобы потом оно не сожалело, что вложено слишком мало, то выбирают критерий Сэвиджа.

9.5. Теория игр

В отличие от рассмотренных выше задач принятия решений в условиях определенности, риска и неопределенности, в которых внешняя среда (природа) предполагалась пассивной, в конфликтных ситуациях имеются противодействующие стороны, интересы которых противоположны. При конфликтных ситуациях решения принимаются в условиях неопределенности двумя и более разумными противниками, каждый из которых стремится оптимизировать свои решения за счет других. Теория, занимающаяся принятием решений в условиях конфликтных ситуаций, называется *теорией* игр. Математическая модель конфликтной ситуации представляет собой игру.

Игра – это совокупность правил, описывающих сущность конфликтной ситуации. Эти правила устанавливают:

- выбор образа действия игроков на каждом этапе игры;
- информацию, которой обладает каждый игрок при осуществлении таких выборов;
- плату для каждого игрока после завершения любого этапа игры.

Игру можно определить следующим образом:

- имеются n конфликтующих сторон (игроков), принимающих решения, интересы которых не совпадают;
- сформулированы правила выбора допустимых стратегий, известные игрокам;
- определен набор возможных конечных состояний игры (например, выигрыш, ничья, проигрыш);
- всем игрокам (участникам игры) заранее известны платежи, соответствующие каждому возможному конечному состоянию. Платежи задаются в виде матрицы $\overline{A} = \|a_{ij}\|$.

В зависимости от числа конфликтующих сторон игры делятся на парные (с двумя игроками) и множественные (имеющие не менее трех игроков). Каждый игрок имеет некоторое множество (конечное или бесконечное) возможных выборов, т. е. стратегий.

Стратегией игры называется совокупность правил, определяющих поведение игрока от начала игры до ее завершения. Стратегии каждого игрока определяют результаты или платежи в игре. Игра называется *игрой с нулевой суммой*, если проигрыш одного игрока равен выигрышу другого, в противном случае она называется *игрой с ненулевой суммой*.

В подразд. 9.5 рассматриваются только игры двух лиц с нулевой суммой. Задание стратегий (A и B) двух игроков в парной игре полностью определяет ее исход, т. е. выигрыш одного или проигрыш другого. Игра называется *конечной*, если у каждого игрока имеется конечное число стратегий. Результаты конечной парной игры с нулевой суммой можно задавать матрицей, строки и столбцы которой соответствуют различным стратегиям, а ее элементы – выигрышам одной стороны (равные проигрышам другой). Эта матрица называется *платежной матрицей* или *матрицей игры.*

Если первый игрок имеет m стратегий, а второй – n стратегий, то говорят, что мы имеем дело с игрой $m \times n$.

Рассмотрим игру $m \times n$. Пусть заданы множество стратегий: для первого игрока $\{A_i\}$, для второго игрока $\{B_j\}$, платежная матрица $\bar{A}\, m{\times}n = ||a_{ij}||$, где a_{ij} – выигрыш первого игрока или проигрыш второго игрока при выборе ими стратегий A_i и B_j соответственно. Каждый из игроков выбирает однозначно с вероятностью 1 некоторую стратегию, т.е. пользуется при выборе решения *чистой стратегией.* При этом *решение* игры будет в чистых стратегиях. Поскольку интересы игроков противоположны, то первый игрок стремится максимизировать свой выигрыш, а второй игрок, наоборот, минимизировать свой проигрыш.

Решение игры состоит в определении наилучшей стратегии каждым игроком. Выбор наилучшей стратегии одним игроком проводится при полном отсутствии информации о принимаемом решении вторым игроком. Следует отметить, что и первый, и второй игрок являются разумными противниками, которые находятся в состоянии конфликта. Поэтому для решения игры двух лиц с нулевой суммой используется очень «пессимистичный» критерий, так называемый критерий *мини-макса-максимина.* Этот критерий рассмотрен в подразд. 9.4. Основное отличие заключается в том, что ранее «природа» не рассматривалась как активный противник, тогда как в теории игр каждый игрок действует разумно и, следовательно, пытается активно помешать своему противнику. Так, если первый игрок применяет стратегию A_i, то второй будет стремиться к тому, чтобы выбором соответствующей стратегии B_j свести выигрыш первого игрока к минимуму, что равнозначно сведению своего проигрыша к минимуму. Величина этого минимума

$$\alpha_i = \min_j a_{ij}, \quad i = \overline{1,m}. \tag{9.34}$$

Первый игрок (при любых ответах противника) будет стремиться найти такую стратегию, при которой α_i обращается в максимум:

$$\alpha = \max_i \alpha_i = \max_i \min_j a_{ij}. \tag{9.35}$$

Величина α называется *нижней ценой игры.* Ей соответствует максиминная стратегия, придерживаясь которой первый игрок при любых стратегиях противника обеспечит себе выигрыш, не меньший α. Другими словами, нижняя цена игры является гарантированным выигрышем первого игрока при любых стратегиях второго игрока.

Аналогично определим по каждому столбцу матрицы $\beta_j = \max_i \alpha_{ij}$, $j = \overline{1, n}$, найдем минимальное значение β_j:

$$\beta = \min_j \beta_j = \min_j \max_i a_{ij}. \tag{9.36}$$

Величина β называется *верхней ценой игры.* Ей соответствует минимаксная стратегия второго игрока. Величина β представляет собой гарантированный проигрыш второго игрока при любой стратегии первого игрока.

Пример 9.8. Дана платежная матрица 3×4, которая определяет выигрыши игрока A. Вычислить нижнюю и верхнюю цены заданной игры.

$$\overline{A} = \left\| a_{ij} \right\| = \begin{pmatrix} 10 & 4 & 11 & 7 \\ 7 & 6 & 8 & 20 \\ 6 & 2 & 1 & 11 \end{pmatrix}.$$

Решение

Представим нашу игру в виде табл. 9.7:

Таблица 9.7

Стратегии первого игрока, A_i	Стратегии второго игрока, B_j				Значение, α_i	α
	B_1	B_2	B_3	B_4		
A_1	10	4	11	7	4	–
A_2	7	6	8	20	6	6
A_3	6	2	1	11	1	–
Значение β_i	10	6	11	20	–	
β	–	6	–	–	–	–

Если игрок A выбирает первую стратегию, он может получить выигрыш в размере 10, 4, 11 или 7 д. е. в зависимости от выбранной стратегии игроком B. При этом выигрыш игрока будет не меньше $\alpha_1 = \min\{10; 4; 11; 7\} = 4$ д. е. независимо от поведения игрока B. Аналогично при выборе игроком A второй стратегии гарантированный выигрыш $\alpha_2 = \min\{7; 6; 8; 20\} = 6$ д. е. При выборе игроком A третьей стратегии выигрыш $\alpha_3 = \min\{6; 2; 1; 11\} = 1$ д. е.

Таким образом, минимальные значения α_i, $i = \overline{1, 3}$ определяют минимально гарантированный выигрыш для игрока A, если он выбирает соответствующую стратегию i. Величина $\max\limits_i \alpha_i = \max\{4; 6; 1\} = 6$ д.е. будет гарантированным выигрышем игрока A при любых стратегиях игрока B. Выбранная игроком A вторая стратегия называется *максиминной стратегией*, а соответствующее ее значение выигрыша $\alpha_2 = 6$ д. е. будет нижней ценой игры.

Второй игрок стремится минимизировать свой проигрыш. Выбрав первую стратегию β_1, игрок B может проиграть не более чем $\beta_1 = \max\{10; 7; 6\} = 10$ д. е. независимо от выбора стратегии игроком A. Аналогично рассуждая, получим следующие результаты (д. е.):

$$\beta_2 = \max\{4; 6; 2\} = 6; \ \beta_3 = \max\{11; 8; 1\} = 11;$$
$$\beta_4 = \max\{7; 20; 11\} = 20.$$

Игрок B выбирает стратегию β_2, которая минимизирует его максимальные проигрыши:

$$\beta = \min_j \beta_j = \min\{10; 6; 11; 20\} = 6 \text{ д. е.}$$

Величина $\beta = 6$ д. е. будет гарантированным проигрышем игрока B при любых стратегиях игрока A. Выбранная игроком B вторая стратегия называется *минимаксной стратегией*, а соответствующее ее значение проигрыша $\beta_2 = 6$ д. е. будет верхней ценой игры.

Следует отметить, что для любой матрицы $\bar{A} = \|a_{ij}\|$ выполняется неравенство

$$\beta \geq \alpha. \tag{9.37}$$

Если $\beta = \alpha$, т. е. верхняя цена равна нижней цене игры, то соответствующие чистые стратегии называются *оптимальными*, а про игру говорят, что она имеет *седловую точку*. Седловая точка является минимальным элементом соответствующей строки и максимальным элементом соответствующего столбца. Эта точка есть точка равновесия игры, определяющая однозначно оптимальные стратегии. Оптимальность здесь означает, что ни один игрок не стремится изменить свою стратегию, так как его противник может на это ответить выбором другой стратегии, дающей худший для первого игрока результат.

Величина $C = \beta = \alpha$ называется *ценой игры*. Она определяет средний выигрыш игрока A и средний проигрыш игрока B при использовании ими оптимальных стратегий. В нашем примере цена игры $C = 6$ д. е., оптимальная пара стратегий – A_2 и B_2.

Если в платежной матрице A все элементы строки $A_i = (a_{i1}, a_{i2}, ..., a_{in})$ не меньше соответствующих элементов строки $A_k = (a_{k1}, a_{k2},$

..., a_{kn}), а по крайней мере один строго больше, то строка A_i называется *доминирующей*, а строка A_k – *доминируемой*.

Аналогичны понятия «доминирующий столбец» и «доминируемая строка».

Первому игроку невыгодно применять стратегии, которым соответствуют доминируемые строки; второму игроку невыгодно применять стратегии, которым соответствуют доминирующие столбцы. Поэтому при решении игры можно уменьшить размеры платежной матрицы путем удаления из нее доминирующих столбцов и доминируемых строк.

Пример 9.9. Для игры с платежной матрицей

$$A = \begin{vmatrix} 2 & -3 & 5 \\ 4 & -2 & 3 \\ 6 & -1 & 4 \end{vmatrix}$$

найдите стратегии игроков и цену игры.

Решение

Элемент $a_{32} = -1$ является наименьшим в третьей строке и наибольшим во втором столбце, т. е. он является седловой точкой. Поэтому цена игры $C = -1$, а оптимальные стратегии игроков: первого – A_3, а второго – B_2.

Используя понятия доминируемых строк и доминирующих столбцов, задачу можно решить следующим образом.

В матрице A третья строка доминирует над второй, поэтому вторую можно изъять из платежной матрицы. В результате получится матрица

$$A_1 = \begin{pmatrix} 2 & -3 & 5 \\ 6 & -1 & 4 \end{pmatrix}.$$

В матрице A_1 первый и третий столбцы доминируют над вторым, следовательно, их можно изъять. В результате платежная матрица принимает вид

$$A_2 = \begin{pmatrix} -3 \\ -1 \end{pmatrix}.$$

В матрице A_2 вторая строка доминирует. После вычеркивания получится матрица A_3, состоящая из одного элемента:

$$A_3 = (-1).$$

Элемент матрицы A_3 и определяет решение нашей задачи.

Отдельные игры могут не иметь седловых точек, т. е. у каждого игрока не существует единственной, наиболее надежной стратегии. В этом случае используют *смешанную* стратегию. Смешанная стратегия состоит в том, что в ходе игры происходит случайный выбор стратегии из некоторого множества смешанных стратегий и для каждой смешанной стратегии указывается вероятность ее выбора. Смешанная стратегия для игрока A представляет собой вектор

$$P = \begin{Vmatrix} P_1 \\ P_2 \\ \vdots \\ P_m \end{Vmatrix}, \qquad (9.38)$$

где P_i – вероятность выбора i-й стратегии игроком и удовлетворяет следующим условиям:

$$P_i \geq 0,\ i = \overline{1,m};$$

$$\sum_{i=1}^{m} P_i = 1. \qquad (9.39)$$

Аналогично смешанная стратегия игрока B представляет собой вектор

$$Q = \begin{Vmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{Vmatrix}, \qquad (9.40)$$

где q_j – вероятность выбора j-й стратегии игроком B – удовлетворяет следующим условиям;

$$q_j \geq 0,\ j = \overline{1,\ n}; \qquad (9.41)$$

$$\sum_{j=1}^{n} q_j = 1. \qquad (9.42)$$

Платежная матрица игры имеет следующий вид (табл. 9.8):

Таблица 9.8

$A \backslash B$	q_1	q_2	q_3	...	q_n
P_1	a_{11}	a_{12}	a_{13}	...	a_{1n}
P_2	a_{21}	a_{22}	a_{23}	...	a_{2n}
P_3	a_{31}	a_{32}	a_{33}	...	a_{3n}
.	.	.	.		.
P_m	a_{m1}	a_{m2}	a_{m3}	...	a_{mn}

Игрок A выбирает стратегию P_i так, чтобы максимизировать наименьший *ожидаемый выигрыш* по столбцам платежной матрицы, тогда как игрок B выбирает стратегию q_j с целью минимизировать наибольший *ожидаемый проигрыш* по строкам. Математически критерий минимакса при смешанных стратегиях может быть описан следующим образом. Игрок A выбирает стратегию P_i, дающую

$$\max_{P_i}\left\{\min\left(\sum_{i=1}^{m} a_{i1}\cdot P_i,\ \sum_{i=1}^{m} a_{i2}\cdot P_i,\ldots,\ \sum_{i=1}^{m} a_{in}\cdot P_i,\right)\right\}. \tag{9.43}$$

Игрок B выбирает стратегию q_j, дающую

$$\min_{q_j}\left\{\max\left(\sum_{j=1}^{n} a_{1j}\cdot q_j,\ \sum_{j=1}^{n} a_{2j}\cdot q_j,\ldots,\ \sum_{j=1}^{n} a_{mj}\cdot q_j\right)\right\}. \tag{9.44}$$

Когда стратегии P_i^0 и q_j^0 оптимальны, то выполняется строгое равенство между максиминным ожидаемым выигрышем и минимаксным проигрышем, а результирующее значение равно оптимальному (ожидаемому) значению игры.

Этот вывод следует из *теоремы фон Неймана о минимаксе*. Теорема утверждает, что выражения (9.43), (9.44) имеют одно и то же значение $M(P_0,Q_0)$, называемое *ценой игры*. Если P_i^0 и q_j^0 – оптимальные решения для обоих игроков, каждому элементу платежной матрицы a_{ij} соответствует вероятность $P_i^0\ .\ q_j^0$. Следовательно, оптимальное ожидаемое значение игры

$$M\left(P_0,Q_0\right)=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\cdot P_i^0\cdot q_j^0. \tag{9.45}$$

Цена игры заключена между нижней и верхней ценами, т. е.

$$a \le M(P_0,\ Q_0) \le \beta.$$

Решить конечную игру – это значит нужно найти векторы P и Q (оптимальные стратегии), удовлетворяющие теореме о минимаксе, а следовательно, получить величину ожидаемого платежа $M(P_0, Q_0)$ – цену игры.

Свойство оптимальности означает, что любое отступление одного из игроков от оптимальной стратегии (при условии, что второй игрок продолжает придерживаться своей оптимальной стратегии) при многократном повторении игры может только уменьшить его средний выигрыш (увеличить средний проигрыш).

Решение игры обладает одним важным свойством: если игрок A использует свою оптимальную стратегию, а игрок B смешивает свои стратегии в любых пропорциях, то средний выигрыш игрока A не уменьшается. Стратегии, которые смешиваются для получения оптимальной стратегии, будем называть *полезными*. Доказано, что у игры $m \times n$ существует такое оптимальное решение, что число полезных стратегий с каждой стороны не превосходит минимального из чисел m и n.

Известно несколько методов нахождения оптимальных стратегий в играх двух лиц с нулевой суммой. Рассмотрим один из методов – *метод линейного программирования для нахождения решения игр.*

Пусть игра $m \times n$ не имеет оптимального решения непосредственно в чистых стратегиях, т. е. отсутствует седловая точка ($\alpha \neq \beta$). Оптимальное решение необходимо искать в области смешанных стратегий. Предположим, что все m стратегий игрока A полезные. Определим вероятности их применения в смешанной оптимальной стратегии. Обозначим эти вероятности через P_i, $i = \overline{1,m}$, а цену игры – через M. Оптимальная смешанная стратегия игрока A определяется из условия (9.43):

$$\max\left\{\min\left(\sum_{i=1}^{m} a_{i1} \cdot P_i, \sum_{i=1}^{m} a_{i2} \cdot P_i, \ldots, \sum_{i=1}^{m} a_{in} \cdot P_i\right)\right\}.$$

Пусть

$$M = \min\left(\sum_{i=1}^{m} a_{i1} \cdot P_i, \sum_{i=1}^{m} a_{i2} \cdot P_i, \ldots, \sum_{i=1}^{m} a_{in} \cdot P_i\right). \tag{9.46}$$

Поскольку при оптимальной стратегии средний выигрыш не меньше M при любой стратегии противника, то справедлива система n неравенств:

$$\begin{array}{l} P_1 \cdot a_{11} + P_2 \cdot a_{21} + \ldots + P_m \cdot a_{m1} \geq M; \\ P_1 \cdot a_{12} + P_2 \cdot a_{22} + \ldots + P_m \cdot a_{m2} \geq M; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ P_1 \cdot a_{1n} + P_2 \cdot a_{2n} + \ldots + P_m \cdot a_{mn} \geq M, \end{array} \tag{9.47}$$

или

$$\sum_{i=1}^{m} a_{ij} \cdot P_i \geq M, \quad j=\overline{1,n}. \tag{9.48}$$

Тогда задача отыскания оптимальной смешанной стратегии игрока A может быть сформулирована в виде задачи линейного программирования.

Для этого необходимо максимизировать $Z = M$ при ограничениях

$$\begin{aligned} &\sum_{i=1}^{m} a_{ij} \cdot P_i \geq M, \quad j=\overline{1,n}; \\ &\sum_{i=1}^{m} P_i = 1, \quad P_i \geq 0, \quad i=\overline{1,m}. \end{aligned} \tag{9.49}$$

Введем новые неизвестные:

$$x_1 = \frac{P_1}{M}, \quad x_2 = \frac{P_2}{M}, \ldots, \quad x_m = \frac{P_m}{M}.$$

Чтобы исключить возможность деления на нуль, увеличим цену игры на положительное число C. Для этого достаточно ко всем элементам матрицы $\|a_{ij}\|$ прибавить одно и то же положительное число C, при этом все элементы a_{ij} сделать положительными. Следует отметить, что эта операция не меняет искомых оптимальных стратегий, поскольку

$$\sum_{i=1}^{m} P_i = 1, \text{ то } \sum_{i=1}^{m} x_i = \frac{1}{M}.$$

Разделим левую и правую части неравенств (9.48) и (9.49) на M, получим:

$$\begin{aligned} &a_{11}x_1 + a_{21}x_2 + \ldots + a_{m1}x_m \geq 1; \\ &a_{12}x_1 + a_{22}x_2 + \ldots + a_{m2}x_m \geq 1; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{mn}x_m \geq 1; \end{aligned} \tag{9.50}$$

$$x_1 + x_2 + \ldots + x_m = \frac{1}{M}. \tag{9.51}$$

В силу того что

$$\max M = \min \frac{1}{M} = \min\{x_1 + x_2 + \ldots + x_m\},$$

задача принимает вид

$$Z = x_1 + x_2 + ... + x_m \qquad (9.52)$$

при ограничениях

$$\sum_{i=1}^{m} a_{ij} \cdot x_i \geq 1, \quad j = \overline{1,n}. \qquad (9.53)$$

Для игрока B оптимальная стратегия определяется из условия (9.44):

$$\min_{qj}\left\{\max\left(\sum_{j=1}^{n} a_{1j}\,q_j, \sum_{j=1}^{n} a_{2j}\,q_j, ..., \sum_{j=1}^{n} a_{mj}\,q_j,\right)\right\}$$

при ограничениях

$$q_1 + q_2 + ... + q_n = 1. \qquad (9.54)$$

Эта задача записывается как симметричная двойственная задача линейного программирования к задаче игрока A(9.52), (9.53): максимизировать

$$L = Y_1 + Y_2 + ... + Y_n \qquad (9.55)$$

при ограничениях

$$\sum_{j=1}^{n} a_{ij} \cdot Y_j \leq 1; \qquad (9.56)$$

$$Y_j \geq 0, j = \overline{1, n},$$

где $L = \dfrac{1}{M}$, $Y_j = \dfrac{q_j}{M}$, $j = \overline{1,n}$.

Задачи игроков A и B решают обычным симплекс-методом.

Пример 9.10. Рассмотрим игру 3×4:

B

		1	2	3	4	α_i
	1	4	3	2	−5	−5
A	2	−2	5	−1	4	−2
	3	−3	2	−3	6	−3
	β_j	4	5	2	6	

Определим α_i и β_j,

где α_i – минимум в i-й строке;

β_j – максимум в j-м столбце.

Нижняя цена игры равна максимину $\alpha = -2$, верхняя цена игры равна минимаксу $\beta = 2$. Так как $\alpha \neq \beta$, то седловая точка игры отсутствует, задача должна решаться в смешанных стратегиях.

Нижняя цена игры – число отрицательное ($\alpha = -2$), поэтому, возможно, значение игры M не будет положительным. Число C, которое необходимо прибавить ко всем элементам матрицы, должно быть не меньше 2. Пусть $C = 6$. Тогда матрица принимает вид

		B 1	2	3	4
	1	10	9	8	1
A	2	4	11	5	10
	3	3	8	3	12

Задача игрока B записывается в форме задачи линейного программирования

$$L_{\max} = Y_1 + Y_2 + Y_3 + Y_4$$

при ограничениях:

$$\begin{cases} 10Y_1 + 9Y_2 + 8Y_3 + \quad Y_4 \le 1; \\ 4Y_1 + 11Y_2 + 5Y_3 + 10Y_4 \le 1; \\ 3Y_1 + 8Y_2 + 3Y_3 + 12\,Y_4 \le 1; \end{cases}$$
$$Y_j \ge 0, \quad j = \overline{1,4}.$$

Решая задачу симплекс-методом, получим:

$$L_{\max} = 0{,}16;\ Y_1 = 0;\ Y_2 = 0;\ Y_3 = 0{,}12;\ Y_4 = 0{,}04.$$

Таким образом, решением исходной задачи будет следующее:

$$M = \frac{1}{L} - C = \frac{1}{0{,}16} - 6 = 0{,}25;$$

$$q_1^0 = \frac{Y_1}{L} = \frac{0}{0{,}16} = 0;$$

$$q_2^0 = \frac{Y_2}{L} = \frac{0}{0{,}16} = 0;$$

$$q_3^0 = \frac{Y_3}{L} = \frac{0{,}12}{0{,}16} = 0{,}75;$$

$$q_4^0 = \frac{Y_4}{L} = \frac{0,04}{0,16} = 0,25.$$

В нашем примере первая и вторая стратегии игрока B бесполезны, так как $q_1^0 = q_2^0 = 0$. При случайном чередовании третьей и четвертой стратегий с относительными частотами $q_3^0 = 0,75$ и $q_4^0 = 0,25$ соответственно игроку B обеспечен средний выигрыш в размере $M = 0,25$.

Оптимальные стратегии игрока A получаются из решения двойственной задачи.

В настоящее время развитие теории игр вышло далеко за пределы рассмотренного нами простейшего случая парных игр с нулевой суммой. Однако ограниченный объем книги не позволяет дать представление о более сложных разделах современной теории игр, таких, как множественные коалиционные игры, дифференциальные игры и др.

Задачи

9.1. Для пяти проектов технических систем определены относительные единичные показатели технического совершенства конструкции и коэффициенты весомости единичных показателей. Численные значения единичных показателей и коэффициентов весомости приведены в следующей таблице:

Варианты технических систем	Относительные единичные показатели					
	сложности	веса	времени подготовки	автоматизации	мощности	унификации
I	1,0	0,088	1,0	1,0	0,72	0,614
II	0,72	1,0	0,8	0,78	0,81	0,420
III	0,658	0,358	0,765	0,782	0,525	0,915
IV	0,425	0,97	0,755	0,70	0,98	0,31
V	0,467	0,555	0,865	0,705	0,865	0,650
Коэффициенты веса	0,157	0,124	0,210	0,195	0,174	0,140

Проведите ранжирование проектов технических систем по комплексному критерию.

9.2. Одной из фирм требуется выбрать оптимальную стратегию по техническому обеспечению процесса управления производством. С помощью статистических данных и информации соответствующих заводов-изготовителей были определены локальные критерии функционирования необходимого оборудования. Исходные данные представлены в следующей таблице:

Варианты оборудования	Локальные критерии эффективности оборудования*			
	производительность, д. е.	стоимость оборудования, д. е.	объем памяти, у. е.	надежность, у. е.
I	100	5	5	8
II	150	6	8	5
III	120	4	6,5	6
IV	200	7	6	4
Коэффициенты веса	0,25	0,20	0,32	0,23

* Значения локальных критериев даны в условных единицах.

9.3. Для шести проектов транспортных устройств определены относительные единичные показатели технического совершенства конструкций. Численные значения единичных показателей и соответствующие весовые коэффициенты приведены в следующей таблице:

Варианты транспортных устройств	Относительные единичные показатели					
	скорости, K_1	прочности, K_2	перегрузки, K_3	устойчивости, K_4	металлоемкости, K_5	мощности, K_6
I	1,0	0,798	0,92	1,0	1,0	0,77
II	1,0	1,0	0,65	0,92	0,94	0,92
III	1,0	0,93	0,924	1,0	0,98	0,95
IV	0,87	0,96	0,91	0,915	0,99	0,85
V	0,85	0,97	1,0	0,90	0,7	0,82
VI	0,88	0,78	0,75	0,967	0,8	1,0
Коэффициенты веса	0,210	0,195	0,174	0,157	0,124	0,140

Проведите ранжировку проектов технических систем по комплексному критерию.

9.4. Абсолютные показатели качества двигателей различных вариантов приведены в следующей таблице:

Варианты двигателей	Показатели качества		
	мощность, л. с.	крутящий момент, кгс · м	масса, кг
1	180	67	850
2	176	70	1000
3	176	68	860
4	181	67	820
5	177	68	860
6	180	66	800
7	175	69	900
8	176	67	850
9	180	68,2	880
10	179	38,5	870
Коэффициенты веса	0,4	0,24	0,36

Найдите оптимальный вариант двигателя.

9.5. Показатели эффективности работы предприятий приведены в следующей таблице:

№ предприятия	Показатели эффективности работы предприятий				
	прибыль, д. е.	себестоимость единицы продукции, д. е.	доходы, д. е.	фондоотдача, у. е.	производительность, у. е.
I	30,0	40,0	20,0	0,2	300
II	25,0	20,0	30,0	0,3	200
III	40,0	45,0	54,0	0,1	250
IV	28,0	30,0	35,0	0,4	160
V	15,0	12,0	20,0	0,25	280
VI	50,0	30,0	40,0	0,21	120
Весовые коэффициенты	0,32	0,23	0,15	0,20	0,10

Выберите наиболее эффективно работающее предприятие.

9.6. Рассмотрим следующую платежную матрицу (матрицу доходов):

	S_1	S_2	S_3	S_4	S_5	S_6
R_1	15	12	1	−3	18	20
R_2	2	15	9	7	1	3
R_3	0	6	15	21	−2	5
R_4	8	20	12	3	0	4

Вероятности состояний природы $\{S_j\}$ неизвестны. Определите оптимальную стратегию R_i, используя критерии Лапласа и максимина. *Сравните* полученные решения.

9.7. *Сравните* решения в задаче 9.6 при использовании критериев Сэвиджа и Гурвица (положите $\alpha = 0{,}4$).

9.8. Один из пяти станков должен быть выбран для изготовления партии изделий, размер которой Q может принимать три значения: 150; 200; 350. Производственные затраты C_i для i станка задаются следующей формулой

$$C_i = P_i + c_i \,.\, Q.$$

Данные P_i и c_i приведены ниже в таблице:

Показатели	Модель станка				
	1	2	3	4	5
P_i	30	80	50	160	100
c_i	14	6	10	5	4

Решите задачу для каждого из следующих критериев Лапласа, Вальда, Сэвиджа и Гурвица (положите $\alpha = 0{,}6$). Полученные решения сравните.

9.9. При выборе стратегии R_j ($j = \overline{1,3}$) каждому возможному состоянию природы $S_i(i = \overline{1,4})$ соответствует один результат (исход) $V_{ji}(j = \overline{1,3};\ i = \overline{1,4})$. Элементы V_{ji}, являющиеся мерой потерь при принятии решения, приведены ниже в таблице (д. е.):

Стратегии	Состояние природы			
	S_1	S_2	S_3	S_4
R_1	2	6	5	8
R_2	3	9	1	4
R_3	5	1	6	2

Выберите оптимальное решение в соответствии с критериями Лапласа, Вальда, Сэвиджа и Гурвица (при $\alpha = 0,5$).

9.10. Намечается крупномасштабное производство легковых автомобилей. Имеются четыре варианта проекта автомобиля $R_j(j = \overline{1,4})$. Определена экономическая эффективность V_{ji} каждого проекта в зависимости от рентабельности производства. По истечении трех сроков $S_i(i = \overline{1,3})$ рассматриваются как некоторые состояния среды (природы). Значения экономической эффективности для различных проектов и состояний природы приведены в следующей таблице (д. е.):

Проекты	Состояние природы		
	S_1	S_2	S_3
R_1	20	25	15
R_2	25	24	10
R_3	15	28	12
R_4	9	30	20

Требуется выбрать лучший проект легкового автомобиля для производства, используя критерии Лапласа, Вальда, Сэвиджа и Гурвица при $\alpha = 0,1$. Сравните решения и сделайте выводы.

9.11. *Определите* тип электростанции, которую необходимо построить для удовлетворения энергетических потребностей комплекса крупных промышленных предприятий. Множество возможных стратегий в задаче включает следующие параметры:

R_1 – сооружается гидростанция;

R_2 – сооружается теплостанция;

R_3 – сооружается атомная станция.

Экономическая эффективность сооружения электростанции зависит от влияния случайных факторов, образующих множество состояний природы $S_i(i = \overline{1,5})$.

Результаты расчета экономической эффективности приведены в следующей таблице:

Тип станции	Состояние природы				
	S_1	S_2	S_3	S_4	S_5
R_1	40	70	30	25	45
R_2	60	50	45	20	30
R_3	50	30	40	35	60

9.12. Фирма рассматривает вопрос о строительстве станции технического обслуживания (СТО) автомобилей. Составлена смета расходов на строительство станции с различным количеством обслуживаемых автомобилей, а также рассчитан ожидаемый доход в зависимости от удовлетворения прогнозируемого спроса на предлагаемые услуги СТО (прогнозируемое количество обслуженных автомобилей в действительности). В зависимости от принятого решения – проектного количества обслуживаемых автомобилей в сутки (проект СТО) R_j и величины прогнозируемого спроса на услуги СТО – построена нижеследующая таблица ежегодных финансовых результатов (доход, д. е.):

Проекты СТО	Прогнозируемая величина удовлетворяемости спроса					
	0	10	20	30	40	50
20	–120	60	240	250	250	250
30	–160	15	190	380	390	390
40	–210	–30	150	330	500	500
50	–270	–80	100	280	470	680

Определите наилучший проект СТО с использованием критериев Лапласа, Вальда, Сэвиджа и Гурвица (при $\alpha = 0{,}5$).

9.13. Найдите седловую точку и значение игры для каждой из двух следующих игр. Платежные матрицы имеют вид:

$$\bar{A} = \left\|a_{ij}\right\| = \begin{pmatrix} 8 & 6 & 2 & 8 \\ 8 & 9 & 4 & 5 \\ 7 & 5 & 3 & 5 \end{pmatrix}; \quad \bar{B} = \left\|b_{ij}\right\| = \begin{pmatrix} 4 & -5 & -4 & 6 \\ -5 & -6 & -7 & -1 \\ 5 & 10 & -3 & -5 \\ 7 & 2 & -10 & 6 \end{pmatrix}.$$

9.14. *Определите* области значений x, для которых стратегии A_2 и B_2 будут оптимальными в следующих играх:

$$\bar{A} = \begin{pmatrix} 1 & 4 & 6 \\ 5 & x & 9 \\ 7 & 3 & 4 \end{pmatrix}; \quad \bar{B} = \begin{pmatrix} 2 & 4 & 5 \\ 10 & x & 6 \\ 4 & 8 & 3 \end{pmatrix}.$$

9.15. *Определите*, будут ли значения следующих игр больше, меньше или равны нулю:

$$\overline{A}=\begin{pmatrix} 2 & 10 & 5 & 0 \\ 3 & 4 & 9 & 6 \\ -5 & 3 & -2 & -4 \\ 8 & 5 & -3 & -5 \end{pmatrix};\quad \overline{B}=\begin{pmatrix} 4 & 8 & -2 & -3 \\ 5 & 9 & 1 & 0 \\ 5 & -8 & -3 & 5 \end{pmatrix};$$

$$\overline{C}=\begin{pmatrix} -2 & 8 & 5 & 7 \\ -3 & 9 & 3 & 5 \\ 4 & 2 & 0 & 6 \\ 8 & -1 & 9 & 3 \end{pmatrix};\quad \overline{D}=\begin{pmatrix} 4 & 8 & 1 \\ 0 & 2 & 5 \\ 5 & 1 & -3 \end{pmatrix}.$$

В задачах **9.16 – 9.30** решите игру с платежной матрицей.

9.16. $$\overline{A}=\begin{pmatrix} 2 & 1 & 3 \\ 4 & -1 & -2 \\ 1 & 0 & -3 \end{pmatrix}.$$

9.17. $$\overline{A}=\begin{pmatrix} 3 & 1 & 2 \\ 2 & 4 & -1 \\ 5 & 7 & 6 \end{pmatrix}.$$

9.18. $$\overline{A}=\begin{pmatrix} 2 & 3 & 6 & 5 \\ 1 & -2 & 7 & 3 \\ 5 & 4 & 3 & 0 \end{pmatrix}.$$

9.19. $$\overline{A}=\begin{pmatrix} 4 & 5 & 6 \\ 7 & 3 & 2 \\ 2 & 1 & 8 \end{pmatrix}.$$

9.20. $$\overline{A}=\begin{pmatrix} 2 & 3 & 1 & 2 \\ 3 & 5 & 2 & 3 \\ 1 & 2 & 4 & 1 \end{pmatrix}.$$

9.21. $$\overline{A}=\begin{pmatrix} 1 & -1 & 3 \\ 5 & 2 & -4 \\ 3 & 7 & 5 \end{pmatrix}.$$

9.22. $$\overline{A}=\begin{pmatrix} -1 & 1 & 1 \\ 2 & -2 & 2 \\ 3 & 3 & -3 \end{pmatrix}.$$

9.23. $$\overline{A}=\begin{pmatrix} 1 & 3 & -4 & 2 \\ -1 & 4 & 8 & 1 \\ 6 & -1 & 2 & 10 \end{pmatrix}.$$

9.24. $\bar{A} = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 2 & -2 \\ 0 & 3 & 5 \end{pmatrix}$.

9.25. $\bar{A} = \begin{pmatrix} 1 & -1 & 3 \\ 5 & 2 & -4 \\ 3 & 4 & 0 \end{pmatrix}$.

9.26. $\bar{A} = \begin{pmatrix} 2 & 6 & 4 & 5 \\ 5 & 3 & 6 & 2 \\ 7 & 2 & 1 & 3 \end{pmatrix}$.

9.27. $\bar{A} = \begin{pmatrix} 2 & 4 & 1 & 5 \\ 1 & -1 & 3 & 2 \\ 5 & 2 & -4 & 0 \end{pmatrix}$.

9.28. $\bar{A} = \begin{pmatrix} 2 & 1 & 3 & 0 \\ 2 & 4 & -1 & 5 \\ 5 & 7 & -4 & 3 \end{pmatrix}$.

9.29. $\bar{A} = \begin{pmatrix} 2 & 3 & -1 & 4 \\ 3 & 2 & 4 & 1 \\ 4 & -1 & 0 & -2 \end{pmatrix}$.

9.30. $\bar{A} = \begin{pmatrix} 3 & 4 & 2 & 1 \\ 5 & 2 & 0 & 3 \\ 0 & 3 & 5 & 4 \end{pmatrix}$.

Глава 10
Нелинейное программирование

10.1. Геометрическая интерпретация задач нелинейного программирования

Если целевая функция или система ограничений (или и та и другая) содержит выражения, не линейные относительно искомых величин, то имеем задачу нелинейного программирования.

Общая формулировка задачи имеет вид:

$$Z_{\max,\min} = F(x_1, x_2, ..., x_n) \tag{10.1}$$

при ограничениях

$$f_i(x_1, x_2, ..., x_n) \geq 0,\ x_j \geq 0,\ i = \overline{1, m},\ j = \overline{1, n}.$$

Если все неравенства преобразовать в уравнения, то получим классическую задачу на условный экстремум, которая может быть решена методом множителей Лагранжа. Практически системы получаются трудно разрешимыми, поэтому ограничения преобразуют в неравенства, и задачу решают методами математического программирования. Существующие методы позволяют решать узкий класс задач.

С помощью большинства вычислительных методов можно найти точку локального оптимума, но нельзя установить, является ли она точкой глобального (абсолютного) оптимума или нет.

Если в задачах линейного программирования точка экстремума является вершиной многогранника, то в задачах нелинейного программирования она может лежать в вершине многогранника, на ребре (грани) или внутри области. Если задача содержит нелинейные ограничения, то область допустимых решений не является выпуклой и кроме глобального оптимума могут существовать точки локального оптимума.

Рассмотрим несколько примеров решения задач графическим методом.

Пример 10.1.

$$Z_{\max,\min} = 2(x-5)^2 + (y-7)^2$$

$$x + 2y \leq 12;$$
$$x + y \leq 9;$$
$$x \geq 0;\ y \geq 0.$$

Решение

Построим область допустимых решений (рис. 10.1):

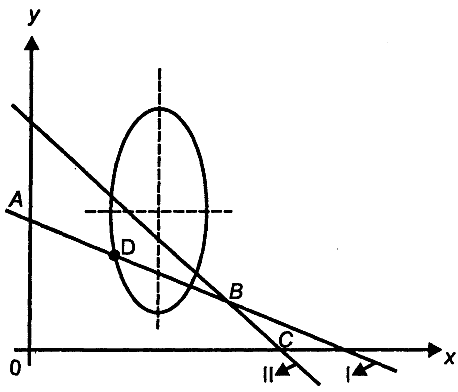


Рис. 10.1. Область допустимых решений

На этом же графике построим одну из семейства целевых функций, для этого преобразуем Z в каноническую форму $\frac{(x-5)^2}{Z/2}+\frac{(y-7)^2}{Z}=1$. Получим уравнение эллипса с полуосями $a=\sqrt{Z/2}$; $b=\sqrt{Z}$. Полагая $Z = 16$, имеем $b = 4$; $a \approx 2{,}8$. Согласно чертежу максимальное значение Z достигается в точке (0;0) $Z_{max} = 2 \cdot 25 + 49 = 99$; Z_{min} достигается в точке D, в которой эллипс касается области 0ABC. Для нахождения координат точки D приравняем угловые коэффициенты первой прямой и целевой функции: $y = -1/2\,x - 6$; $0 = 4(x - 5) + 2(y - 7)y^1$; $y^1=-\frac{2(x-5)}{y-7}$;

$-1/2=-\frac{2(x-5)}{y-7}$; упрощая, получим: $4(x - 5) = y - 7$.

Решая совместно систему

$$\begin{cases}4(x-5)=y-7\\ x+2y=12,\end{cases}$$

находим координаты точки

$$D\ (38/9;\ 35/9).$$

$$Z_{min} = 2(38/9 - 5)^2 + (35/9 - 7)^2 \approx 11.$$

Пример 10.2.

$$Z_{max,\,min} = x_1^{\,2} + x_2^{\,2}$$

$$\begin{array}{ll} x_1 + x_2 \le 4; & \\ x_1 + x_2 \ge 5; & x_1 \ge 0; \\ x_1 \le 7; & \\ x_2 \ge 6; & x_2 \ge 0. \end{array}$$

Решение

В этом случае область допустимых решений не является выпуклой и состоит из двух отдельных частей (рис. 10.2).

Минимальное значение функции $Z = 11$ достигается в точках A (1; 4) и L (4; 1). Функция Z имеет два локальных максимума: в точках D (2/3; 6); $Z = 328/9$ и в точке M (7; 4/7); $Z=\frac{2417}{49}$. Точка M является точкой глобального максимума.

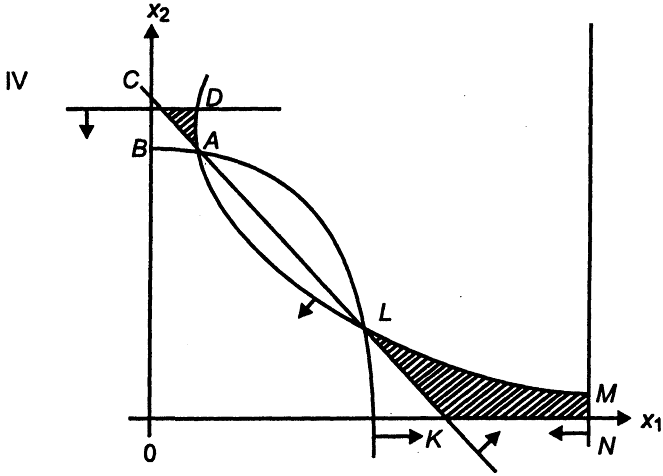


Рис. 10.2. Область допустимых решений

10.2. Метод множителей Лагранжа

Пусть задана задача математического программирования: максимизировать функцию

$$Z = f(x_1, x_2, ..., x_n) = 0 \tag{10.2}$$

при ограничениях

$$g_i = (x_1, x_2, ..., x_n) = 0; \; i = 1, \overline{m}. \tag{10.3}$$

Ограничения в задаче заданы уравнениями, поэтому для ее решения можно воспользоваться классическим методом отыскания условного экстремума функций нескольких переменных. При этом полагаем, что функции (10.2) и (10.3.) непрерывны вместе со своими первыми частными производными.

Для решения задачи составим функцию

$$\begin{gathered} F(x_1, x_2, ..., x_n, \lambda_1, ..., \lambda_m) = \\ = f(x_1, x_2, ..., x_n) + \Sigma f\lambda_i g_i (x_1, x_2, ..., x_n). \end{gathered} \tag{10.4}$$

Определим частные производные $\dfrac{\partial F}{\partial x_j}$ и $\dfrac{\partial F}{\partial \lambda_i}$ и приравняем их нулю. В результате получим систему уравнений

$$\begin{cases} \dfrac{\partial F}{\partial x_j} = \dfrac{\partial f}{\partial x_j} + \sum \lambda_i \dfrac{\partial g_i}{\partial x_j} = 0; \quad j = \overline{1,n}; \\ \dfrac{\partial F}{\partial \lambda_j} = g_i(x_1, x_2, ..., x_n) = 0; \quad i = \overline{1,m}. \end{cases} \tag{10.5}$$

Функция (10.4) называется *функцией Лагранжа*, а числа – *множителями Лагранжа*. Если функция (10.2) в точке $x^0 = (x_1^0, x_2^0, ..., x_n^0)$ имеет экстремум, то существует такой вектор $\lambda^0 = (\lambda_1^0, \lambda_2^0, ..., \lambda_m^0)$, что точка $(x_1^0, x_2^0, ..., x_n^0, \lambda_1^0, \lambda_2^0, ..., \lambda_m^0)$ является решением системы (10.5). Следовательно, решая систему (10.5), получим множество точек, в которых функция Z имеет экстремальные значения. При этом неизвестен способ определения точек глобального минимума или максимума. Однако, если решения системы найдены, то для определения глобального максимума (минимума) достаточно найти значения функции Z в соответствующих точках. Метод множителей Лагранжа имеет ограниченное применение, так как система (10.5), как правило, имеет несколько решений. Рассмотрим пример применения метода множителей Лагранжа.

Пример 10.3. Найти точку условного экстремума функции

$$Z = x_1x_2 + x_2x_3$$

при ограничениях

$$x_1 + x_2 = 2;$$
$$x_2 + x_3 = 2.$$

Решение

Составим функцию Лагранжа

$$F = (x_1, x_2, x_3, \lambda_1, \lambda_2) =$$
$$= x_1x_2 + x_2x_3 + \lambda_1(x_1 + x_2 - 2) + \lambda_2(x_2 + x_3 - 2)$$

и продифференцируем ее по всем переменным: x_1, x_2, x_3, λ_1, λ_2. Приравнивая производные к нулю, получим систему уравнений

$$\begin{array}{llllll} & x_2 & & +\lambda_1 & & = 0; \\ x_1 + & & x_3 & +\lambda_1 & +\lambda_2 & = 0; \\ & x_2 & & & +\lambda_2 & = 0; \\ x_1 + & x_2 & & & -2 & = 0; \\ & x_2 + & x_3 & & -2 & = 0; \end{array}$$

Из первого и третьего уравнения следует, что $\lambda_1 = \lambda_2 = -x_2$, тогда

$$x_1 - 2x_2 + x_3 = 0;$$
$$x_1 + x_2 = 2;$$
$$x_2 + x_3 = 2.$$

Решая систему, получим $x_1 = x_2 = x_3 = 1$;

$$Z = 2.$$

10.3. Градиентный метод

Возьмем для простоты целевую функцию с двумя независимыми переменными:

$$Z = F(x_1, x_2). \tag{10.6}$$

Изобразим линии уровня этой функции на рис. 10.3.

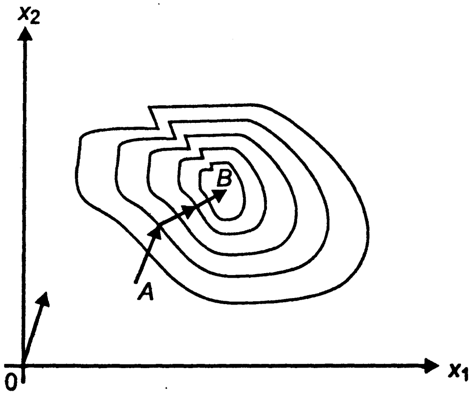


Рис. 10.3. Линии уровня функции Z

Выберем некоторую точку A и вычислим для нее значение Z. Если сдвинуться из точки A на одно и то же расстояние в разных направлениях, то функционал Z изменится на разную величину.

Чтобы быстрее достичь экстремума, надо двигаться в направлении наибольшего изменения Z.

Найдем частные производные функции Z и вычислим их значения в точке A. Получим

$$\left(\frac{\partial Z}{\partial x_1}\right)_A, \left(\frac{\partial Z}{\partial x_2}\right)_A.$$

Примем найденные числа за координаты некоторого вектора и построим этот вектор на рис. 10.3. Оказывается, что если из точки A перемещаться вдоль прямой, параллельной указанному вектору, то функция Z будет изменяться быстрее всего: в направлении вектора увеличиваться, в обратном – уменьшаться.

Вектор, координатами которого служат значения частных производных функции $Z = F(x_1, x_2)$, называется *градиентом этой функции* и обозначается

$$\mathrm{grad}Z\left\{\frac{\partial Z}{\partial x_1}_A, \frac{\partial Z}{\partial x_2}_A\right\}\ldots$$

В каждой точке градиент будет свой. Он показывает направление наибольшего возрастания функции в данной точке, которое всегда перпендикулярно проходящей через нее линии уровня.

На этом свойстве и основан градиентный метод решения задач нелинейного программирования. Выбрав в области допустимых решений первоначальную точку A и убедившись, что она не является оптимальной, дальнейшее движение осуществляем целенаправленно. Для этого вычисляем в выбранной точке координаты градиента целевой функции

$$\left(\frac{\partial Z}{\partial x_1}\right)_A, \left(\frac{\partial Z}{\partial x_2}\right)_A\ldots$$

Составляем параметрические уравнения прямой, проходящей через точку А параллельно градиенту:

$$\begin{aligned}
x_1 &= x_1^{(A)} + \left(\frac{\partial Z}{\partial x_1}\right)_A t;\\
x_2 &= x_2^{(A)} + \left(\frac{\partial Z}{\partial x_2}\right)_A t;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
x_n &= x_n^{(A)} + \left(\frac{\partial Z}{\partial x_n}\right)_A t.
\end{aligned} \tag{10.7}$$

В случае нахождения максимума перемещаемся по прямой в направлении градиента ($t > 0$), для минимизации функции движемся в противоположном направлении ($t < 0$).

При этом возможны два способа перемещения.

1. Из начальной точки движемся в направлении градиента до тех пор, пока целевая функция перестанет возрастать. Это движение осуществляется постепенным увеличением параметра t, причем, для каждого значения находится величина функционала, а также проверяется допустимость полученного решения. Достигнув наивысшей в данном направлении точки (точка B на рис. 10.3), снова определяем градиент и по новому направлению перемещаемся на максимально возможную величину и т. д.

Такой порядок расчетов соответствует наискорейшему подъему и применяется тогда, когда значение функционала вычислить проще, чем его градиент.

В задачах на минимум движение происходит в сторону убывания функции, и такой метод носит название *наискорейшего спуска*.

2. Из начальной точки по линии, параллельной градиенту, перемещаемся на некоторый интервал h, характеризуемый значением параметра t. Во второй точке определяем свой градиент и по новому направлению снова перемещаемся на величину h. Получаем третью точку и т. д. Так небольшими переходами, корректируя всякий раз направление движения, достигаем точки, соответствующей оптимальному значению функционала (максимум, минимум). Признаком достижения оптимума является получение нулевых составляющих градиента.

При перемещении по направлению градиента может оказаться, что найденная точка лежит вне области (т. е. на данном шаге пересечена граница области). В этом случае экстремум, как правило, достигается на границе, и его нахождение усложняется.

Возврат в область простым уменьшением последнего шага неэффективен, поскольку нет уверенности, что экстремальная точка находится в пересечении с границей именно этого направления. Необходим возврат с одновременным смещением в сторону оптимума.

Рассмотрим рис. 10.4, где область допустимых решений показана жирными линиями, а линия уровня функционала – тонкими.

Из точки A перемещаемся по нормали к линии уровня и попадаем за границу области в точку B. Так как оптимум достигается в точке M, возврат в область по линии BA бесполезен.

Рассмотрим подробно участок нарушенной границы, представленный на рис. 10.5.

Направление первоначального движения AB перпендикулярно касательной в точке A. При возвращении из точки B в заданную область будем перемещаться по прямой BQ, которая перпендикулярна к касательной, проведенной через точку Q. Иными словами, BC должно быть нормалью к границе y_i. Новое направление BC отклонится от прежнего AB на угол α. Отклонение будет в ту сторону, где

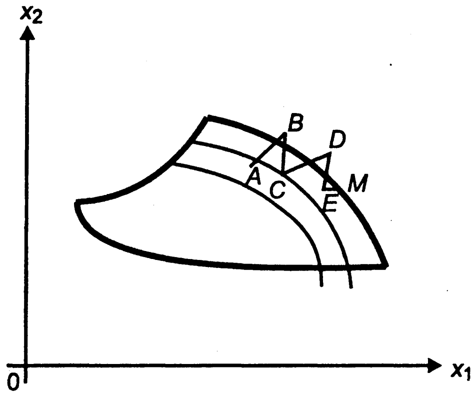


Рис. 10.4. Область допустимых решений и линии уровня функционала

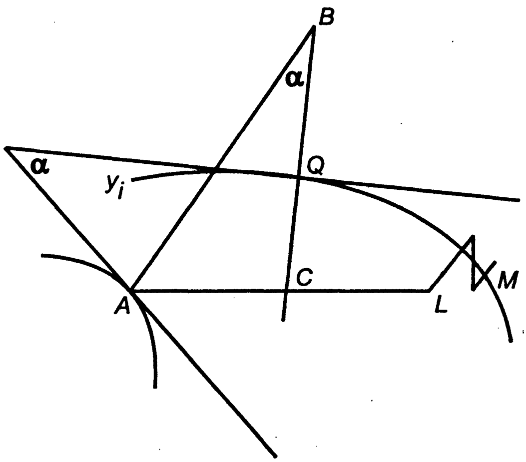


Рис. 10.5. Участок области допустимых решений

находится точка, соответствующая оптимальному плану. Так как точка B находится вблизи границы, за направление возврата в область можно принять направление градиента функции y_i в точке B.

Придя в точку C, движемся по направлению grad Z, попадаем в точку D и т. д. (рис. 10.4). Признаком получения оптимума является коллинеарность векторов grad Z и grad y_i. Коллинеарность векторов проверяется пропорциональностью соответствующих координат.

Для ускорения сходимости процесса, после одного-двух зигзагов, делают шаг в направлении AC, т. е. почти параллельно границе, при этом параметр t берут больше обычного. Из новой точки продолжается опять зигзагообразное движение к оптимуму.

Пример 10.4. Найти максимум функции $Z = x_1^2 + 4x_2$ при условии

$$4x_1^2 + 9x_2^2 \le 36; \quad x_1 \ge 0; x_2 \ge 0.$$

Решение

Уравнение границы перепишем в другом виде

$$y = 36 - 4x_1^2 - 9x_2^2 \ge 0.$$

Возьмем произвольную точку A (2; 1): так как $y_A = 11 > 0$, следовательно, она принадлежит области допустимых решений

$$Z_A = 2^2 + 4 \cdot 1 = 8.$$

Частные производные функции Z имеют вид:

$$\frac{\partial Z}{\partial x_1} = 2x_1; \quad \frac{\partial Z}{\partial x_2} = 4;$$

в точке A:

$$\left(\frac{\partial Z}{\partial x_1}\right)_A = 4; \quad \left(\frac{\partial Z}{\partial x_2}\right)_A = 4.$$

Следовательно, grad Z_A (4; 4) и в точке A функции Z не достигает максимума. Запишем уравнение прямой, проходящей через точку A, параллельную градиенту:

$$\begin{cases} x_1 = 2 + 4t; \\ x_2 = 1 + 4t. \end{cases}$$

Полагая $t = 0{,}05$, получим

$$\begin{cases} x_1 = 2 + 4 \cdot 0{,}05 = 2{,}2; \\ x_2 = 1 + 4 \cdot 0{,}05 = 1{,}2. \end{cases}$$

Точка B (2,2; 1,2) лежит в заданной области, так как

$$y_B = 36 - 4 \cdot 2{,}2^2 - 9 \cdot 1{,}2^2 = 3{,}68 > 0;$$

$$Z_B = 2{,}2^2 + 4 \cdot 1{,}2 = 9{,}64.$$

Для перехода в новую точку находим координаты градиента для точки B:

$$\left(\frac{\partial Z}{\partial x_1}\right)_B = 2{,}2 = 4{,}4; \quad \left(\frac{\partial Z}{\partial x_2}\right)_B = 4.$$

Составляем уравнение линии перемещения из точки B:

$$x_1 = 2{,}2 + 4{,}4t;$$
$$x_2 = 1{,}2 + 4t.$$

Придавая параметру t прежнее значение, получаем:

$$x_1 = 2{,}42;$$
$$x_2 = 1{,}40,$$

попадаем в точку C (2,42; 1,40), так как

$$y_c = 36 - 4 \cdot 2{,}42^2 - 9 \cdot 1{,}40^2 = -5{,}04 < 0.$$

Следовательно, точка C лежит вне области допустимых решений. Необходимо вернуться в область и сдвинуться в сторону оптимума. Находим градиент граничной функции в точке C:

$$\frac{\partial y}{\partial x_1} = 8x_1; \quad \frac{\partial y}{\partial x_2} = 18x_2;$$

$$\left(\frac{\partial y}{\partial x_1}\right)_C = -19{,}4; \quad \left(\frac{\partial y}{\partial x_2}\right)_C = -25{,}2;$$

$$\text{grad}\, y_C(-19{,}4;\ -25{,}2).$$

Движение из точки C будем осуществлять вдоль линии

$$x_1 = 2{,}42 - 19{,}4t;$$
$$x_2 = 1{,}40 - 25{,}2t.$$

Учитывая большую абсолютную величину координат градиента, уменьшим параметр t, допустим $t = 0{,}01$.

Получим точку D (2,23; 1,15).

Проверяем принадлежность точки D заданной области:

$$y_D = 36 - 4 \cdot 2{,}23^2 - 9 \cdot 1{,}15^2 = 4{,}22 > 0.$$

Условие выполняется, решение допустимо:

$$Z_D = 2{,}23^2 + 4 \cdot 1{,}15 = 9{,}57,$$

которое, по сравнению с точкой B, несколько уменьшилось. В найденных соседних точках B и D значения Z близки между собой, это указывает на смещение оптимальной точки вдоль линии уровня целевой функции. Для ускорения сходимости процесса сделаем шаг в направлении BD. Координатами вектора, параллельного BD, будут разности соответствующих координат точек B и D:

$$x_1^{(\overline{BD})} = x_1^{(D)} - x_1^{(B)} = 2{,}23 - 2{,}20 = 0{,}03;$$

$$x_2^{(\overline{BD})} = x_2^{(D)} - x_2^{(B)} = 1{,}15 - 1{,}20 = -0{,}05.$$

Уравнение прямой, проходящей через точку D в направлении BD, имеет вид:

$$x_1 = 2{,}23 + 0{,}03t;$$
$$x_2 = 1{,}15 - 0{,}05t.$$

Параметр t надо брать большим, чтобы сдвиг вдоль границы был существенным. Полагая $t = 5$, получим точку E (2,38; 0,90)

$$y_E = 6{,}09 > 0; \quad Z_E = 2{,}38^2 + 4 \cdot 0{,}90 = 9{,}26.$$

Точка E лежит в заданной области. Из точки E переместимся параллельно градиенту Z. Для этого находим

$$\left(\frac{\partial Z}{\partial x_1}\right)_E = 4{,}76; \quad \left(\frac{\partial Z}{\partial x_2}\right)_E = 4; \quad \operatorname{grad} Z_E(4{,}76;\ 4);$$

$$x_1 = 2{,}38 + 4{,}76t;$$
$$x_2 = 0{,}90 + 4t.$$

При $t = 0{,}05$ переходим в точку F (2,62; 1,10). Точка F лежит за границей области, так как

$$y_E = 36 - 4 \cdot 2{,}62^2 - 9 \cdot 1{,}10^2 = -2{,}31 < 0.$$

Возврат в допустимую область необходимо вести по направлению вектора

$$Y_F: \left(\frac{\partial y}{\partial x_1}\right)_F = -21{,}0; \quad \left(\frac{\partial y}{\partial x_2}\right)_F = -19{,}8;$$

Прежде чем делать переход, вычислим отношения одноименных координат grad Z_E и grad y_F:

$$\frac{4,76}{-21,0} = -0,226; \quad \frac{4}{-19,8} = -0,202.$$

Отношения близки друг к другу, т. е. векторы grad Z_E и grad y_F практически коллинеарны и противоположно направлены. Движение по ним будет происходить взад-вперед через границу. Поэтому, чтобы попасть в область ближе к границе, положим $t = 0,005$ и перейдем в точку G (2,52; 1,00)

$$y_G = 36 - 4 \cdot 2,52^2 - 9 \cdot 1,00^2 = 1,6 > 0.$$

Следовательно, точка лежит в заданной области:

$$Z_G = 2,52^2 + 4 \cdot 1,00 = 10,35.$$

Полученный результат $x_1 = 2,52$; $x_2 = 1,00$; $Z = 10,35$ можно считать оптимальным: его можно уточнить дальнейшими шагами до полной коллинеарности градиентов целевой и граничной функций.

Движение точки во время поиска можно представить на рис. 10.6.

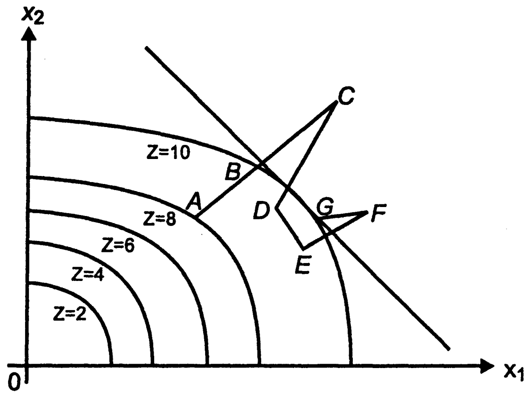


Рис. 10.6. Пример нахождения оптимального решения

10.4. Многоцелевые задачи линейного программирования

Известно, что экономическая эффективность производства количественно измеряется системой экономических показателей. Так, в промышленном производстве важными показателями эффективности являются: рост производительности труда, прибыль, рентабельность и др. Максимальное значение одного из показателей еще не означает, что то или иное предприятие работает лучше. Только система показателей может характеризовать эффективность всего производства.

Задачи, решаемые с учетом множества показателей или критериев, носят название *многоцелевых*. Планы задач, полученные по разным критериям, будут отличаться один от другого. Можно для каждого плана определить все интересующие нас показатели, и после их сравнения остановиться на том, который позволяет получить достаточно высокую эффективность всего производства. Если же из полученных планов нельзя выбрать план, который обеспечивал бы необходимую эффективность, можно воспользоваться специальным методом, так называемым *методом уступок*.

Рассмотрим ***первый прием*** на конкретной задаче.

С помощью двух операций производится два вида продукции – A и B. При производстве продукта A на каждой операции затрачивается 3 часа, а при производстве продукта B соответственно 4 и 5 часов. Общее наличие времени для первой и второй операции 18 и 21 час. Стоимость единицы продукции A – 3 руб., а продукции B – 8 руб.

Провести анализ работы предприятия с учетом различных целей:

а) максимум прибыли;

б) максимум выпуска продукции;

в) наилучшее использование оборудования.

Обозначим через x_1 и x_2 количество продукции видов A и B, выпущенных предприятием.

Используя введенные переменные, ограничения можно записать

$$3x_1 + 4x_2 \le 18;$$
$$3x_1 + 5x_2 \le 21;$$
$$x_1 \ge 0;\ \ x_2 \ge 0.$$

Полученные ограничения можно записать в виде уравнений

$$3x_1 + 4x_2 + x_3 = 18;$$
$$3x_1 + 5x_2 + x_4 = 21,$$

где x_3 и x_4 выражают неиспользованное время на первой и второй операциях.

Целевые функции согласно трем сформулированным целям можно представить в виде:

а) $Z_{max} = 3x_1 + 8x_2 + 0 \cdot x_3 + 0 \cdot x_4$;

б) $Z_{max} = x_1 + x_2 + 0 \cdot x_3 + 0 \cdot x_4$;

в) $Z_{min} = 0 \cdot x_1 + 0 \cdot x_2 + x_3 + x_4$.

Решение задачи «а» представлено в табл. 10.1.

Таблица 10.1

Решение задачи по а) критерию

			0	3	8	0	0	
				x_1	x_2	x_3	x_4	Σ
	0	x_3	18	3	4	1	0	26
	0	x_4	21	3	$\boxed{5}$	0	1	30
			0	–3	–8	0	0	–11
	0	x_3	6/5	3/5	0	1	–4/5	2
	8	x_2	21/5	3/5	1	0	1/5	6
			168/5	9/5	0	0	8/5	37

В индексной строке нет отрицательных элементов, следовательно, получено решение (план), соответствующее максимальной прибыли:

$$X = (0;\ 21/5;\ 6/5;\ 0);\ Z_{max} = 33{,}6.$$

Анализ решения

1. Продукция вида *А* не производится $x_1 = 0$, а продукция вида *В* производится 21/6 единиц.

2. $x_3 = 6/5$ указывает на то, что по этому плану производственное оборудование при выполнении первой операции простаивает 6/5 ч.

3. Максимальная прибыль равна 33,6 руб.

4. Элемент 9/5 в индексной строке указывает на то, что дополнительное производство одной единицы продукции *А* уменьшает прибыль на 9/5 руб.

5. Значение индекса 3/5 указывает на то, что дополнительный час работы оборудования при выполнении второй операции увеличивает прибыль на 8/5 руб. Соответственно сокращение времени работы на один час уменьшает прибыль на 8/5 руб.

Решение задачи «б» симплекс-методом представлено в табл. 10.2.

Таблица 10.2

Решение задачи по б) критерию

		0	1	1	0	0	
			x_1	x_2	x_3	x_4	Σ
0	x_3	18	$\boxed{3}$	4	1	0	26
0	x_4	21	3	5	0	1	30
		0	–1	–1	0	0	–2
1	x_1	6	1	4/3	1/3	0	29/3
0	x_4	3	3/05	1	–1	1	4
		6	9/05	1/3	1/3	0	20/3

Анализ решения

1. Производится шесть единиц продукции A $x_1 = 6$, продукция вида B не производится $x_2 = 0$.

2. $x_3 = 0$ производственные мощности при выполнении первой операции используются полностью. $x_4 = 3$ – производственные мощности при выполнении второй операции не использованы в течение 3 часов.

3. 1/3 в столбце x_2 показывает, насколько сократился объем продукции, если изготовить единицу продукции вида B. 1/3 в столбце x_3 означает, что увеличение времени работы на первой операции (на 1 час) позволит выпустить добавочно 1/3 изделия вида.

4. Общая прибыль равна $6 \cdot 3 = 18$ руб.

Решение задачи «в» симплекс-методом представлено в табл. 10.3.

Анализ решения

1. Нуль в столбце констант показывает, что имеющиеся мощности используются полностью. (Время простоя оборудования равно 0).

2. Продукции вида A производятся две единицы ($x_1 = 2$), а продукции вида B три единицы ($x_2 = 3$).

3. Прибыль при осуществлении программы максимального использования оборудования равна $3 \cdot 2 + 3 \cdot 8 = 30$ руб.

Для сравнения полученных решений сведем результаты в таблицу (табл. 10.4).

Если оценивать работу предприятия по трем показателям (табл. 10.4), то вполне очевидно, что третий вариант наилучший.

Рассмотрим ***второй способ*** (метод уступок), вначале сформулируем алгоритм метода.

Таблица 10.3

Решение задачи по в) критерию

		0	0	0	−1	−1	
			x_1	x_2	x_3	x_4	Σ
−1	x_3	18	3	4	1	0	26
−1	x_4	21	3	$\boxed{5}$	0	1	30
		−39	−6	−9	0	0	−54
−1	x_3	6/5	3/5	0	1	−4/5	2
0	x_2	21/5	3/5	1	0	1/5	6
		−6/5	−3/5	0	0	9/5	0
0	x_1	2	1	0	5/3	−4/3	10/3
0	x_2	3	0	1	−1	1	4
		0	0	0	1	1	2

$$x = (2;\ 3;\ 0;\ 0),$$
$$Z_{\min} = 0$$

Таблица 10.4

Решение задачи по трем критериям

Критерий	Количество изделий, ед.		Общее количество, ед.	Прибыль	Неиспользованная мощность		
	A	B			I	II	всего
Максимум прибыли	0	4,2	4,2	33,6	1,2	0	1,2
Максимум продукции	6	0	6	18,0	0	3	3
Минимум времени простоя оборудования	2	3	5	30,0	0	0	0

1. Расположить критерии по их значимости (наиболее важный считается первым).

2. Решить задачу по первому критерию, $Z_1 = Z_1^*$, т. е. отыскать экстремальное значение Z_1^* целевой функции Z_1.

3. Сделать уступку по первому критерию, иными словами, уменьшить величину Z_1 до значения $Z_1 = K_1 Z_1^*$; $0 < K_1 < 1$.

4. В задачу ввести дополнительное ограничение $Z_1 \geq K_1Z_1^*$.

5. Решить задачу по второму критерию $Z_2 = Z_2^*$.

6. Обратиться к пункту 3, сделать уступку для второго критерия $Z_2 = K_2Z_2^*$; $0 < K_2 < 1$.

7. Ввести в задачу дополнительное ограничение $Z_2 \geq K_2Z_2^*$.

8. Новую задачу, уже с двумя дополнительными ограничениями, решить по третьему критерию и т. д.

9. Процесс итерации заканчивается, когда решение будет получено по всем критериям.

Проведем геометрическую интерпретацию этого метода.

Решить задачу по двум критериям, считая первый наиболее предпочтительным. Его отклонение от максимального составляет не более 10%.

$$Z_{1\max} = x_1 + 2x_2;\ Z_{2\min} = x_1 + x_2;$$
$$x_1 + 2x_2 \geq 6;$$
$$x_1 \leq 4;$$
$$x_2 \leq 5;$$
$$x_1 \geq 0;\ x_2 \geq 0.$$

Решение

Используя первый критерий, находим область допустимых решений *АВСД*. Максимальное значение целевой функции достигается в точке *С* (4,5), представленной на графике (рис. 10.7).

Делаем уступку на 10% $f_1 = 0{,}9 \cdot 14 = 12{,}6$; получаем дополнительное ограничение $x_1 + 2x_2 > 12{,}6$ (на рис. 10.7 – прямая). С учетом дополнительного ограничения областью решения является треугольник ЕСК. Минимальное значение целевой функции Z_2 достигается в точке *Е* (2,6; 5).

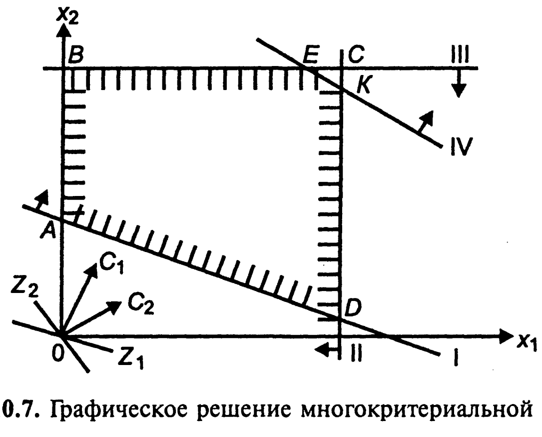


Рис. 10.7. Графическое решение многокритериальной задачи

Таким образом, план в точке E (2,6; 5) будет наиболее эффективным при данных условиях

$$x = (2{,}6;5),\ Z_{1\max} = 12{,}6;\ Z_{2\min} = 7{,}6.$$

Задачи

Используя графический метод, найти наибольшее и наименьшее значения следующих функций:

10.1. $Z_{\min} = (x_1 - 6)^2 + (x_2 - 2)^2$
$x_1 + 2x_2 \le 8$
$3x_1 + x_2 \le 15$
$x_1 + x_2 \ge 1$
$x_1 \ge 0;\ x_2 \ge 0.$

10.2. $Z_{\min} = 2(x_1 - 7)^2 + 4(x_2 - 3)^2$
$x_1 + 2x_2 \ge 2$
$x_1 + x_2 \le 6$
$2x_1 + x_2 \le 11$
$x_1 \ge 0;\ x_2 \ge 0.$

10.3. $Z_{\max} = 2x_1 - 0{,}2x_1^2 + 3x_2 - 0{,}2x_2^2$
$2x_1 + 3x_2 \le 13$
$2x_1 + x_2 \le 10$
$x_1 \ge 0;\ x_2 \ge 0.$

10.4. $Z_{\max} = 3x_1 - 0{,}3x_1^2 + 6x_2 - 0{,}3x_2^2$
$9x_1 + 8x_2 \le 72$
$x_1 + 2x_2 \le 10$
$x_1 \ge 0;\ x_2 \ge 0.$

Решить градиентным методом, начиная процесс оптимизационного поиска с указанной точки $\overline{x}_0$ и сопровождая решение графиком:

10.5. $Z_{\max} = 3x_1 - 0{,}2x_1^2 + x_2 - 0{,}2x_2^2$
$x_1 + x_2 \le 7$
$x_1 + 2x_2 \le 10$
$x_1 \ge 0;\ x_2 \ge 0$
$\overline{x}_0 = (1;\ 2).$

10.6. $Z_{\max} = 2x_1 - 0{,}1x_1^2 + 3x_2 - 0{,}1x_2^2$
$5x_1 + 2x_2 \ge 30$
$3x_1 + 4x_2 \le 24$
$x_1 \ge 0;\ x_2 \ge 0$
$\overline{x}_0 = (3;\ 1).$

10.7. Обработка деталей А, В и С может производиться на трех станках (I, II, III). В таблице указаны нормы затрат времени на обработку станком соответствующей детали, цена одной детали (в руб.), стоимость 1 часа работы станка и предельное время работы станка.

Детали / Станки	Норма времени			Стоимость	Предельное время
	A	B	C		
I	0,2	0,1	0,05	30	40
II	0,6	0,3	0,2	10	60
III	0,2	0,1	0,4	20	30
Цена	10	16	12		

Составить математическую модель по определению оптимальной производственной программы, предполагая, что каждая деталь при ее изготовлении должна последовательно обрабатываться на каждом из станков:

а) максимум товарной продукции;

б) максимум прибыли;

в) максимум продукции в натуральном выражении;

г) максимум товарной продукции при заданном ассортименте 3:2:1.

Провести сравнительный анализ полученных решений.

10.8. Найти компромиссное решение с учетом отклонения от максимального значения по первому критерию на 40%.

$$Z_{1\max} = x_1 + x_2;$$
$$Z_{2\min} = x_1 + 3x_2;$$
$$3x_1 + 2x_2 \geq 9;$$
$$2x_1 - 3x_2 \leq 8;$$
$$-x_1 + 3x_2 \leq 2;$$
$$x_2 \leq 5;$$
$$x_1 \geq 0;\ x_2 \geq 0.$$

10.9. Найти компромиссное решение задачи, считая второй критерий наиболее предпочтительным. Его отклонение от минимального значения 20%.

$$Z_{1\max} = 2x_1 + 4x_2;\ Z_{2\min} = x_1 + x_2;$$
$$4x_1 + 4x_2 \leq 20;$$
$$12x_1 + 3x_2 \geq 24;$$
$$x_1 \leq 3,$$
$$x_2 \leq 3,$$
$$x_1 \geq 0;\ x_2 \geq 0.$$

Глава 11
Динамическое программирование

11.1. Общие понятия о динамическом программировании

Развитие производительных сил и усложнение связей в обществе привели к необходимости научного подхода к проблеме планирования и управления. Оказалось, что многие задачи планирования и управления можно трактовать как процесс поэтапного выбора (принятия решений).

Задачи оптимального планирования (управления) обычно состоят в следующем. Рассматривается некоторая система S, состояние которой со временем меняется. Процесс изменения состояний системы – управляемый, т. е. можно влиять на его ход посредством выбора управляющих факторов U_i. С процессом связан некоторый критерий W, характеризующий качество управления и зависящий от него. Оптимальность планирования означает выбор наилучшего управления для достижения поставленной цели, т.е. выбор такого U, чтобы $W(U)$ было оптимальным.

Задачи динамического программирования имеют ряд особенностей.

1. В них рассматривается процесс поведения системы во времени.

2. Состояние системы в каждый момент времени однозначно определяется численными значениями небольшого набора параметров.

3. Операция выбора решения состоит в преобразовании этого набора параметров в такой же набор с другими числовыми значениями.

4. Если система в рассматриваемый момент времени находится в некотором состоянии, то ее поведение в дальнейшем определяется этим состоянием и выбираемым управлением, но не зависит от предыстории (т. е. от того, в каких состояниях находилась система до этого момента).

Пусть имеется система S, состояние которой характеризуется вектором $\overline{x} = (x_1, x_2, ..., x_n)$.

Численное значение набора параметров в момент времени i обозначим через $\overline{x}_i$.

Пусть в момент i на систему S воздействует управляющий фактор U^i, в результате которого в следующий момент $i + 1$ она переходит в новое состояние. Символически такой переход можно представить в виде

$$S(\overrightarrow{x^i})^{U^i} \to S(\overrightarrow{x^{i+1}}).$$

Пример 11.1. Пусть имеется сеть дорог, соединяющих населенные пункты. Рассмотрим в качестве системы человека, передвигающегося из одного пункта в другой.

Процесс движения естественно рассматривать как многоэтапный. Каждый этап состоит в переходе из одного пункта в соседний и совершается (условно) за единицу времени. Человек, находящийся в момент i в каком-то пункте K_i, вообще говоря, может из него попасть в различные соседние пункты. Управление U^i, принимаемое системой – человеком, и будет определять выбор конкретного соседнего пункта K_{i+1}, в который человек будет двигаться из K_i. Очевидно, что следующие состояния человека (пункт K_{i+1}), в котором он окажется через единицу времени, однозначно определяется его предыдущим состоянием K_i (пунктом, где он был) и принятым управлением U^i – решением куда двигаться.

Если в нашем примере человек поставил цель попасть из пункта A в пункт B, то естественно возникает задача выбрать управление так, чтобы движение из A в B проходило по кратчайшему пути. Оптимальное управление – это последовательность решений о том, через какие пункты надо двигаться, чтобы пройденный путь из A в B был самым коротким из всех возможных.

Специфика метода динамического программирования состоит в том, что для отыскания оптимального управления изучаемый процесс расчленен на этапы. Он представляется как развивающийся по этапам, причем каждый раз оптимизируется управлением только на одном этапе.

Динамическое программирование (планирование) – это планирование дальновидное, с учетом будущего, а не близорукое, когда руководствуются принципом «лишь бы хорошо сейчас, а там – что будет».

Как же находить оптимальное управление в многошаговом процессе? Общее правило состоит в том, что управление на каждом шаге надо выбирать с учетом будущего. Из этого правила есть исключение – это последний шаг, где можно действовать без оглядки на будущее: его на последнем шаге нет.

Управление на последнем шаге надо выбирать так, чтобы получить наибольший эффект без учета будущего (так как его нет). Поэтому процесс динамического планирования естественно разворачивается с конца, сначала планируется последний шаг. Как же это планировать, если неизвестно, чем кончился предпоследний шаг.

Очевидно, надо сделать различные предположения о том, чем кончился предпоследний шаг, и для каждого из них выбрать управление на последнем. Таким образом, приходим к понятию условно оптимального управления – оптимального управления, найденного в предположении, что предыдущий шаг окончился так-то.

Динамическое программирование основывается на принципе нахождения на каждом шаге условно оптимального управления для каждого возможного исхода предшествующего шага.

Руководствуясь этим принципом, мы разворачиваем процесс нахождения оптимального управления с конца, находя сначала условно оптимальное управление для каждого возможного исхода предпоследнего шага, а затем на основе его условно оптимальное управление на предпоследнем шаге и т. д., пока не дойдем до первого шага. На первом шаге нам надо делать гипотезы о состоянии системы – мы знаем, с чего начинается процесс, и можем с учетом найденных условно оптимальных управлений на последующих шагах найти безусловно оптимальное управление на первом шаге, т. е. такое управление, которое с учетом всех условно оптимальных управлений на последующих шагах дает оптимальное управление для всего процесса.

Итак, методология динамического программирования состоит в расчленении задачи на этапы и поэтапном построении оптимального управления на каждом шаге.

11.2. Задача о замене оборудования

Идеи и принципы динамического программирования рассмотрим на простом примере, подобранном так, чтобы основные идеи не загромождались техническими деталями. В связи с этим наша задача будет предельно простой, но за счет некоторых изменений может быть сделана адекватной действительности.

Известно, что проблема замены старого парка машин новыми, устаревших орудий – современными – одна из основных проблем индустрии.

Оборудование со временем изнашивается, стареет физически и «морально»; в процессе эксплуатации либо падает производительность оборудования, либо растут эксплуатационные расходы, либо и то и другое. Поэтому задача о замене оборудования весьма актуальна.

Формулировка задачи: рассматривается плановый период из нескольких лет, в начале которого имеется одна машина фиксированного возраста. В процессе работы машина дает ежегодно доход, требует эксплуатационных затрат и имеет остаточную стоимость, причем все перечисленные характеристики зависят от возраста машины. В любой год машину можно сохранить или продать по остаточной стоимости и купить новую по известной цене (которая может меняться со временем). Задача состоит в следующем: для каждого года в плановом периоде надо решить – сохранять имеющуюся в этот момент машину или продать ее и купить новую с тем, чтобы суммарная прибыль за весь плановый период была максимальной.

Переход системы S из одного состояния в другое за 1 год в зависимости от принятого решения можно изобразить графически (рис. 11.1).

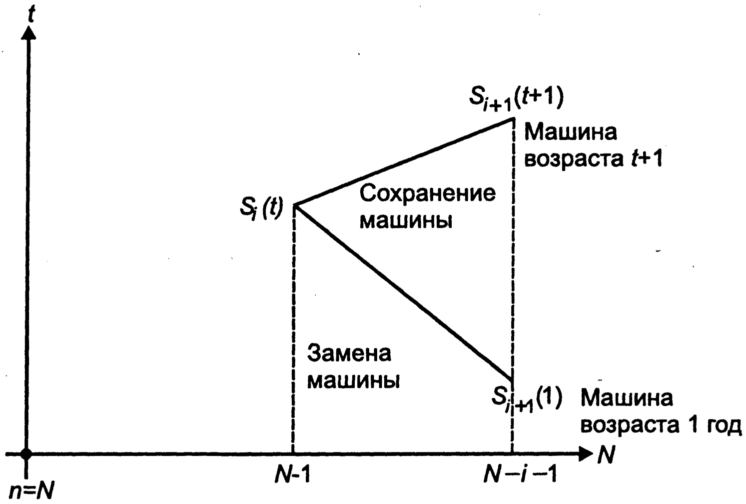


Рис. 11.1. Переход системы S из состояния в «новое» состояние за 1 год

Введем в рассмотрение функцию $f_n(t)$ – величину суммарного дохода (прибыли) за последние n лет планового периода при условии, что в начале этого периода из n лет имеется машина возраста t.

Функции $f_1(t)$, $f_2(t)$, ..., $f_n(t)$, учитывающие вклад последующих шагов в общий эффект, называются *функциями Беллмана* – по фамилии американского математика Р. Беллмана, создателя метода динамического программирования. С помощью этих функций ведется анализ задач динамического программирования. Очевидно, если мы сумеем вычислить $f_n(t_0)$ и найти политику замен, то это и будет решение задачи.

Предположим, что к началу последнего года планового периода $n = 1$ у нас имеется машина возраста t. В нашем распоряжении две возможности. Рассмотрим их.

Возможность 1-я: сохранить машину и, следовательно, получить за последний год доход

$$Z(t) - U(t). \tag{11.1}$$

Введем следующие обозначения:

t — возраст машины: $t = 0, 1, 2, ...,$ ($t = 0$ – соответствует использованию новой машины, $t = 1$ – соответствует использованию машины возраста 1 год и т.д.);

$Z(t)$ — стоимость продукции, производимой за 1 год на машине возраста t;

$U(t)$ — эксплуатационные затраты за 1 год на машину возраста t;

$S(t)$ — остаточная стоимость машины возраста t;

T — текущее время в плановом периоде;

$P(T)$ — цена новой машины в году T (может меняться со временем от изменения цен; для простоты будем считать, что $P(T) = P$, т. е. не зависит от времени);

t_0 — начальный возраст машины;

N — длина планового периода.

Мы сделали ряд упрощений, чтобы не осложнять анализа задачи. Так, например, мы считаем, что функции $Z(t)$, $U(t)$, $S(t)$ не зависят от текущего времени.

В нашей задаче в качестве системы S рассматривается машина, набор параметров, характеризующих ее состояние, состоит из единственного параметра – возраста машины. В качестве возможных управлений рассматриваются два решения – о сохранении имеющейся машины и о ее замене на новую. Условимся считать, что решения принимаются в моменты $n = N$; $N - 1$:... 1. Таким образом, плановый период разбит на шаги длиной в 1 год и в каждый из них решается – сохранить или заменить машину.

Возможность 2-я: продать имеющуюся машину и купить новую, что обеспечит в последний год доход

$$S(t) - P + Z(0) - U(0). \tag{11.2}$$

Для принятия решения необходимо вычислить функцию Беллмана $f_1(t)$, которая для нашего случая имеет вид:

$$f_1(t) = \max \begin{cases} Z(t) - U(t) & \text{сохранение машины;} \\ S(t) - P + Z(0) - U(0) & \text{замена машины.} \end{cases} \tag{11.3}$$

Задача будет решена, если мы определим прибыль за весь плановый период, т.е. найдем значение функции $f_N(t)$. Вначале попытаемся установить связь между выражениями

$$f_{n+1} \text{ и } f_n. \tag{11.4}$$

Если связь между ними будет найдена, то последовательно, двигаясь с конца, где $n = 1$, и зная $f_1(t)$, сможем найти $f_2(t)$, ..., $f_n(t)$, ..., $f_N(t)$ и тем самым решить задачу.

Итак, предположим, что с конца планового периода остается $n + 1$ год; в нашем распоряжении имеется машина возраста t и мы ищем оптимальную политику для периода длиной $n + 1$ год. Этот период разбивается на две части (рис. 11.2).

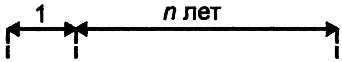


Рис. 11.2.

Рассмотрим все возможные решения в «первом году» для машины возраста t и для каждого состояния системы найдем оптимальную политику в оставшейся части из n последних лет. Так мы получим политики на весь период из $n + 1$ последних лет, лучшая из которых и будет условно оптимальной для всего периода.

В случае сохранения машины доход за рассматриваемый период определяется выражением:

$$Z(t) - U(t) + f_n(t + 1). \tag{11.5}$$

В случае замены машины аналогичной имеем:

$$S(t) - P + Z(0) - U(0) + f_n(1). \tag{11.6}$$

Для принятия окончательного решения вычислим функцию Беллмана вида

$$f_{n+1} = \max\begin{cases} Z(t) - U(t) + f_n(t+1) & \text{сохранение машины;} \\ S(t) - P + Z(0) - U(0) + f_n(1) & \text{замена машины.} \end{cases} \tag{11.7}$$

Рекуррентные формулы (11.3.; 11.7) позволяют реализовать концепцию динамического программирования и развернуть процесс нахождения оптимальной политики с конца, последовательно находя $f_1(t)$, $f_2(t)$, ..., $f_n(t)$, ..., $f_N(t)$ для различных значений t.

Пример 11.1. Пусть функции $Z(t)$, $U(t)$ и значения $\Delta = Z(t) - U(t)$ заданы табл. 11.1.

Мы ограничились машиной возраста $t \le 10$ лет, так как из табл. 11.1 видно, что машина возраста $t > 10$ лет невыгодна. Будем считать условно (для простоты вычислений), что:

1) остаточная стоимость машины равна 0;

2) цена новой машины со временем не меняется и равна 10 условным единицам;

3) длина планового периода N равна 10 годам, т. е. $S(t) = 0$; $P = 10$.

Таблица 11.1

Исходные данные

t	0	1	2	3	4	5	6	7	8	9	10
$Z(t)$	20	20	20	19	19	18	18	17	17	16	15
$U(t)$	10	11	12	12	13	13	14	14	15	15	15
Δ	10	9	8	7	6	5	4	3	2	1	0

Тогда формулы (11.3 и 11.6) принимают вид

$$f_1(t) = Z(t) - U(t) \text{ сохранение машины.} \tag{11.8}$$

$$f_{n+1}(t) = \max \begin{cases} Z(t) - U(t) + f_n(t+1) & \text{сохранение машины;} \\ f_n(1) & \text{замена машины.} \end{cases} \tag{11.9}$$

Используя полученные формулы, вычислим значения функций Беллмана $f_n(t)$ при различных n и t. Значения функций будем вписывать в табл. 11.2.

Таблица 11.2

Решение задачи о замене оборудования с использованием метода динамического программирования

$f_n(t)$ \ t	0	1	2	3	4	5	6	7	8	9	10
$f_1(t)$	10	9	8	7	6	5	4	3	2	1	0
$f_2(t)$	19	17	15	13	11	9	9	9	9	9	9
$f_3(t)$	27	24	21	18	17	17	17	17	17	17	17
$f_4(t)$	34	30	26	24	24	24	24	24	24	24	24
$f_5(t)$	40	35	32	31	30	30	30	30	30	30	30
$f_6(t)$	45	41	39	37	36	35	35	35	35	35	35
$f_7(t)$	51	48	45	43	41	41	41	41	41	41	41
$f_8(t)$	58	54	51	48	48	48	48	48	48	48	48
$f_9(t)$	64	60	56	55	54	54	54	54	54	54	54
$f_{10}(t)$	70	65	63	61	60	60	60	60	60	60	60

Заполнение таблицы будем производить по строкам: сначала заполним первую строку, потом вторую и т. д. Заметим, что согласно формуле (11.8.) первая строка табл. 11.2 совпадает с последней строкой табл. 11.1.

Теперь перейдем к заполнению второй строки таблицы:

$$f_2(0) = \max\begin{cases} Z(0) - U(0) + f_1(1) \\ f_1(1) \end{cases} = \max\begin{cases} 10+9 \\ 9 \end{cases} = 19 \text{ сохранение машины,}$$

$$f_2(1) = \max\begin{cases} Z(1) - U(1) + f_1(2) \\ f_1(1) \end{cases} = \max\begin{cases} 9+8 \\ 9 \end{cases} = 17 \text{ сохранение машины,}$$

$$f_2(5) = \max\begin{cases} Z(5) - U(5) + f_1(6) \\ f_1(1) \end{cases} = \max\begin{cases} 5+4 \\ 9 \end{cases} = 9 \text{ сохранение машины.}$$

Здесь обе политики – сохранения и замены – обеспечивают одинаковую прибыль – 9 единиц, выбираем в этом случае «старую» машину, к которой «привыкли». Далее

$$f_2(6) = \max\begin{cases} Z(6) - U(6) + f_1(7) \\ f_1(1) \end{cases} = \max\begin{cases} 4+3 \\ 9 \end{cases} = 9 \text{ замена машины.}$$

Для того чтобы в таблице различать, в результате какой политики получается оптимальный доход (т. е. в данном случае $f_2(6)$), мы будем величину оптимального дохода, соответствующую политике замены, записывать особым цветом, например, красным. Итак, значение $f_2(6)$ в таблице будет записано красным цветом.

Можно показать, что $f_2(7) = f_2(8) = f_2(9) = f_2(10) = 9$ и соответствует замене машины.

После заполнения второй строки таблицы (11.2) заполняем третью, четвертую и т.д. В итоге черный цвет в таблице соответствует политике сохранения машины, а красный – политике ее замены (в табл. 11.2 жирная черта отделяет черные и красные числа).

Построенная табл. 11.2 содержит очень много ценной информации и позволяет решать целый ряд однотипных задач.

Допустим, вначале имеется машина возраста 7 лет. Посмотрим, какова будет оптимальная политика действий для получения максимальной прибыли за 10 лет планового периода.

Величина максимальной прибыли (согласно табл. 11.2) определяется функцией Беллмана $f_{10}(7) = 60$. Теперь найдем оптимальную политику, обеспечивающую эту прибыль.

Так как $f_{10}(7)$ вписано в таблицу красным цветом, то для достижения максимальной прибыли необходимо в первом году рассматриваемого периода заменить машину на новую. По истечении одного года мы за 9 лет до конца планового периода будем иметь машину возраста один год. Теперь надо действовать оптимально в оставшийся период, располагая машиной возраста один год, т. е. найти $f_9(1)$ из девяти лет. Из табл. 11.2 видно, что $f_9(1)$ – черное, следовательно, во втором году надо сохранить машину. Рассматривая процесс по годам, замечаем: $f_8(2)$ – черное, $f_7(3)$ – черное, $f_6(4)$ – черное, $f_5(5)$ – красное. Последнее выражение ($f_5(5)$ – красное) указывает на то, что по истечении пяти лет планового периода машину надо менять на новую. Действуя далее оптимально, найдем последовательно: $f_4(1)$ – черное, $f_3(2)$ – черное, $f_2(3)$ – черное, $f_1(4)$ – черное. Итак, используя табл. 11.2, мы найдем оптимальную политику, которую можно представить схемой

$$f_{10}(7) \xrightarrow[\text{замена}]{} f_9(1) \xrightarrow[\text{сохранение}]{} f_8(2) \xrightarrow[\text{сохранение}]{} f_7(3) \xrightarrow[\text{сохранение}]{} f_6(4) \xrightarrow[\text{сохранение}]{}$$

$$f_5(5) \xrightarrow[\text{замена}]{} f_4(1) \xrightarrow[\text{сохранение}]{} f_3(2) \xrightarrow[\text{сохранение}]{} f_2(3) \xrightarrow[\text{сохранение}]{} f_1(4) \longrightarrow .$$

В рассмотренной задаче число возможных решений, принимаемых ежегодно, равно двум (сохранить машину или заменить). На практике часто применяют решение о покупке неновой машины; в этом случае необходимо включить в число возможных решений (управлений) решение: заменить имеющуюся машину возраста t на машину возраста $t_1 < 5$ (на старую заменять невыгодно).

В этом случае рекуррентные формулы (11.3 и 11.7) существенно усложняются. Динамическое программирование позволяет учесть все решения (управления), которые вызывают практический интерес.

Эффективность динамического программирования обусловлена использованием рекуррентных формул (11.3; 11.7), позволяющих осуществить рациональный процесс поиска оптимальных вариантов, чем полный перебор вариантов. Это делается при помощи функций Беллмана, несущих информацию об оптимальном продолжении процесса.

Глава 12
Параметрическое программирование

12.1. Постановка задачи и геометрический метод ее решения

Задачи параметрического программирования являются обобщением задач линейного программирования. Это обобщение состоит в том, что исходная информация задач параметрического программирования изменяется линейно в зависимости от некоторого параметра. Если предположить, например, что произведенная предприятием продукция подлежит хранению, то ее стоимость будет складываться из двух частей:

постоянной – стоимости продукции на момент изготовления;

переменной – зависящей от срока хранения продукции.

Целевая функция, если она отражает стоимость, будет выражена через коэффициенты, линейно зависящие от параметра t (время).

Часто на практике встречаются задачи, в которых значение коэффициентов целевой функции известны лишь приближенно. Представив их в виде линейных функций параметра t, можно проанализировать решения задач при различных значениях этих коэффициентов. Аналогично можно провести исследования для случая, когда изменяются коэффициенты системы ограничений. Остановимся только на целевой функции. Итак, задача заключается в следующем. Дана система ограничений:

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le a_1 \\ &a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le a_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le a_m \\ &x_j \ge 0; \quad j = \overline{1, n}. \end{aligned} \tag{12.1}$$

Задана целевая функция Z:

$$Z_t = \sum_{j=1}^{n} (c_j + d_j t)x_j \to \max, \tag{12.2}$$

в которой коэффициенты при переменных являются линейными функциями параметра t, $t \in [\alpha, \beta]$.

В данном случае определить просто оптимальный план нельзя, так как он может изменяться при различных значениях t. Поэтому задача формулируется следующим образом: требуется разбить отрезок $[\alpha, \beta]$. на конечное число интервалов, в которых целевая функция Z_t достигает максимума (минимума) в одной и той же вершине многогранника решений.

Проведем геометрическую интерпретацию задачи параметрического программирования. Полагая $t = \alpha$ и ограничиваясь только двумя переменными, получаем обычную задачу линейного программирования (рис. 12.1).

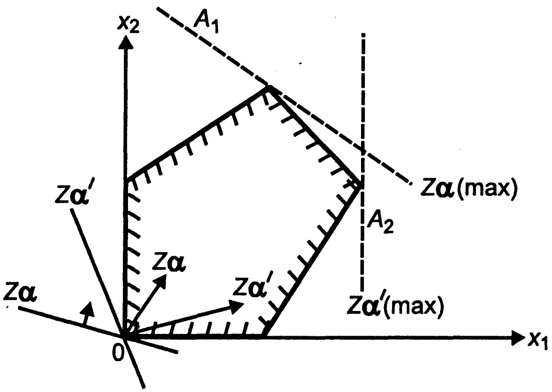


Рис. 12.1. Геометрическая интерпретация задачи

При $t = \alpha'$ изменяются коэффициенты целевой функции, геометрически это значит, что прямая, соответствующая Z_α, повернулась на некоторый угол. Из рис. 12.1 следует, что $Z_{\alpha\,\max}$ достигается в точке A_1, а $Z'_{\alpha\max}$ – в точке A_2. Естественно предположить, что при некотором значении t′ максимальное значение достигается одновременно в точках A_1 и A_2. В этом случае прямая, соответствующая целевой функции, параллельна стороне A_1A_2 многоугольника решений. Фиксированное значение t' является граничной точкой между двумя соседними интервалами отрезка $[\alpha, \beta]$. Исходя из изложенного, задачу параметрического программирования с двумя переменными можно решить графически.

Пример 12.1. Разбить отрезок $t \in [0; 8]$ на такие интервалы, в которых целевая функция $Z_t = 4x_1 + (2 + t)x_2$ достигает максимума в одной и той же вершине многогранника решений, и найти соответствующее значение переменных:

$$2x_1 - 5x_2 \le 10;$$
$$x_1 + x_2 \ge 5;$$
$$-x_1 - x_2 \le 4;$$
$$4x_1 + 5x_2 \le 40;$$
$$x_1 \ge 0;\ x_2 \ge 0.$$

Решение

Положив $t = 0$ (наименьшее значение t из заданного промежутка), находим геометрически $Z_{0\max}$ (рис. 12.2) в точке C.

Далее приравняем Z_t к нулю и находим уравнение разрешающей прямой при любом t:

$$x_2 = -\frac{4}{2+t}x_1.$$

Выпишем угловой коэффициент K_z этой прямой и исследуем его поведение при изменении параметра t: $K_z = -4/(2 + t)$ при $t = 0$; $K_Z = -2$.

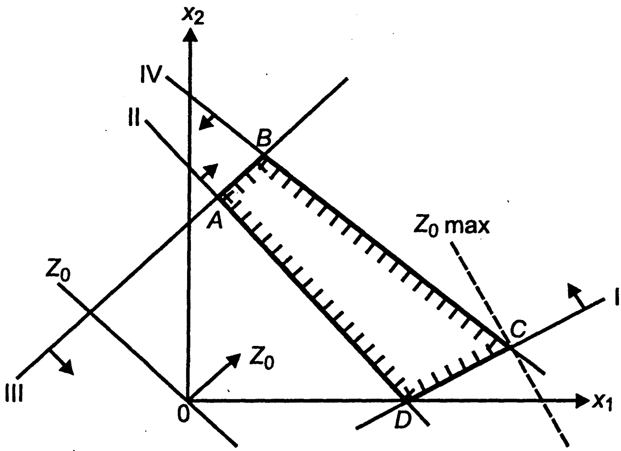


Рис. 12.2. Геометрическое решение задачи параметрического программирования

Найдем производную углового коэффициента по параметру t:

$$(K_z)' = -4/(2 + t)^2.$$

Так как производная при любом t положительна, угловой коэффициент при увеличении t возрастает. Найдем предел его возрастания

$$\lim_{t\to\infty} K_z = \lim(-4/(2+t)) = -0.$$

При $t \to +\infty$ угловой коэффициент K_z приближается к нулю со стороны отрицательных значений, следовательно, разрешающая прямая поворачивается против часовой стрелки до предельного горизонтального положения. В процессе указанного анализа необходимо понять, что при вертикальном положении прямой угловой коэффициент как функция терпит разрыв. При вращении прямой против часовой стрелки от оси абсцисс до вертикального положения угловой коэффициент возрастает от 0 до $+\infty$, при дальнейшем вращении прямой он возрастает от $-\infty$ до 0.

В рассматриваемом примере при изменении параметра t от нуля до некоторого значения максимальное значение целевой функции будет в вершине C, затем в некоторый фиксированный момент времени оптимум будет достигаться на отрезке BC, а затем он перейдет в точку B и останется в ней для всех значений t отрезка [0; 8]. Определим значение параметра t, при котором Z_{max} будет соответствовать всем точкам отрезка BC.

Поскольку в этот момент прямая BC и разрешающая прямая должны быть параллельны, приравняем их угловые коэффициенты. Угловой коэффициент прямой BC, $K_{BC} = -4/5$, следовательно, $-4/(2 + t) = -4/5$, откуда $t = 3$. Итак, при $0 \le t < 3$ оптимальное решение будет в вершине C (8,3; 1,3), при $t = 3$ оно достигается на всем отрезке BC, а при $3 < t \le 8$ – в точке B (2,2; 6,2).

12.2. Аналитический метод решения задач параметрического программирования

Алгоритм решения задачи в общей форме состоит из двух частей.

1. Фиксируем значение t из заданного промежутка, допустим самое меньшее $t = \alpha$. Тогда в целевой функции все коэффициенты станут постоянными. Решаем полученную задачу симплекс-методом и находим вершину, в которой Z_α достигает максимума.

2. Определяем все значения параметра t, для которых максимум достигается в той же вершине. Найденный интервал исключаем из заданного отрезка $[\alpha, \beta]$. Для оставшегося интервала снова решаем задачу, т. е. переходим к п.1 алгоритма. Так продолжаем до тех пор, пока не будет разделен на частичные интервалы весь отрезок$[\alpha, \beta]$.

Проанализируем пункты алгоритма более подробно.

1. Полагая $t = \alpha$, целевая функция принимает вид:

$$Z_\alpha = \sum_{j=1}^{n} (c_j + d_j\alpha)x_j.$$

Составляем первую симплекс-таблицу, добавляя две строки: для коэффициентов c_j и d_j (табл. 12.1):

Таблица 12.1

Форма симплекс-таблицы для решения задачи параметрического программирования

Свободный / Базисный	Свободные члены	$x_1, x_2, \ldots \quad x_n$	$x_{n+1} \ldots \quad x_{n+m}$
x_{n+1} . . . x_{n+m}	a_1 . . . a_m	a_{ij}	1 1 1 1
$Z\alpha$	0	$-(c_1 + d_1\alpha) \ldots - (c_n + d_n\alpha)$	00... 0
Z_t		$-c_1, -c_2 \ldots \quad -c_n$	00... 0
		$-d_1, -d_2 \ldots \quad -d_n$	00... 0

В строке Z_α табл. 12.1 в каждом столбце фактически будет стоять одно число, равное $c_j + d_j\alpha$. В последних двух строках записаны индексы линейной формы Z_t для произвольного параметра t.

Обычным симплекс-методом находим оптимальный план, причем этим же преобразованиям подвергаются элементы двух последних строк. В итоге получаем данные табл. 12.2.

Необходимо отметить, что в табл. 12.2 столбцы, соответствующие базисным переменным, являются нулевыми (т. е. содержат 1 на пересечении S-столбца и S-строки, а остальные нули).

Поскольку план, полученный в табл. 12.2, оптимален, все коэффициенты Z_α строки неотрицательны:

$$(p_j + q_j a) \geq 0.$$

После этого переходим ко второму пункту алгоритма. Надо выделить интервал времени t, в котором максимум Z_t достигается в той же вершине. Иными словами, надо найти такие значения t, для

Таблица 12.2

Общий вид оптимального плана

Свободный / Базисный	Свободные члены	$x_1, x_2, \dots \quad x_n$	$-x_{n+1} \dots$ x_{n+m}
x_1 x_S x_{n+m}	b_1 b_S	b_{ij}	1 1 1 1
$Z\alpha$	$P + Q_\alpha$	$p_1 + q_1\alpha \dots \quad p_n + q_n\alpha$	$p_{n+m} + q_{n+m}\alpha$
Z_t	P	$p_1 \dots \quad p_n$	p_{n+m}
	Q	$q_1 \dots \quad q_n$	q_{n+m}

которых план, полученный в табл. 12.2, будет оставаться оптимальным. План оптимален, если все коэффициенты функции Z_t будут неотрицательны:

$$\begin{aligned} &p_1 + q_1 t \geq 0 \\ &p_2 + q_2 t \geq 0 \\ &\dots\dots\dots\dots\dots \\ &p_n + q_n t \geq 0. \end{aligned} \tag{12.3}$$

Столбцы, соответствующие базисным переменным, можно не рассматривать. Из решения полученной системы неравенств можно найти интересующий нас интервал.

Рассмотрим несколько случаев:

а) Пусть все коэффициенты $q_j > 0$, $j = \overline{1,n}$. Тогда $t \geq -p_j/q_j$, параметр t, удовлетворяющий системе (12.3), определяется выражением $t \geq \max\left(-\dfrac{p_j}{q_j}\right)$. Верхняя граница для t стремится к $+\infty$.

Таким образом, окончательно

$$\alpha_1 = \max\left(-\frac{p_j}{q_j}\right) \leq t \langle +\infty = \alpha_2. \tag{12.4}$$

В интервале (12.4) целевая функция достигает максимума в той же вершине.

б) Пусть все коэффициенты $q_j < 0$, тогда из (12.3) имеем $t \leq -p_j/q_j$. Для того чтобы значение t удовлетворяло всей системе (12.3), необходимо положить $t \leq \min\left(\frac{-p_j}{q_j}\right)$. Нижняя граница для t стремится к $-\infty$, поэтому

$$\alpha_1 = -\infty \langle t \leq \min(-p_j/q_j) = \alpha_2. \qquad (12.5)$$

в) Пусть среди элементов q_j имеются как положительные, так и отрицательные числа.

Разделив неравенства системы (12.3) на две группы соответственно знакам q_j, получаем два интервала (12.4) и (12.5), объединяя эти интервалы, получим

$$\alpha_1 = \max_{q_j > 0}(-p_j/q_j) \leq t \leq \min_{q_j < 0}(-p_j/q_j) = \alpha_2. \qquad (12.6)$$

г) Если $q_j = 0$, то соответствующий коэффициент функции будет неотрицательным при любом t, поэтому на такой столбец можно не обращать внимания.

Сравниваем полученный интервал $(\alpha_1; \alpha_2)$ с заданным $[\alpha, \beta]$. Независимо от значения α_1, левой границей первого интервала будет α, так как α_1 больше α быть не может. Если $\alpha_2 > \beta$, то весь отрезок попадает внутрь интервала $[\alpha_1, \alpha_2]$, и задача решена. Для любого значения параметра $t \in [\alpha, \beta]$ максимум Z_t достигается в одной и той же вершине

Если $\alpha_2 < \beta$, то найденный интервал исключаем из рассмотрения и решаем задачу для оставшегося интервала (α_2, β). Для этого полагаем $t = \alpha_2$ и заменяем строку Z_α строкой $Z_{\alpha 2}$. За разрешающий столбец в новой таблице выбираем тот, по которому определено значение $t = \alpha_2$ (в этом столбце на пересечении с $Z_{\alpha 2}$-строкой находится элемент, равный нулю). Если нули находятся в нескольких столбцах, то за разрешающий можно брать любой из них.

Для найденного решения снова определяем интервал изменения параметра t и т. д. Если в разрешающем столбце не окажется положительных элементов, задача на оставшемся интервале не имеет решения.

Пример 12.2. Решить задачу параметрического программирования

$$2x_1+3x_2+x_3\le 1;$$
$$3x_1+2x_2-2x_3\le 3;$$
$$4x_1+2x_2+x_3\le 4;$$
$$x_j\ge 0\quad j=\overline{1;\ 3};$$
$$Z_t=tx_1+(1+t)x_2+(6-2t)x_3\to \max;$$
$$t\in[1;\ 8].$$

Решение

Приведем задачу к виду, допускающему применение алгоритма, и найдем выражение целевой функции при $t=1$ (левый конец заданного промежутка):

$$\begin{aligned}
&2x_1+3x_2+x_3+x_4 &&=1;\\
&3x_1+x_2-2x_3+\quad x_5+x_6 &&=3;\\
&4x_1+2x_2+x_3+\qquad\quad x_6 &&=4;\\
&Z_1=x_1+2x_2+4x_3\to \max.
\end{aligned}$$

Составим первую симплекс-таблицу с дополнительными строчками для функции Z_t и выполним преобразование симплекс-таблицы для решения задачи (табл. 12.3–12.4).

Таблица 12.3

	0	1	2	4	0	0	0	
Свободные / Базисные	св. ч	x_1	x_2	x_3	x_4	x_5	x_6	Σ
x_4	1	2	3		1	0	0	
x_5	3	3	1	–2	0	1	0	
x_6	4	4	2	1	0	0	1	
Z_t	0	–1	–2	–4	0	0	0	
Z_t	0	0	–1	–6	0	0	0	
	0	–1	–1	2	0	0	0	

Таблица 12.4

	0	1	2	4	0	0	0	
Свободные / Базисные	св. ч	x_1	x_2	x_3	x_4	x_5	x_6	Σ
x_3	1	2	3	1	1	0	0	
x_5	5	7	7	0	2	1	0	
x_6	3	2	–1	0	–1	0	1	
Z_1	4	7	10	0	4	0	0	
Z_t	6	12	17	0	6	0	0	
	–2	–5	–7	0	–2	0	0	

Таблица 12.5

	0	1	2	4	0	0	0	
Свободные / Базисные	св. ч	x_1	x_2	x_3	x_4	x_5	x_6	Σ
x_3	1	$\boxed{2}$	3	1	1	0	0	
x_5	5	7	7	0	2	1	0	
x_6	3	2	–1	0	–1	0	1	
$Z_{12/5}$	6/5	0	1/5	0	6/5	0	0	
Z_t	6	12	17	0	6	0	0	
	–2	–5	–7	0	–2	0	0	

Таблица 12.6

	0	1	2	4	0	0	0	
Свободные / Базисные	св. ч	x_1	x_2	x_3	x_4	x_5	x_6	Σ
x_1	1/2	1	3/2	1/2	1/2	0	0	
x_5	3/2	0	–7/2	–7/2	–3/2	1	0	
x_6	2	0	–4	–1	–2	0	1	
$Z_{12/5}$	6/5	0	1/5	0	6/5	0	0	
Z_t	0	0	–1	–6	0	0	0	
	1/2	0	1/2	5/2	1/2	0	0	

В табл. 12.4 получено оптимальное решение, так как все коэффициенты Z_1-строки неотрицательны – $x = (0;\ 0;\ 1;\ 0;\ 5;\ 3)$. Определим интервал для t, в котором оптимальное решение будет в той же вершине. Так как все $q_j < 0$ (табл. 12.4), то согласно (12.5) $\alpha_1 = -\infty$, $\alpha_2 = 12{,}5$.

Для нашего случая $\alpha_1 = 1$ (левый конец заданного отрезка). Таким образом, на отрезке [1; 12/5] решение будет в точке $x = (0;\ 0{,}1;\ 0;\ 5;\ 3)$. Исключаем найденный отрезок из рассмотрения и решаем задачу для оставшегося отрезка [12/5; 8]. Для этого, полагая $t = 12/5$, вычисляем для него строку $Z_{12/5}$. В первом столбце получим нуль; остальные коэффициенты $Z_{12/5}$ строки положительны, а другие элементы остаются прежними (табл. 12.5). За разрешающий столбец в табл. 12.5 выбираем тот, по которому определено значение $t = \alpha_2$. Если нули находятся в нескольких столбцах, то за разрешающий можно брать любой из них. Новый план $x = (1/2;\ 0;\ 0;\ 0;\ 3/12;\ 0)$ оптимален, так как в строке $Z_{12/5}$ нет отрицательных элементов (практически элементы строки $Z_{12/5}$ остались без изменения).

В последней строке табл. 12.6 все $q_j > 0$, поэтому согласно (табл. 12.6) $\alpha_1 = 12/5$, $\alpha_2 = +\infty$; $t \in [12/5;\ 8]$. Таким образом, заданный промежуток разбился на два; при $t = 12/5$ оптимальное значение достигается в обеих вершинах и в любой точке, являющейся их выпуклой линейной комбинацией

$$A = \lambda_1 A_1 + \lambda_2 A_2;\quad \lambda_1 \geq 0;\quad \lambda_2 \geq 0;\quad \lambda_1 + \lambda_2 = 1.$$

Задачи

12.1. Решить графическим методом

$$
\begin{gathered}
Z_t = (2 + 2t)x_1 + (4 - 2t)x_2 \to \max; \\
x_1 + x_2 \le 6; \\
x_2 \le 4; \\
2x_1 + x_2 \le 10; \\
x_1 \ge 0;\ x_2 \ge 0;\ t \in [1;\ 15].
\end{gathered}
$$

12.2.

$$
\begin{gathered}
Z_t = 4x_1 + (2 + t)x_2 \to \max; \\
2x_1 - 5x_2 \le 10; \\
x_1 + x_2 \le 5; \\
-x_1 + x_2 \le 4; \\
4x_1 + 5x_2 \le 40; \\
x_1 \ge 0;\ x_2 \ge 0;\ t \in [0;\ 8].
\end{gathered}
$$

12.3. Решить аналитически

$$
\begin{gathered}
Z_t = (1 + t)x_1 + (2 - t)x_2 + (2 - 3t)x_3 + (1 - 2t)x_4 \to \max; \\
x_1 + x_2 + 2x_3 - x_4 \le 5; \\
x_1 + 2x_2 + x_4 \le 7; \\
x_1 + x_3 + 2x_4 \le 3; \\
x_1 \ge 0;\ j = \overline{1,4};\ t \in [1;\ 20].
\end{gathered}
$$

12.4.

$$
\begin{gathered}
Z_t = (10 - 10t)x_1 + (9 + t)x_2 + (7 - 2t)x_3 \to \max; \\
x_1 + x_2 \le 5; \\
2x_1 + x_3 \le 7; \\
x_1 + 2x_2 + 3x_3 \le 3; \\
x_j \ge 0;\ j = \overline{1,3};\ t \in [1;\ 10].
\end{gathered}
$$

12.5.

$$
\begin{gathered}
Z_t = tx_1 + (1 + t)x_2 \to \max; \\
-3x_1 + 4x_2 \le 12; \\
4x_1 + x_2 \le 8; \\
x_1 \ge 0;\ x_2 \ge 0;\ t \in [1;\ 7].
\end{gathered}
$$

Глава 13
Целочисленное программирование

Постановка задачи и алгоритм решения

По смыслу многих задач требуется, чтобы значение неизвестных в оптимальном плане выражались целыми числами. К ним относятся задачи, у которых переменные величины означают количество единиц неделимой продукции, например, распределение производственных заданий между предприятиями, раскрой материалов, загрузка оборудования, распределение судов по линиям, самолетов по рейсам, а также задачи по производству неделимой продукции. Если единица составляет малую часть всего объема производства, то оптимальное решение находят обычным симплекс-методом, округляя его до целых единиц, исходя из смысла задачи. В противном случае округление может привести к решению, не удовлетворяющему системе ограничений (особенно, если значения неизвестных выражены малыми числами). Поэтому для нахождения оптимального целочисленного решения нужен особый алгоритм.

Идея такого алгоритма заключается в следующем: вначале условие целочисленности не принимается во внимание и симплекс-методом отыскивается оптимальный план. Если этот план нецелочисленный, составляется дополнительное ограничение, которому удовлетворяет любое целочисленное решение, но заведомо не удовлетворяет найденное.

После введения этого ограничения в задачу план становится недопустимым, и решение продолжается опять до получения оптимума. Если новый оптимальный план также окажется нецелочисленным, снова формулируется дополнительное ограничение и для расширенной задачи снова находится оптимум и т. д.

Геометрически дополнительному ограничению соответствует гиперплоскость, которая отсекает от многогранника решений какую-то часть вместе с вершиной, которой соответствует нецелочисленное решение. При этом все точки с целочисленными координатами остаются в новом многограннике, так что через несколько операций такая точка становится вершиной с целочисленными координатами (рис. 13.1).

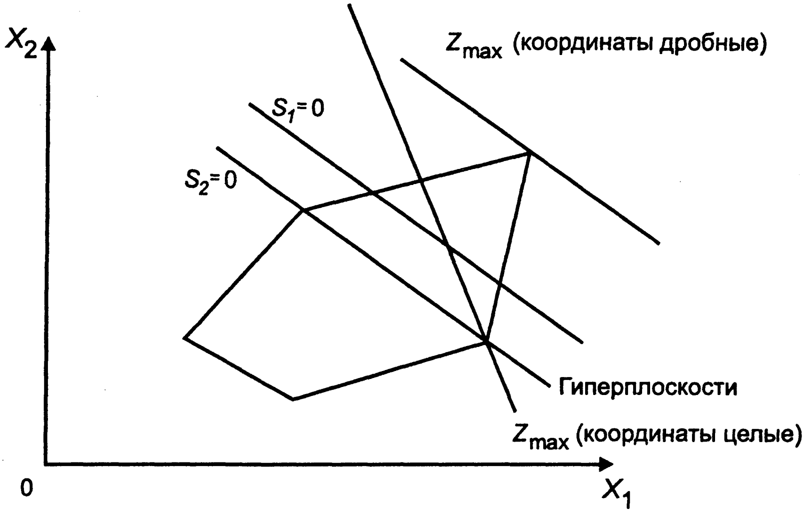


Рис. 13.1. Геометрическая интерпретация решения задачи целочисленного программирования

Пусть требуется найти максимум целевой функции

$$Z_{\max} = P_1x_1 + P_2x_2 + \ldots + P_nx_n \tag{13.1}$$

при ограничениях

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le a_1 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le a_m \\ &x_j \ge 0 \quad j = \overline{1,\ n}. \end{aligned} \tag{13.2}$$

Не нарушая общности, можно предположить, что все коэффициенты системы (13.2) выражаются целыми числами.

Приводим систему к виду

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + y_1 = a_1 \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots , \\ &a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n + y_m = a_m \end{aligned} \tag{13.3}$$

где y_i – выравнивающие неизвестные.

Отметим, что целочисленность переменных x_j влечет за собой целочисленность переменных y_i.

Составим первую симплекс-таблицу и после l шагов получим оптимальный план, который можно представить последней симплекс-таблицей.

Таблица 13.1

Общий вид первой симплекс-таблицы решения задачи целочисленного программирования

$Б$ \ C	Свободный член	$y_1 \dots$	y_ℓ $\chi_{\ell+1}$	$\dots$	χ_n
χ_1	b_1	$b_{11}\dots$	$b_{1\ell}$	$\dots$	b_{1n}
—					
χ_ℓ	$b_{\ell+1}$	$b_{\ell 11}\dots$	$b_{\ell\ell}$	$\dots$	$b_{\ell n}$
$y_{\ell+1}$	$b_{\ell+1}$	$b_{(\ell+1)1}$	$b_{(\ell+1)\ell}$	$\dots$	$b_{\ell+1;n}$
—					
y_m	b_m	$b_{m1}\dots$	$b_{m\ell}$	$\dots$	b_{mn}
$Z_{\max}$	Q	$q_1\dots$	q_ℓ	$\dots$	q_n

Если все свободные члены целые – задача решена. Для удобства все базисные переменные обозначим $\eta_i (i = \overline{1, m})$, а все свободные переменные ξ_j, $j = \overline{1, n}$, тогда таблица примет вид (табл. 13.2):

Таблица 13.2

$Б$ \ C	Свободный член	$\xi_1 \dots$	ξ_j	$\dots$	ξ_n
η_1	b_1	$b_{11}\dots$	b_{1ij}	$\dots$	b_{1n}
—					
η_i	b_i	$b_{i1}\dots$	b_{ij}	$\dots$	b_{in}
—					
η_m	b_m	$b_{m1}\dots$	b_{mj}	$\dots$	b_{mn}
$Z_{\max}$	Q	$q_1\dots$	q_j	$\dots$	q_n

С целью упрощения в таблице опущены:

а) столбцы, соответствующие базисным переменным;

б) целевая строка;

в) контрольный столбец.

Пусть свободный член b_i является дробным, в этой же строке среди коэффициентов b_{ij} могут быть как дробные, так и целые. Обозначим буквой n с соответствующими индексами антье от b_{ij} и b_i:

$$E(b_{ij}) = n_{ij};\ E(b_i) = n_i.$$

Вычислим разности между соответствующими коэффициентами и E:

$$\begin{aligned} \beta_{ij} &= b_{ij} - n_{ij}; \\ \beta_i &= b_i - n_i. \end{aligned} \tag{13.4}$$

Эти разности удовлетворяют условиям

$$\begin{aligned} 0 &\le \beta_{ij} \langle 1; \\ 0 &\le \beta_{ij} \langle 1. \end{aligned} \tag{13.5}$$

Выпишем из таблицы выражение для η_i и заменим в нем коэффициенты согласно условию (13.4), тогда получим:

$$\begin{aligned} \eta_i &= (\beta_{i1} + n_{i1}) \cdot (-\xi_1) + \ldots + (\beta_{in} + n_{in})(-\xi_n) + \beta_i + n_i; \\ \eta_i &= \sum_{j=1}^{n} n_{ij}(-\xi_j) + n_i + \sum_{j=1}^{n} \beta_{ij}(-\xi_j) + \beta_i. \end{aligned} \tag{13.6}$$

Перепишем последнее выражение в другом виде

$$S_i = -\sum_j \beta_{ij}(-\xi_j) - \beta_i = \sum_j n_{ij}(-\xi_j) + n_i - \eta_i. \tag{13.7}$$

Если ξ_i и η_i – целые, то из правой части (13.7) следует, что S_i – целое; так как $\xi \ge 0$; $\eta_i \ge 0$, а из средней части (13.7) следует $S_i \ge -\beta_i$; учитывая выражение (13.5), заключаем, что S_i может принимать значения: 0, 1, 2, ... и т. д.

Вывод. При любых целых неотрицательных η_i и ξ_i, S_i принимает целые неотрицательные значения.

Введем в задачу дополнительное ограничение (табл. 13.3), используя подчеркнутую часть выражения (13.7).

Дополнительному ограничению удовлетворяет любой целочисленный план издания, в то же время найденный ранее оптимальный план для расширенной задачи является недопустимым, так как $S_i = -b_i < 0$.

Для расширенной задачи обычным путем отыскивается сначала допустимый план, а затем оптимальный.

Таблица 13.3

$Б \backslash C$	Свободный член	$\xi_1 \ldots$	ξ_j	$\ldots$	ξ_n
η_1	b_1	$b_{11} \ldots$	b_{1ij}	$\ldots$	b_{1n}
—					
η_m	$-b_{m1}$	$-b_{i1} \ldots$	$-b_{mj}$	$\ldots$	$-b_{mn}$
S_i	$-\beta_i$	$-\beta_{i1} \ldots$	$-\beta_{ij}$	$\ldots$	$-\beta_{in}$
—					
Z	Q	$q_1 \ldots$	q_j	$\ldots$	q_n

Если в какой-либо строке таблицы (кроме строки Z) свободный член дробный, а все прочие коэффициенты целые, целочисленных решений нет.

Пример

$$Z_{\max} = 2x_1 + x_2 - 3x_3;$$
$$x_1 + 3x_2 - 2x_3 + y_1 = 4;$$
$$5x_1 - x_3 + y_2 = 12;$$
$$2x_1 - x_2 + 3x_3 + y_3 = 4;$$
$$x_j \geq 0;\ j = \overline{1,\ 3};$$
$$y_i \geq 0;\ i = \overline{1,\ 3}.$$

После двух шагов применения симплекс-метода приходим к оптимальному плану, представленному в табл. 13.4.

Решение оптимальное, но дробное $x = (16/7;\ 4/7;\ 0)$.

Выбираем среди свободных членов дробный, например, 4/7 из первой строки. По формулам (13.6) находим коэффициент дополнительного ограничения:

$$\beta_{11} = -\frac{1}{7} - (-1) = \frac{6}{7}; \qquad \beta_{12} = \frac{2}{7} - 0 = \frac{2}{7};$$
$$\beta_{13} = -1 - (-1) = 0; \qquad \beta_1 = \frac{4}{7} - 0 = \frac{4}{7}.$$

Ограничение S_1 имеет вид:

$$S_1 - \frac{6}{7}y_3 - \frac{2}{7}y_1 - 0 \cdot X_3 = \frac{4}{7}.$$

Таблица 13.4

	y_3	y_1	x_3	Свободный член
x_2	−1/7	2/7	−1	4/7
y_2	−15/7	−5/7	−6	4/7
x_1	3/7	1/7	1	16/7
Z	5/7	4/7	4	36/7

Введем это ограничение в табл. 13.4 и получим следующее:

Таблица 13.5

	y_3	y_1	x_3	Свободный член
x_2	−1/7	2/7	−1	4/7
y_2	−15/7	−5/7	−6	4/7
x_1	3/7	1/7	1	16/7
S_1	−6/7	$\boxed{-2/7}$	0	−4/7
Z	5/7	4/7	4	36/7

Применяя теорему, находим допустимое решение, представленное в табл. 13.6.

Таблица 13.6

	y_3	S_1	x_3	Свободный член
x_2	−1	1	−1	0
y_2	0	−5/2	−6	2
x_1	0	1/2	1	2
y_1	3	−7/2	0	2
Z	−1	2	4	4

Полученный план неоптимален (в последней строке имеются отрицательные элементы): после еще одного шага получаем оптимальный план, но нецелочисленный, поэтому составляем дополнительное ограничение по первой строке и получаем табл. 13.7.

Таблица 13.7

	y_3	S_1	x_3	Свободный член
x_2	1/3	−1/6	−1	2/3
y_2	0	−5/2	−6	2
x_1	0	1/2	1	2
y_3	1/3	−7/6	0	2/3
S_2	$\boxed{-1/3}$	−5/6	0	−2/3
Z	1/3	5/6	4	14/3

$$\beta_{11} = 1/3$$
$$\beta_{12} = 5/6$$
$$\beta_{13} = 0$$
$$\beta_1 = 2/3$$

Применяя теорему о нахождении допустимого плана, получаем одновременно оптимальный и целочисленный план.

ПРИЛОЖЕНИЯ

Приложение 1

Процентные точки распределения Стьюдента (t_k)

Число степеней свободы	Уровень значимости					
	20	10	5	2	1,0	0,05
1	3,078	6,314	12,706	31,821	63,657	127,3
2	1,886	2,920	4,303	6,965	9,925	14,089
3	1,638	353	3,152	4,541	5,841	7,453
4	533	132	2,776	3,747	4,604	5,598
5	476	0,15	571	365	0,32	4,773
6	4440	1,943	447	143	3,307	317
7	415	895	365	299	499	29
8	397	860	306	890	355	3,833
9	383	833	262	821	250	690
10	372	812	226	764	169	581
11	363	796	201	718	108	497
12	356	872	179	681	0,55	428
13	350	771	160	650	0,12	372
14	345	716	145	624	2,977	326
15	341	553	131	602	947	286
16	337	746	120	583	921	252
17	333	740	111	567	898	222
18	1,330	734	101	552	878	197
19	1,328	729	93	539	861	174
20	325	725	86	528	845	153
21	323	721	80	518	831	135
22	321	717	74	503	819	119
23	319	714	69	500	807	104
24	318	711	64	492	2,80	91
25	1,316	1,706	2,60	2,485	2,787	3,078
26	215	706	56	479	779	67
27	314	703	52	473	771	57
28	313	701	43	467	763	47
29	311	699	45	462	756	38
30	310	697	42	467	750	20
40	303	684	21	423	704	2,971
60	296	671	000	390	660	2,915
120	289	658	1,980	358	617	2,860
∞	282	645	1,960	326	576	2,807

Приложение 2

Процентные точки распределения Фишера (при $\alpha = 0{,}05$, $f_1 = m$, $f_2 = n - m - 1$)

Число степеней свободы f_2	f_1															
	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	20
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243	244	245	246	246	248,0
2	18,51	19,0	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	14,4	19,41	19,42	19,43	19,41	19,45
3	10,13	9,55	9,25	9,12	9,05	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,71	8,70	8,69	8,66
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,59	5,93	5,91	5,87	5,86	5,84	5,80
5	6,65	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,7	4,68	4,64	4,62	4,6	4,56
6	5,99	5,14	4,76	4,53	4,39	4,26	4,21	4,15	4,10	4,06	4,03	4,0	3,96	3,94	3,92	3,87
7	5,59	4,74	4,35	4,13	3,97	3,87	3,79	3,73	3,68	3,64	3,6	3,57	3,52	3,51	3,49	3,44
8	5,32	4,46	4,07	3,84	3,69	3,53	3,5	3,44	3,39	3,35	3,31	3,28	3,23	3,23	3,2	3,15
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,1	3,07	3,02	3,01	2,93	2,94
10	4,96	4,1	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,86	2,85	2,82	2,77
11	4,84	3,89	3,59	3,36	3,2	3,09	3,01	2,95	2,9	2,35	2,82	2,79	2,74	2,74	2,7	2,65
12	4,75	3,89	3,49	3,26	3,11	3,0	2,91	2,85	2,8	2,75	2,72	2,69	2,64	2,62	2,6	2,54
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,6	2,55	2,53	2,51	2,46
14	4,6	8,74	3,34	3,11	2,96	2,85	2,76	2,7	2,65	2,6	2,56	2,53	2,48	2,46	2,44	2,39
15	4,54	3,68	3,29	3,06	2,9	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,43	2,4	2,39	2,33
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	2,37	2,35	2,33	2,28
17	4,45	3,59	3,20	2,96	2,81	2,7	2,61	2,55	2,49	2,43	2,41	2,38	2,33	2,31	2,29	2,23
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,29	2,27	2,25	2,19
19	4,38	3,52	3,13	2,9	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,26	2,23	2,21	2,16
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	3,31	2,38	2,23	2,2	2,18	2,12
21	4,32	3,47	3,07	2,84	2,63	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,2	2,13	2,15	2,10
22	4,3	3,44	3,05	2,82	2,66	2,55	2,46	2,4	2,34	2,3	2,26	2,28	2,18	2,15	2,13	2,07
23	4,23	3,42	3,03	2,8	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,2	2,14	2,13	2,10	2,05
24	4,26	3,4	3,01	2,78	2,62	2,61	2,42	2,36	2,3	2,25	2,22	2,18	2,13	2,11	2,09	2,03
25	4,24	3,39	2,99	2,76	2,6	2,49	2,4	2,34	2,28	2,24	2,2	2,16	2,11	2,09	2,06	2,01

Приложение 3

Критические значения коэффициента циклической автокорреляции

f	Положительные значения		Отрицательные значения	
	5%	1%	5%	1%
5	0,253	0,297	–0,753	–0,798
6	0,345	0,447	–0,708	–0,863
7	0,370	0,510	–0,674	–0,799
8	0,371	0,531	–0,625	–0,764
9	0,366	0,533	–0,593	–0,737
10	0,360	0,525	–0,564	–0,705
11	0,353	0,515	–0,539	–0,679
12	0,348	0,505	–0,516	–0,655
13	0,341	0,495	–0,497	–0,634
14	0,335	0,485	–0,479	–0,615
15	0,328	0,475	–0,462	–0,597
20	0,299	0,432	–0,399	–0,524
25	0,276	0,398	–0,356	–0,473
30	0,257	0,370	–0,325	–0,433
35	0,242	0,347	–0,300	–0,401
40	0,229	0,326	–0,279	–0,376
45	0,218	0,314	–0,262	–0,356
50	0,208	0,301	–0,248	–0,339
55	0,199	0,289	–0,236	–0,324
60	0,191	0,278	–0,225	–0,310
65	0,184	0,268	–0,216	–0,298
70	0,178	0,259	–0,207	–0,287
75	0,173	0,250	–0,199	–0,276

Приложение 4

Критические значения статистики Дарбина–Уотсона при 5%-ном уровне значимости

T	m									
	1		2		3		4		5	
	d	$d_в$	d	$d_в$	d	$d_в$	d	$d_в$	d	$d_в$
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	1,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	1,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,66	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80

Приложение 5

Таблица критических уровней RS-критерия

Количество наблюдений	Граница RS-критерия	
	нижняя	верхняя
10	2,67	3,69
15	2,96	4,14
20	3,18	4,49
25	3,34	4,71
30	3,47	4,89
...	...	...

Приложение 6

Функция распределения для закона Гаусса

$$\Phi^*(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x} e^{-\frac{\lambda^2}{2}}\cdot d\lambda$$

x	$\Phi^*(x)$	x	$\Phi^*(x)$	x	$\Phi^*(x)$	x	$\Phi^*(x)$	x	$\Phi^*(x)$
−3,00	0,001	−1,18	0,119	−0,38	0,352	0,42	0,663	1,22	0,889
−2,90	0,002	−1,16	0,123	−0,36	0,359	0,44	0,670	1,24	0,893
−2,80	0,003	−1,14	0,127	−0,34	0,367	0,46	0,677	1,26	0,896
−2,70	0,003	−1,12	0,131	−0,32	0,374	0,48	0,684	1,28	0,900
−2,60	0,005	−1,10	0,136	−0,30	0,382	0,50	0,691	1,30	0,903
−2,50	0,006	−1,08	0,140	−0,28	0,390	0,52	0,698	1,32	0,907
−2,40	0,008	−1,06	0,145	−0,26	0,397	0,54	0,705	1,34	0,910
−2,30	0,011	−1,04	0,149	−0,24	0,405	0,56	0,712	1,36	0,913
−2,25	0,012	−1,02	0,154	−0,22	0,413	0,58	0,719	1,38	0,916
−2,20	0,014	−1,00	0,159	−0,20	0,421	0,60	0,726	1,40	0,919
−2,15	0,016	−0,98	0,164	−0,18	0,429	0,62	0,732	1,42	0,922
−2,10	0,018	−0,96	0,169	−0,16	0,436	0,64	0,739	1,44	0,925
−2,05	0,020	−0,94	0,174	−0,14	0,444	0,66	0,745	1,46	0,928
−2,00	0,023	−0,92	0,179	−0,12	0,452	0,68	0,752	1,48	0,931
−1,95	0,026	−0,90	0,184	−0,10	0,460	0,70	0,758	1,50	0,933
−1,90	0,029	−0,88	0,189	−0,08	0,468	0,72	0,764	1,54	0,938
−1,85	0,032	−0,86	0,195	−0,06	0,476	0,74	0,770	1,58	0,943
−1,80	0,036	−0,84	0,200	−0,04	0,484	0,76	0,776	1,62	0,947
−1,75	0,040	−0,82	0,206	−0,02	0,492	0,78	0,782	1,66	0,952
−1,70	0,045	−0,80	0,212	0,00	0,500	0,80	0,788	1,70	0,955
−1,66	0,048	−0,78	0,218	0,02	0,508	0,82	0,794	1,75	0,960
−1,62	0,053	−0,76	0,224	0,04	0,516	0,84	0,800	1,80	0,964
−1,58	0,057	−0,74	0,230	0,06	0,524	0,86	0,805	1,85	0,968
−1,54	0,062	−0,72	0,236	0,08	0,532	0,88	0,811	1,90	0,971
−1,50	0,067	−0,70	0,242	0,10	0,540	0,90	0,816	1,95	0,974
−1,48	0,069	−0,68	0,248	0,12	0,548	0,92	0,821	2,00	0,977
−1,46	0,072	−0,66	0,255	0,14	0,556	0,94	0,826	2,05	0,980
−1,44	0,075	−0,64	0,261	0,16	0,564	0,96	0,831	2,10	0,982
−1,42	0,078	−0,62	0,268	0,18	0,571	0,98	0,836	2,15	0,984
−1,40	0,081	−0,60	0,274	0,20	0,579	1,00	0,841	2,20	0,986
−1,38	0,084	−0,58	0,281	0,22	0,587	1,02	0,846	2,25	0,988
−1,36	0,087	−0,56	0,288	0,24	0,595	1,04	0,851	2,30	0,989
−1,34	0,090	−0,54	0,295	0,26	0,603	1,06	0,855	2,40	0,992
−1,32	0,093	−0,52	0,302	0,28	0,610	1,08	0,860	2,50	0,994
−1,30	0,097	−0,50	0,309	0,30	0,618	1,10	0,864	2,60	0,995
−1,28	0,100	−0,48	0,316	0,32	0,626	1,12	0,869	2,70	0,997
−1,26	0,104	−0,46	0,323	0,34	0,633	1,14	0,873	2,80	0,997
−1,24	0,107	−0,44	0,330	0,36	0,641	1,16	0,877	2,90	0,998
−1,22	0,111	−0,42	0,337	0,38	0,648	1,18	0,881	3,00	0,999
−1,20	0,115	−0,40	0,345	0,40	0,655	1,20	0,885		

Приложение 7

Значения $P_m = \dfrac{a^m}{m!} \cdot e^{-a}$ (распределение Пуассона)

m	a															
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	1,5	2	2,5	3	3,5	4
0	0,904	0,818	0,740	0,670	0,606	0,548	0,496	0,449	0,406	0,367	0,223	0,135	0,082	0,049	0,030	0,018
1	0,090	0,163	0,222	0,268	0,303	0,329	0,347	0,359	0,365	0,367	0,334	0,270	0,205	0,149	0,105	0,073
2	0,004	0,016	0,033	0,053	0,075	0,098	0,121	0,143	0,164	0,183	0,251	0,270	0,256	0,224	0,184	0,146
3	0,000	0,001	0,003	0,007	0,012	0,019	0,028	0,038	0,049	0,061	0,125	0,180	0,213	0,224	0,215	0,195
4	–	0,000	0,000	0,000	0,001	0,002	0,004	0,007	0,011	0,015	0,047	0,090	0,133	0,168	0,188	0,195
5	–	–	–	–	0,000	0,000	0,000	0,001	0,002	0,003	0,014	0,036	0,066	0,100	0,132	0,156
6	–	–	–	–	–	–	–	0,000	0,000	0,000	0,003	0,012	0,027	0,050	0,077	0,104
7	–	–	–	–	–	–	–	–	–	–	0,000	0,003	0,009	0,021	0,038	0,059
8	–	–	–	–	–	–	–	–	–	–	–	0,000	0,003	0,008	0,016	0,029
9	–	–	–	–	–	–	–	–	–	–	–	–	0,000	0,002	0,006	0,013
10	–	–	–	–	–	–	–	–	–	–	–	–	–	0,000	0,002	0,005
11	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0,000	0,001
12	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0,000

Продолжение

m	a															
	4,5	5	5,5	6	6,5	7	7,5	8	8,5	9	9,5	10	12	14	16	18
0	0,011	0,006	0,004	0,002	0,001	0,000	0,000	0,000	0,000	0,000	–	–	–	–	–	–
1	0,049	0,033	0,022	0,014	0,009	0,006	0,004	0,002	0,001	0,001	0,000	0,000	–	–	–	–
2	0,112	0,084	0,061	0,044	0,031	0,022	0,015	0,010	0,007	0,004	0,003	0,002	0,000	–	–	–
3	0,168	0,140	0,113	0,089	0,068	0,052	0,038	0,028	0,020	0,014	0,010	0,007	0,001	0,000	–	–
4	0,189	0,175	0,155	0,133	0,111	0,091	0,072	0,057	0,044	0,033	0,025	0,018	0,005	0,001	0,000	–
5	0,170	0,175	0,171	0,160	0,145	0,127	0,109	0,091	0,075	0,060	0,048	0,037	0,012	0,003	0,001	–
6	0,128	0,146	0,157	0,160	0,157	0,149	0,136	0,122	0,106	0,091	0,076	0,063	0,025	0,008	0,002	0,000
7	0,082	0,104	0,123	0,137	0,146	0,149	0,146	0,139	0,129	0,117	0,103	0,090	0,043	0,017	0,005	0,001
8	0,046	0,065	0,084	0,103	0,118	0,130	0,137	0,139	0,137	0,131	0,123	0,112	0,065	0,030	0,011	0,004
9	0,023	0,036	0,051	0,068	0,085	0,101	0,114	0,124	0,129	0,131	0,130	0,125	0,087	0,047	0,021	0,008
10	0,010	0,018	0,028	0,041	0,055	0,070	0,085	0,099	0,110	0,118	0,123	0,125	0,104	0,066	0,034	0,014
11	0,004	0,008	0,014	0,022	0,032	0,045	0,058	0,072	0,085	0,097	0,106	0,113	0,114	0,084	0,049	0,024
12	0,001	0,003	0,006	0,011	0,017	0,026	0,036	0,048	0,060	0,072	0,084	0,094	0,114	0,098	0,066	0,036
13	0,000	0,001	0,002	0,005	0,008	0,014	0,021	0,029	0,039	0,050	0,061	0,072	0,105	0,105	0,081	0,050
14	–	0,000	0,001	0,002	0,004	0,007	0,011	0,016	0,023	0,032	0,041	0,052	0,090	0,105	0,093	0,065
15	–	–	0,000	0,000	0,001	0,003	0,005	0,009	0,013	0,019	0,026	0,034	0,072	0,098	0,099	0,078
16	–	–	–	–	0,000	0,001	0,002	0,004	0,007	0,010	0,015	0,021	0,054	0,086	0,099	0,088
17	–	–	–	–	–	0,000	0,001	0,002	0,003	0,005	0,008	0,012	0,038	0,071	0,093	0,093
18	–	–	–	–	–	–	0,000	0,000	0,001	0,002	0,004	0,007	0,025	0,055	0,083	0,093
19	–	–	–	–	–	–	–	–	0,000	0,001	0,002	0,003	0,016	0,040	0,069	0,088
20	–	–	–	–	–	–	–	–	–	0,000	0,001	0,001	0,009	0,028	0,055	0,079
21	–	–	–	–	–	–	–	–	–	–	0,000	0,000	0,005	0,019	0,042	0,068
22	–	–	–	–	–	–	–	–	–	–	–	–	0,003	0,012	0,030	0,055
23	–	–	–	–	–	–	–	–	–	–	–	–	0,001	0,007	0,021	0,043
24	–	–	–	–	–	–	–	–	–	–	–	–	0,000	0,004	0,014	0,032
25	–	–	–	–	–	–	–	–	–	–	–	–	–	0,002	0,009	0,023

Приложение 8

Критические точки распределения χ^2 Пирсона

r	$P = 0{,}99$	$P = 0{,}95$	$P = 0{,}90$	$P = 0{,}10$	$P = 0{,}05$	$P = 0{,}01$
1	◡0	◡0	0,01	2,70	3,84	6,63
2	0,02	0,10	0,21	4,60	5,99	9,21
3	0,11	0,35	0,58	6,25	7,81	11,3
4	0,29	0,71	1,06	7,77	9,48	13,2
5	0,55	1,14	1,61	9,23	11,0	15,0
6	0,87	1,63	2,20	10,6	12,6	16,8
7	1,23	2,16	2,83	12,0	14,1	18,4
8	1,64	2,73	3,48	13,3	15,5	20,0
9	2,08	3,32	4,16	14,6	16,9	21,6
10	2,55	3,94	4,86	15,9	18,3	23,2
11	3,05	4,57	5,57	17,2	19,7	24,7
12	3,57	5,22	6,30	18,5	21,0	26,2
13	4,10	5,89	7,04	19,8	22,4	27,6
14	4,66	6,57	7,78	21,0	23,7	29,1
15	5,22	7,26	8,54	22,3	25,0	30,5
16	5,81	7,96	9,31	23,5	26,3	31,9
17	6,40	8,67	10,0	24,7	27,6	33,4
18	7,01	9,39	10,8	25,9	28,8	34,8
19	7,63	10,1	11,6	27,2	30,1	36,1
20	8,26	10,8	12,4	28,4	31,4	37,5
21	8,89	11,5	13,2	29,6	32,6	38,9
22	9,54	12,3	14,0	30,8	33,9	40,2
23	10,1	13,0	14,8	32,0	35,1	41,6
24	10,8	13,8	15,6	33,1	36,4	42,9
25	11,5	14,6	16,4	34,3	37,6	44,3
26	12,1	15,3	17,2	35,5	38,8	45,6
27	12,8	16,1	18,1	36,7	40,1	46,9
28	13,5	16,9	18,9	37,9	41,3	48,2
29	14,2	17,7	19,7	39,0	42,5	49,5
30	14,9	18,4	20,5	40,2	43,7	50,5

Приложение 9

Виды распределений и их основные характеристики

А. Непрерывные распределения

Распределение	Область значений	Плотность распределения	Математическое ожидание	Дисперсия	Мода
Равномерное	$(a,\ b)$	$\dfrac{1}{b-a}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	–
Нормальное	$(-\infty,\ \infty)$	$\dfrac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$	a	σ^2	a
Логарифмически нормальное	$(0,\ \infty)$	$\dfrac{1}{x\sigma\sqrt{2\pi}}\cdot e^{-\frac{(\log x-a)^2}{2\sigma^2}}$	$e^{a+\frac{\sigma^2}{2}}$	$e^{2a+\sigma^2}(e^{\sigma^2}-1)$	$e^{a-\sigma^2}$
Вейбулла	$(0,\ \infty)$	$acx^{\alpha-1}e^{-cx^2}$	$\dfrac{\Gamma\left(1+\frac{1}{\alpha}\right)}{c^{\frac{1}{\alpha}}}$	$\dfrac{\Gamma\left(1+\frac{2}{\alpha}\right)-\Gamma^2\left(a+\frac{1}{\alpha}\right)}{c^{\frac{2}{\alpha}}}$	(при $\alpha > 1$) $\sqrt[\alpha]{\dfrac{\alpha-1}{c\alpha}}$
Гамма	$(0,\ \infty)$	$\dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$	$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$	$(a \le 1)$ $\dfrac{\alpha-1}{\beta}$ $(a > 1)$
Частные случаи: Показательное	$(0,\ \infty)$	$\lambda e^{-\lambda x}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	–

Продолжение

Распределение	Область значений	Плотность распределения	Математическое ожидание	Дисперсия	Мода
χ^2 (хи-квадрат)	$(0, \infty)$	$\dfrac{z^{\frac{k}{2}-1}}{\Gamma\left(\frac{k}{2}\right)2^{\frac{k}{2}}}\cdot e^{-\frac{z}{2}}$	k	$2k$	$k-1$
Бета	$(0, 1)$	$x^{a-1}(1-x)^{b-1}$	$\dfrac{a}{a+b}$	$\dfrac{ab}{(a+b)^2(a+b+1)}$	$\dfrac{a-1}{a+b-2}$
Стьюдента	$(-\infty, \infty)$	$\left[2^{\frac{n-1}{2}}\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}\right]^{-1}\cdot\left(1+\dfrac{x^2}{1}\right)^{-\frac{n+1}{2}}$	0	$\dfrac{1}{\frac{n}{2}-2}$	0
Фишера	$(0, \infty)$	$\dfrac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(\dfrac{n_1}{n_2}\right)^{\frac{n_1}{2}}x^{\frac{n_1}{2}-1}\cdot\left(1+\dfrac{n_1}{n_2}x\right)^{-\frac{n_1+n_2}{2}}$	$\dfrac{n_2}{n_2-2}$	$\dfrac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$	$\dfrac{(n_1-2)n_2}{2n_1^2+n_2}$

Продолжение

Б. Дискретные распределения

Распределение	Область значений	Вероятность	Математическое ожидание	Дисперсия
Биномиальное	0, 1, 2, ..., n	$P_n(m) = C_n^m p^m q^{n-m}$	np	npq
Гипергеометрическое	0, 1, ..., $\min(M, n)$	$P_m = \dfrac{C_M^m C_{N-M}^{n-m}}{C_N^n}$	$n\dfrac{M}{N}$	$\dfrac{M(N-M)n(N-n)}{N^2(N-1)}$
Пуассона	0, 1, 2, ...	$P_m = \dfrac{\lambda^m}{m!}e^{-\lambda}$	λ	λ
Геометрическое	0, 1, 2, ...	$P_m = pq^{m-1}$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$
Отрицательное биномиальное	r, $r+1$, ...	$P_m = C_{m-1}^{r-1} p^r q^{m-r}$	$\dfrac{r}{p}$	$\dfrac{q^r}{p^2}$

Равномерно распределенные случайные числа

10 09 73 25 33	76 52 01 35 86	34 67 35 48 76	80 95 90 91 17
37 54 20 48 05	64 89 47 42 96	24 80 52 40 37	20 63 61 04 02
08 42 26 89 53	19 64 50 93 03	23 20 90 25 60	15 95 33 47 64
99 01 90 25 29	09 37 67 07 15	38 31 13 11 65	88 67 67 43 97
12 80 79 99 70	80 15 73 61 47	64 03 23 66 53	98 95 11 68 77
66 06 57 47 17	34 07 27 68 50	36 69 73 61 70	65 81 33 98 85
31 06 01 08 05	45 57 18 24 06	35 30 34 36 14	86 79 90 74 39
85 26 97 76 02	02 05 16 56 92	68 66 57 48 18	73 05 38 52 47
63 57 33 21 35	05 32 54 70 48	90 55 35 75 48	28 46 82 87 09
73 79 64 57 53	03 52 96 47 78	35 80 83 42 82	60 93 52 03 44
98 52 01 67 67	14 90 56 86 07	22 10 94 05 58	60 97 09 34 33
11 80 50 54 31	39 80 82 77 32	50 72 56 82 48	29 40 52 42 01
83 45 29 96 34	06 28 78 80 83	13 74 67 00 78	18 47 54 06 10
88 68 54 02 00	86 50 75 84 01	36 76 66 79 51	90 36 47 64 93
99 59 46 73 48	87 51 76 49 69	91 82 60 89 28	93 78 56 13 68
65 48 11 76 74	17 46 85 09 50	58 04 77 69 74	73 03 95 71 86
80 12 43 56 35	17 72 70 80 15	45 31 82 23 74	21 11 57 82 53
74 35 09 98 17	77 40 27 72 14	43 23 60 02 10	45 52 16 42 37
69 91 62 68 03	66 25 22 91 48	36 93 68 72 03	76 62 11 39 90
09 89 32 05 05	14 22 56 85 14	46 4 75 67 88	96 29 77 88 22
91 49 91 45 23	68 47 92 76 86	46 16 28 35 54	94 75 08 99 23
80 33 69 45 98	26 94 03 68 58	70 29 73 41 35	53 14 03 33 40
44 10 48 19 49	85 15 74 79 54	32 97 92 65 75	57 60 04 08 81
12 55 07 37 42	11 10 00 20 40	12 86 07 46 97	93 64 48 94 39
63 60 64 93 29	16 50 53 44 84	40 21 95 25 63	43 65 17 70 82
61 19 69 04 46	26 45 74 77 74	51 92 43 37 29	65 39 45 95 93
15 47 44 52 66	95 27 07 99 53	59 36 78 38 48	82 39 61 01 18
94 55 72 85 73	67 89 75 43 87	54 62 24 44 31	91 19 04 25 92
42 48 11 62 13	97 34 40 87 21	16 86 84 87 67	03 07 11 20 59
23 52 37 83 17	73 20 88 98 37	68 93 59 14 16	26 25 22 96 63
04 49 35 24 94	75 24 63 38 24	45 86 25 10 25	61 96 27 93 35
00 54 99 76 54	64 05 18 81 59	96 11 96 38 96	54 59 28 23 91
35 96 31 53 07	26 89 80 93 54	33 35 13 54 62	77 97 45 00 24

Продолжение

59 80 80 83 91	45 42 72 68 42	83 60 94 97 00	13 02 12 48 92
46 05 88 52 36	01 39 09 22 86	77 28 14 40 77	93 91 08 36 47
23 17 90 05 97	87 38 92 52 41	05 56 70 70 07	86 74 31 71 57
69 23 46 14 06	20 11 74 52 04	15 95 66 00 00	18 74 39 24 23
19 56 54 14 30	01 75 87 53 79	40 41 92 15 85	66 67 43 68 06
45 15 51 49 38	19 47 60 72 46	43 66 79 45 43	59 04 79 00 33
94 86 43 19 94	36 16 81 08 51	34 88 88 15 53	01 54 03 54 56
98 08 62 48 26	45 24 02 84 04	44 99 90 88 96	39 09 47 34 07
33 18 51 62 32	41 94 15 09 49	89 43 54 85 81	88 69 54 19 94
80 95 10 04 06	96 38 27 07 74	20 15 12 33 87	25 01 62 52 98
79 75 24 91 40	71 96 12 82 96	69 86 10 25 91	74 85 22 05 38
18 63 33 25 37	98 14 50 65 71	31 01 02 46 74	05 45 56 14 27
74 02 94 39 02	77 55 73 22 70	97 79 01 71 19	52 52 75 80 21
54 17 84 56 11	80 99 33 71 43	05 33 51 29 69	56 12 71 92 55
11 66 44 98 83	52 07 98 48 27	59 38 17 15 39	09 97 33 34 40
48 32 47 79 28	31 24 96 47 10	02 29 53 68 70	32 30 75 75 46
69 07 49 41 38	87 63 79 19 76	35 58 40 44 01	10 51 82 16 15

Приложение 11

Задание для расчетно-графической работы «Моделирование показателей надежности технических систем с использованием аппарата марковских случайных процессов»

Вариант 1

Представим автомобиль как некоторую систему S с дискретными состояниями S_1, S_2, ..., S_n, которая переходит из состояния S_i в состояние $S_j (i = 1, 2, ..., n, j = 1, 2, ..., n)$ под воздействием пуассоновских потоков событий (отказов) с интенсивностями λ_{ij}. Будем рассматривать следующие состояния автомобиля, в которых он может находиться в процессе эксплуатации и которые характеризуются целодневными простоями:

S_1 – исправен, работает;
S_2 – проходит техническое обслуживание;
S_3 – находится в текущем ремонте;
S_4 – находится в капитальном ремонте;
S_5 – проводится замена агрегата;
S_6 – исправен, не работает по организационным причинам;
S_7 – исправен, не работает, выходные и праздничные дни;
S_8 – списывается.

Рассматриваемые состояния S_j автомобиля характеризуются средним числом дней пребывания автомобиля в каждом j-м состоянии $(j = 1, 2, ..., n)$ $Д_j$. Отношение

$$P_j = \frac{Д_j}{Д_K},$$

где $Д_K$ – число календарных дней в году, можно трактовать как вероятность нахождения автомобиля в j-м состоянии P_j.

Вероятности P_j являются функциями пробега автомобиля $P_j(L)$.

Вероятность нахождения автомобиля в состоянии S_1 (исправен, работает) $P_1(L)$ представляет собой коэффициент выпуска автомобиля – один из основных показателей работы автопредприятия.

Возможные переходы автомобиля из состояния S_i в состояние S_j описаны матрицей переходов.

Соответствующие интенсивности потоков событий λ_{ij}, переводящих автомобиль из состояния S_i в состояние S_j, определяются по формулам, приведенным в табл. П.11.1.

Требуется:

1. Построить размеченный граф состояний системы S-автомобиль по заданной матрице переходов.

2. Определить интенсивности λ_{ij}, используя формулы табл. П.11.2.

Таблица П.11.1

Интенсивности перехода

Интенсивность	Формула для расчета	Примечание
1	2	3
Исправен – проходит ТО-2	$\lambda_{12}(L)=\frac{1}{L_{\text{то}}}$	$L_{\text{то}}$ – пробег автомобиля до ТО-2, тыс. км
Исправен – находится в текущем ремонте	$\lambda_{13}(L) = \exp(-0{,}8 + 0{,}08 \cdot L)$	
Исправен – находится в капитальном ремонте	$\lambda_{14}(L) = \exp(-0{,}3 + 0{,}002 \cdot L)$	
Исправен – замена агрегата	$\lambda_{15}(L) = \exp(-0{,}4 + 0{,}004 \cdot L)$	
Исправен – не работает по организационным причинам	$\lambda_{16}(L)=\frac{1}{\ell_{\text{сс}} \cdot T_{\text{орг}}}$	ℓ_{cc} – среднесуточный пробег, тыс. км; $T_{\text{орг}}$ – дни простоя по организационным причинам
Исправен – не работает (праздничные, выходные дни)	$\lambda_{17}(L)=\frac{1}{\ell_{\text{сс}} \cdot T_{\text{вых}}}$	$T_{\text{вых}}$ – праздничные и выходные дни
Исправен – списание автомобиля	$\lambda_{18}(L)=\frac{L-L_0}{S}$	L_0 = 400 тыс. км; S = 300 тыс. км

Исходные данные для определения λ_{ij}:
среднесуточный пробег – 0,2 тыс. км;
среднее время простоя автомобиля в текущем ремонте – 1,5 дня;
среднее время простоя по организационным причинам – 2 дня;
остальные данные выбираются, исходя из профессиональных соображений.

3. Составить систему дифференциальных уравнений Колмогорова и решить ее методом Рунге–Кутта с использованием стандартной программы на ЭВМ при следующих условиях:

а) пределы интегрирования: нижний – 0, верхний – 20;
б) шаг интегрирования – 0,5;
в) начальные условия: $P_1(L) = 1$, $P_j(L) = 0$, $j = 2, 3, ..., n$;

г) результаты вывести на печать в точках 1, 5, 10, 15, 20 с точностью $E = 10^{-3}$.

4. Получить значения коэффициента выпуска автомобиля $P_1(L)$ и построить график зависимости коэффициента выпуска от пробега.

5. Определить влияние на изменение коэффициента выпуска среднесуточного пробега и среднего времени простоя в ремонте; изменение значения этих показателей на 40–60 %. Построить графики.

Примечания:

Интенсивность

$$\lambda_{ij} = \frac{1}{T_i},$$

где T_i – среднее время пребывания в i-м состоянии.

Интенсивность $\lambda_{18} = 0$ для $L \leq 400$ тыс. км.

Для приведения всех интенсивностей перехода λ_{ij} к единым единицам измерения $\left(\frac{1}{\text{день (сутки)}}\right)$ интенсивности перехода из 1-го состояния (автомобиль исправен) во все j-е состояния λ_{ij} при составлении дифференциальных уравнений умножаются на коэффициент ℓ_{cc} (среднесуточный пробег).

Вариант 2

В процессе эксплуатации ЭВМ может рассматриваться как физическая система S, которая в результате проверки может оказаться в одном из следующих состояний:

S_1 – ЭВМ полностью исправна;

S_2 – ЭВМ имеет незначительные неисправности в оперативной памяти, при которых она может решать задачи;

S_3 – ЭВМ имеет существенные неисправности и может решать ограниченный класс задач;

S_4 – ЭВМ полностью вышла из строя;

S_5 – ЭВМ находится на профилактике;

S_6 – ЭВМ не работает по организационным причинам;

S_7 – ЭВМ не работает, праздничные и другие нерабочие дни;

S_8 – ЭВМ списывается.

Рассматриваемые состояния S_j ЭВМ характеризуются средним временем T_j пребывания ЭВМ в каждом j-м состоянии.

Отношение

$$P_j = T_j/T,$$

где T – возможное время работы ЭВМ в данный период (месяц, квартал, год и т. д.), можно трактовать как вероятность P_j нахождения ЭВМ в j-м состоянии.

Вероятности P_j являются функциями времени $P_j(t)$. Вероятность нахождения ЭВМ в состоянии $P(t) = P_1(t) + P_2(t)$ может быть истолкована как вероятность безотказной работы ЭВМ, т. е. как один из показателей надежности технической системы.

Возможные переходы системы S-ЭВМ из состояния S_i в состояние $S_j(i = 1, 2, ..., n; j = 1, 2, ..., n)$ описаны матрицей переходов.

Соответствующие интенсивности потоков событий λ_{ij}, переводящих ЭВМ из состояния S_i в состояние S_j, определяются по формулам, приведенным в табл. П.11.2.

Требуется:

1. Построить размеченный граф состояний системы S-ЭВМ по заданной матрице переходов.

Таблица П.11.2

Интенсивности перехода

Интенсивность	Формула для расчета	Примечание
$\lambda_{12}(t)$	$\lambda_{12}(t)=\dfrac{0,25}{T_н}$	$T_н$ – среднее время работы ЭВМ до появления незначительной неисправности; $T_н = 0,1 \cdot T$, где T – общее возможное время работы ЭВМ за данный период
$\lambda_{13}(t)$	$\lambda_{13}(t) = 0,25 \cdot \exp(-0,8 + 0,08 \cdot t)$	
$\lambda_{14}(t)$	$\lambda_{14}(t) = 0,22 \cdot \exp(-0,3 + 0,002 \cdot t)$	
$\lambda_{15}(t)$	$\lambda_{15}(t) = 0,24 \cdot \exp(-0,4 + 0,004 \cdot t)$	
$\lambda_{16}(t)$	$\lambda_{16}(t)=\dfrac{1}{T_{орг}}$	$T_{орг}$ – среднее время простоя ЭВМ по организационным причинам
$\lambda_{17}(t)$	$\lambda_{17}(t)=\dfrac{1}{T_{пр}}$	$T_{пр}$ – среднее время простоя в праздничные дни
$\lambda_{18}(t)$	$\lambda_{18}(t)=\dfrac{(t-t_0)}{S}$	$t_0 = 1\,200$ тыс. ч $S = 72\,000$ тыс. ч $\lambda_{18}(t) = 0$ при $t \leq 1\,200$ тыс. ч

2. Определить интенсивности λ_{ij}, используя формулы табл. П.11.2. Интенсивности λ_{ij} определяются по формулам $\lambda_{ij} = 1/T_i$, где T_i – среднее время пребывания системы S в i-м состоянии за данный период (месяц, квартал, год и т. п.).

3. Составить систему дифференциальных уравнений Колмогорова и решить ее методом Рунге–Кутта с использованием стандартной программы на ЭВМ при следующих условиях:

а) пределы интегрирования: нижний – 0, верхний – 50;

б) шаг интегрирования – 0,5;

в) начальные условия: $P_1(t) = 1$, $P_j(t) = 0, j = 2, 3, ..., n$;

г) результаты вывести на печать в точках 1, 5, 10, 15, ...,50 с точностью $E = 10^{-3}$.

4. Получить значения вероятности безотказной работы ЭВМ $P(t)$ и построить график зависимости вероятности от времени.

Для решения необходимо использовать следующие варианты исходных данных

(0 – нет перехода; 1 – возможен переход)

№ 1

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ

	1	2	3	5	6	8
1	0	1	1	1	1	1
2	1	0	0	0	0	0
3	1	1	0	1	1	0
5	1	1	1	0	1	1
6	1	1	1	1	0	1
8	0	0	0	0	0	0

№ 2

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ

	1	2	3	4	8
1	0	1	1	1	1
2	1	0	1	0	0
3	1	1	0	0	0
4	1	1	0	0	0
8	0	0	0	0	0

№ 3

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ

	1	2	3	5	7	8
1	0	1	1	1	1	1
2	1	0	1	1	1	0
3	1	1	0	1	1	0
5	1	1	0	0	1	0
7	1	1	1	1	0	1
8	0	0	0	0	0	0

№ 4

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	3	5	7	8
1	0	1	1	1	1	1
2	1	0	1	0	1	1
3	1	1	0	1	1	0
5	1	1	0	0	1	0
7	1	1	1	1	0	1
8	0	0	0	0	0	0

№ 5

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	3	4	7	8
1	0	1	1	1	1	1
2	1	0	0	0	0	0
3	1	1	0	1	1	0
4	1	1	0	0	0	0
7	1	1	1	1	0	1
8	0	0	0	0	0	0

№ 6

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	5	8
1	0	1	1	1
2	1	0	1	0
5	1	1	0	0
8	0	0	0	0

№ 7

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	3	4	8
1	0	1	1	1	1
2	1	0	0	0	0
3	1	0	0	0	0
4	1	1	1	0	0
8	0	0	0	0	0

№ 8

Состояния автомобиля — Матрица возможных переходов
Состояния ЭВМ

	1	2	3	7
1	0	1	1	1
2	1	0	0	0
3	1	0	0	0
7	1	1	1	0

№ 9

Состояния автомобиля — Матрица возможных переходов
Состояния ЭВМ

	1	2	3	4	7
1	0	1	1	1	1
2	1	0	0	0	0
3	1	1	0	1	1
4	1	1	1	0	1
7	1	1	1	1	0

№ 10

Состояния автомобиля — Матрица возможных переходов
Состояния ЭВМ

	1	2	3	5	6
1	0	1	1	1	1
2	1	0	0	1	1
3	1	0	0	1	1
5	1	1	0	0	1
6	1	1	1	1	0

№ 11

Состояния автомобиля — Матрица возможных переходов
Состояния ЭВМ

	1	2	3	4	6	7
1	0	1	1	1	1	1
2	1	0	1	1	1	1
3	1	0	0	0	1	1
4	1	1	1	0	1	1
6	1	1	1	1	0	1
7	1	1	1	1	1	0

№ 12

Состояния автомобиля
Состояния ЭВМ

Матрица возможных переходов

	1	2	3	8
1	0	1	1	1
2	1	0	1	0
3	1	1	0	0
8	0	0	0	0

№ 13

Состояния автомобиля
Состояния ЭВМ

Матрица возможных переходов

	1	2	3	5	8
1	0	1	1	1	1
2	1	0	1	0	0
3	1	1	0	1	0
5	1	1	0	0	0
8	0	0	0	0	0

№ 14

Состояния автомобиля
Состояния ЭВМ

Матрица возможных переходов

	1	3	5	7
1	0	1	1	1
3	1	0	1	1
5	1	1	0	1
7	0	0	0	0

№ 15

Состояния автомобиля
Состояния ЭВМ

Матрица возможных переходов

	1	3	5	6	7
1	0	1	1	1	1
3	1	0	1	1	1
5	1	0	0	1	1
6	1	1	1	0	1
7	1	1	1	1	0

№ 16

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	5	7	8
1	0	1	1	1	1
2	1	0	1	1	1
5	1	1	0	1	0
7	1	1	1	0	0
8	0	0	0	0	0

№ 17

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	5	7
1	0	1	1	1
2	1	0	1	1
5	1	1	0	1
7	1	1	1	0

№ 18

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	5	8
1	0	1	1	1
2	1	0	1	0
5	1	1	0	0
8	0	0	0	0

№ 19

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	5	6	8
1	0	1	1	1	1
2	1	0	1	1	0
5	1	1	0	1	0
6	1	1	1	0	0
8	0	0	0	0	0

№ 20

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	5	6	7	8
1	0	1	1	1	1	1
2	1	0	1	1	1	0
5	1	1	0	1	1	0
6	1	1	1	0	1	0
7	1	1	1	1	0	0
8	0	0	0	0	0	0

№ 21

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	4	8
1	0	1	1	1
2	1	0	1	0
4	1	1	0	0
8	0	0	0	0

№ 22

Состояния автомобиля — Матрица возможных переходов

Состояния ЭВМ	1	2	4	7	8
1	0	1	1	1	1
2	1	0	1	1	0
4	1	1	0	1	0
7	1	1	1	0	0
8	0	0	0	0	0

Пример. Предположим, что система S-автомобиль может находиться в следующих состояниях: S_1, S_3, S_5, S_7. Возможные переходы системы S из состояния в состояние указаны в матрице:

Матрица возможных переходов

	1	3	5	7
1	0	1	1	1
3	1	0	1	1
5	1	1	0	1
7	0	0	0	0

(0 – нет перехода; 1 – возможен переход).

Используя матрицу возможных переходов, построим размеченный граф состояний системы S-автомобиль (рис. П.11.1).

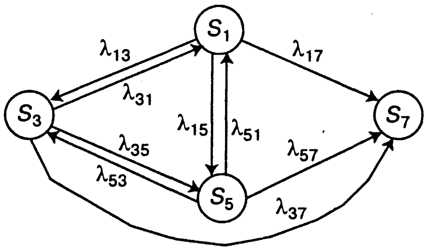


Рис. П.11.1. Граф состояний системы S

Определим интенсивности λ_{ij}, используя табл.П.11.1.

Исходные данные:

среднесуточный пробег $\ell_{cc} = 0{,}25$ тыс. км;

среднее время простоя автомобиля в текущем ремонте $T_3 = 1$ дню. Остальные данные выбираются, исходя из профессиональных соображений:

количество выходных и праздничных дней

$$T_{вых} = 60 \text{ дней};$$

среднее время замены агрегата

$$T_5 = 5{,}0 \text{ дням}.$$

$$\lambda_{13} = \exp(-0{,}8 + 0{,}08 \cdot L);$$

$$\lambda_{15} = \exp(-0{,}4 + 0{,}004 \cdot L);$$

$$\lambda_{17} = \frac{1}{\ell_{cc} \cdot T_{вых}} = \frac{1}{0{,}25 \cdot 60} = 0{,}0667;$$

$$\lambda_{31} = \frac{1}{T_3} = \frac{1}{1} = 1;$$

$$\lambda_{35} = 0{,}1;$$

$$\lambda_{37} = 0{,}01;$$

$$\lambda_{51} = \frac{1}{T_5} = \frac{1}{5} = 0{,}2;$$

$$\lambda_{53} = 0{,}02;$$

$$\lambda_{57} = 0{,}002.$$

Составим систему дифференциальных уравнений Колмогорова для вероятностей состояний P_i, где $i = 1, 3, 5, 7$:

$$\frac{dP_1}{dL} = -P_1 \cdot (\lambda_{13} + \lambda_{15} + \lambda_{17}) \cdot \ell_{cc} + \lambda_{31} \cdot P_3 + \lambda_{51} \cdot P_5;$$

$$\frac{dP_3}{dL} = -P_3 \cdot (\lambda_{31} + \lambda_{35} + \lambda_{37}) + \lambda_{13} \cdot P_1 \cdot \ell_{cc} + \lambda_{53} \cdot P_5;$$

(П. 11.1)

$$\frac{dP_5}{dL} = -P_5 \cdot (\lambda_{51} + \lambda_{53} + \lambda_{57}) + \lambda_{15} \cdot P_1 \cdot \ell_{cc} + \lambda_{35} \cdot P_3;$$

$$\frac{dP_7}{dL} = \lambda_{17} \cdot P_1 \cdot \ell_{cc} + \lambda_{37} \cdot P_3 + \lambda_{57} \cdot P_5.$$

Решим эту систему методом Рунге—Кутта с использованием стандартной программы на ЭВМ при следующих условиях:

а) пределы интегрирования: нижний – 0, верхний – 35;

б) шаг интегрирования – 0,5;

в) начальные условия: $P_1(L) = 1$; $P_3(L) = 0$; $P_5(L) = 0$; $P_7(L) = 0$;

г) результаты выведения на печать с точностью $E = 10^{-3}$ в точках 1, 5, ..., 35.

Для решения задач можно воспользоваться программным продуктом MATHCAD 6.0 PLUS.

Получим значение коэффициента выпуска автомобиля $P_1(L)$ в зависимости от пробега автомобиля L (табл. П. 11.3).

Таблица П.11.3

Значения коэффициента выпуска автомобиля

L	$P_1(L)$	$P_3(L)$	$P_5(L)$	$P_7(L)$
1	0,787	0,0622	0,136	0,0152
5	0,512	0,0824	0,345	0,0614
10	0,419	0,0972	0,376	0,108
15	0,364	0,123	0,363	0,150
20	0,312	0,156	0,344	0,188
25	0,260	0,193	0,323	0,224
30	0,162	0,269	0,218	0,288

Для того чтобы определить влияние на изменение коэффициента выпуска среднесуточного пробега ℓ_{cc} и среднего времени простоя в ремонте T_3, изменим эти показатели. Увеличим ℓ_{cc} и T_3 на 50%. Тогда

$$\ell_{cc} = 0{,}375 \text{ тыс. км}; \; T_3 = 1{,}5 \text{ дня}.$$

С изменением этих показателей:

$$\lambda_{17} = 0{,}0444;$$

$$\lambda_{31} = \frac{1}{T_3} = \frac{2}{3} = 0{,}6667;$$

$$\lambda_{35} = 0{,}0667;$$

$$\lambda_{37} = 0{,}0067.$$

Остальные λ_{ij} остаются без изменения.

Решим систему методом Рунге–Кутта с использованием стандартной программы на ПЭВМ с учетом новых значений интенсивностей и ℓ_{cc}.

Получим новые значения коэффициента выпуска автомобиля $P_1(L)$ (табл. П.11.4).

Таблица П.11.4

Значения коэффициента выпуска автомобиля

L	$P_1(L)$	$P_3(L)$	$P_5(L)$	$P_7(L)$
1	0,694	0,102	0,189	0,0144
5	0,387	0,142	0,418	0,0534
10	0,314	0,162	0,433	0,0915
15	0,266	0,200	0,408	0,126
20	0,221	0,246	0,375	0,158
25	0,177	0,295	0,341	0,187
30	0,137	0,341	0,308	0,214
35	0,103	0,383	0,276	0,239

Вывод. При увеличении среднего времени простоя автомобиля в ремонте и среднесуточного пробега на 50% коэффициент выпуска уменьшается. Полученная математическая модель (П.11.1) функционирования автомобиля позволяет проследить влияние различных условий эксплуатации на коэффициент выпуска автомобиля.

Метод Жордана–Гаусса

Пусть задача линейного программирования записана в каноническом виде:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases}$$

$$Z_{\min} = c_1x_1 + c_2x_2 + \ldots + c_nx_n,$$

$$b_i \geq 0, \quad i = \overline{1,\ m}, \quad x_j \geq 0, \quad j = \overline{1,\ n}.$$

Для компактности составим для записанной задачи жордановскую таблицу, в которой припишем строку, соответствующую целевой функции Z:

$$\left|\begin{array}{cccc|c|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 & \Sigma_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m & \Sigma_m \\ c_1 & c_2 & \cdots & c_n & 0 & \Sigma_n \end{array}\right|$$

Алгоритм метода

1. Записать задачу в форме жордановской таблицы, при этом все элементы столбца свободных членов b_i должны быть неотрицательны $b_i \geq 0$, $i = \overline{1, m}$. Уравнения системы, в которых свободные члены отрицательны, предварительно нужно умножить на -1.

2. Таблицу преобразуем шагами жордановских исключений. При этом на каждом разрешающем шаге может быть выбран любой столбец, содержащий хотя бы один положительный элемент. Строка целевой функции на выбор разрешающих столбцов влияния не оказывает.

3. Разрешающая строка определяется по наименьшему из отношений свободных членов к положительным элементам разрешающего столбца.

4. В процессе преобразований вычеркиваются строки, состоящие из одних нулей.

5. Если в процессе преобразований встречается строка, все элементы которой нули, а свободный член отличен от нуля, то задача не имеет решения. Если встретится строка, в которой кроме свободного члена других положительных элементов нет, то говорят, что задача не имеет положительных решений.

Пример. Найти опорный план задачи линейного программирования.

$$\min L = x_1 + 2x_2 + 3x_3 + 4x_4$$

$$\begin{cases} x_1 + x_2 - 2x_3 + x_4 = 2 \\ x_1 - 2x_2 + x_3 + 2x_4 = 4 \\ x_1 + 2x_2 + 2x_3 + 2x_4 = 8, \end{cases}$$

$$x_j \geq 0; \quad j = \overline{1,\ 4}.$$

Решение

1. Строим первую жордановскую таблицу, последняя строка которой является строкой целевой функции (табл. П.12.1)

Таблица П.12.1

x_1	x_2	x_3	x_4	СЧ
(1)	1	–2	1	2
1	–2	1	2	4
1	2	2	2	8
–1	–2	–3	–4	0

2. Табл. П.12.1 и все последующие таблицы преобразуем шагами жордановских преобразований. При этом на каждом шаге (табл. П.12.1, табл. П.12.2, табл. П.12.3) разрешающим выбираем любой столбец, содержащий хотя бы один положительный элемент. Разрешающая строка выбирается по наименьшему положительному отношению свободных членов к элементам разрешающего столбца. На пересечении разрешающего столбца и разрешающей строки находится разрешающий элемент: в табл. П.12.1 он равен 1, в табл. П.12.2 – 1, в табл. П.12.3 – 5.

3. Переходим к следующей таблице (табл. П.12.2), в которой столбец, соответствующий разрешающему столбцу табл. П.12.1, является нулевым, а элементы строки, соответствующие элементы разрешающей строки табл. П.12.1, получаются путем деления соответствующих элементов табл. П.12.1 на разрешающий элемент. Остальные элементы табл. П.12.2 вычисляются по правилу прямоугольника: искомый элемент из табл. П.12.2 будет равен соответствующему элементу табл. П.12.1 минус дробь, в числителе которой стоит произведение элементов из двух противоположных вершин прямоугольника, а в знаменателе – разрешающий элемент.

В одной и той же строке дважды выбирать разрешающий элемент нельзя.

4. Когда после ряда преобразований будет получено столько же нулевых столбцов, сколько и уравнений в системе ограничений, преоб-

разования необходимо закончить. В нашем случае преобразования заканчиваются на 4-м шаге (табл. П.12.4), так как число нулевых столбцов в последней таблице равно 3, и число уравнений в системе ограничений равно 3.

Таблица П.12.2

x_1	x_2	x_3	x_4	СЧ
1	1	–2	1	2
0	–3	3	(1)	2
0	1	4	1	6
0	–1	–5	–3	2

Таблица П.12.3

x_1	x_2	x_3	x_4	СЧ
1	4	–1	0	0
0	–3	3	1	2
0	4	(5)	0	4
0	–10	–8	0	8

Таблица П.12.4

x_1	x_2	x_3	x_4	СЧ
1	24/5	0	0	4/5
0	–11/5	0	1	14/5
0	4/5	1	0	4/5
0	–18/5	0	0	72/5

Из табл. П.12.4 получим начальный опорный план:

$$L\min = -\frac{18}{5}x_2 + \frac{72}{5};$$

$$\begin{cases} x_1 + \dfrac{24}{5}x_2 = \dfrac{4}{5}; \\ -\dfrac{11}{5}x_2 + x_4 = \dfrac{14}{5}; \\ \dfrac{4}{5}x_2 + x_3 = \dfrac{4}{5}. \end{cases}$$

Приняв $x_2 = 0$, получили первое допустимое решение:

$$L\min = \frac{72}{5} \quad \text{при} \quad x_1 = \frac{4}{5}; \quad x_2 = 0; \quad x_3 = \frac{4}{5}; \quad x_4 = \frac{14}{5}.$$

Дополнительная литература

Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики: Учебник. – М.: ЮНИТИ, 1998.

Акулич И. Л. Математическое программирование в примерах и задачах. – М.: Высшая школа, 1993.

Бережной В. И. Методы и модели логистического подхода к управлению автотранспортным предприятием. – Ставрополь: Интеллект-сервис, 1997.

Бережной В. И., Бережная Е. В. Методы и модели управления материальными потоками микрологистической системы автопредприятия. – Ставрополь: Интеллект-сервис, 1996.

Вагнер Г. Основы исследования операций. – М.: Мир, 1972.

Васильков Ю. В., Василькова И. Н. Компьютерные технологии в математическом моделировании: Учеб. пособие для вузов. – М.: Финансы и статистика, 2004.

Вентцель Н. С. Исследование операций. – М.: Советское радио, 1972.

Вентцель Е. С., Овчарова Л. А. Прикладные задачи теории вероятностей. – М.: Радио и связь, 1983.

Горчаков А. А., Орлова И. В. Компьютерные экономико-математические модели. – М.: Компьютер, 1995.

Дубров А. М. и др. Моделирование рисковых ситуаций в экономике и бизнесе: Учеб. пособие для вузов. / А. М. Дубров, Б. А. Лагоша, Е. Ю. Хрусталев, Т. П. Барановская. Изд. 2-е, перераб. и доп. – М.: Финансы и статистика, 2003.

Елисеева И. И., Юзбашев М. М. Общая теория статистики. – 5-е изд., перераб. и доп. – М.: Финансы и статистика, 2004.

Калинина В. Н., Панкин В. Ф. Математическая статистика. – М.: Высшая школа, 1994.

Лабскер Л. Г., Бабешко Л. О. Теория массового обслуживания в экономической сфере: Учеб. пособие. – М.: ЮНИТИ, 1998.

Лукинский В. С., Зайцев Е. И., Бережной В. И. Модели и алгоритмы управления обслуживанием и ремонтом автотранспортных средств. – СПб.: СПГИЭА, 1997.

Математические методы и модели в экономической деятельности: Учебник – М.: Финансы и статистика, 2001.

Практикум по эконометрике / Под ред. И. И. Елисеевой. – М.: Финансы и статистика, 2004.

Пьяных С. М. Экономико-математические методы оптимального планирования работы речного транспорта. – М.: Транспорт, 1998.

Рабочая книга по прогнозированию / Под ред. И. В. Бестужева. – М.: Мысль, 1985.

Таха Х. Введение в исследование операций: В 2-х кн. – М.: Мир, 1995.

Теория статистики: Учебник / Под ред. Р. А. Шмойловой. – М.: Финансы и статистика, 2004.

Федосеев В. В. Экономико-математические методы и модели в маркетинге. – М.: Финстатинформ, 1996.

Ферстер Э., Ренц Б. Методы корреляционного и регрессионного анализа. – М.: Финансы и статистика, 1983.

Четыркин Е. М. Статистические методы прогнозирования. – М.: Финансы и статистика, 1979.

Чуев Ю. В., Михайлов Ю. Б., Кузьмин В. И. Прогнозирование количественных характеристик процессов. – М.: Советское радио, 1975.

Шикин Е. В., Чхартишвили А. Г. Математические методы и модели в управлении.: Учеб. пособие. – М.: Дело, 2000.

Экономико-математические методы и модели в управлении морским транспортом. – М.: Транспорт, 1998.

Экономико-математические методы и прикладные модели: Учеб. пособие / Под ред. В. В. Федосеева. – М.: ЮНИТИ, 2000.

Эконометрика: Учебник / Под ред. И. И. Елисеевой. – М.: Финансы и статистика, 2004.

Предметный указатель

А

Автокорреляция 180
Адекватность модели 180
Алгоритм симплекс-метода 216
– отыскания опорного плана 220, 224
– метода потенциалов 272
Анализ модели на чувствительность 208
Аппроксимация 156
Аппроксимирующая функция 168

Б

Базисное решение задачи линейного программирования 215, 220
Базисные переменные 216
Биномиальное распределение 24

В

Вероятность перехода 44
– события 8
– состояния 43
– отказа в обслуживании заявки 88
Верхняя цена игры 330
Весовой коэффициент прогноза 186
Виды регрессий 139
– корреляции 140
Входной поток требований 83
Выравнивание 167
Выбор математической модели прогнозирования 179

Г

Гамма-распределение 31
Гармонический фильтр 162
Гистограмма распределения 19, 38
Граф состояний 42, 47, 61
– автомобиля 63
Границы доверительного интервала 185

Д

Двойственная задача 238
Двойственные оценки 247
Дисперсия случайной величины 15
Дисциплины очереди 83
Дискретные законы распределения 24
Длина очереди 85
Доминируемая строка 332
Доминируемый столбец 332
Доминирующий столбец 332
Допустимое решение 220
Достаточность оценки 149

З

Задача об ассортименте продукции 191
– об использовании мощностей оборудования 193

– минимизации дисбаланса на линии сборки 194
– о диете 196
– составление жидких смесей 197
– о раскрое или минимизации обрезков 198
Задачи анализа на чувствительность оптимального решения 247
– корреляционного анализа 141
– регрессионного анализа 141
Закон распределения 10, 24
Заявка на обслуживание (требование) 82

И

Игра вида mxn 335
– с нулевой суммой 328
– конечная 328
– двух лиц 328
Интенсивность потока событий 49, 53
– перехода 48, 55
Интерполяция 141
Искусственные переменные 221

К

Качество модели 179
Классификация моделей массового обслуживания 85
Корреляция 140
Коэффициент автокорреляции 181
– асимметрии 182
– вариации 17
– детерминации 147
– корреляции 146
– множественной корреляции 147
– эксцесса 182
– эластичности 157
Комбинированный прогноз 186
Корреляционное отношение 147
Критерий Вальда 324
– Гурвица 326
– Дарбина–Уотсона 180
– Лапласа 322
– наиболее вероятного исхода 320
– ожидаемого значения 317
– предельного уровня 319
– «пиков» и «впадин» 181
– Стьюдента 154, 183
– Сэвиджа 325
– Фишера 155, 156
– хи-квадрат 35
–эффективности функционирования системы массового обслуживания 85

М

Марковский процесс 42
– с дискретными состояниями 42
Марковские цепи 43, 48
Матрица переходных вероятностей 44
Медиана случайной величины 15
Мера мультиколлинеарности 152
Метод динамик средних 64
– Монте-Карло 118
– авторегрессии 162
– аддитивной оптимизации 312
– наименьших квадратов 169
– потенциалов 272
– северо-западного угла 273
– экспоненциального сглаживания 170
Механизм обслуживания 84
Многоканальная система массового обслуживания с отказами 97
– с ожиданием 101
Мода случайной величины 15
Модель Брауна 172
– замкнутой системы массового обслуживания 104
– «гармонический фильтр» 162
Моделирование случайных величин 121

– случайных событий 125
– случайных функций 128
– СМО с отказами 128
Мультиколлинеарность 151

Н

Непрерывная цепь Маркова 48
Неопределенность 320
Непрерывные распределения вероятностей 28
Несовместные события 6
Несмещенность оценки 149
Нижняя цена игры 328
Нормальное распределение 28
Нормировочное условие 50
Нормализация критериев 313

О

Область допустимых решений 202
Общая дисперсия 147
Одноканальная СМО с отказами 86
– с ожиданием 89
Оптимальное решение 229
Ординарность потока событий 53
Остаточная дисперсия 147
Отбор главных факторов 151
– видов аппроксимирующей функции 168
Ошибка аппроксимации 157
Оценка точности 169
– математической модели прогноза 169

П

Параметр сглаживания 174
Плотность распределения 12
– вероятности перехода 48
– восстановления 132
Полигон распределения 19
Показательное распределение 32
Поток событий 48
Правило построения двойственной пары 238
Признаки оптимальности двойственной пары 239
Простейший поток событий 53
Простейшая одноканальная модель 86
Процесс гибели и размножения 55
Пропускная способность системы 87–88
Процедура построения вариационного ряда 18
Прогноз 158
Прогнозирование 158
Пуассоновский поток событий 53

Р

Равномерное распределение 33
Распределение Пуассона 26
– Эрланга 31
Регрессия 139
Регрессионный анализ 141
Риск 309
Решение задачи методом статистического моделирования 120
RS-критерий 183
Ряд распределения 18

С

Свойства случайного потока событий 53
– вероятностей событий 8
Сглаживание 165
Седловая точка 331
Система дифференциальных уравнений 49, 64, 65
Симплекс-метод 215
Симплекс-таблица 216
Случайная величина 10
– функция 41
Событие 6
Состоятельность оценок 149
Средняя случайной величины 14

Среднее квадратическое отклонение 16
– линейное отклонение 184
Среднее число каналов, занятых обслуживанием 106
Статус ресурсов 230
Стратегия игры 328
– смешанная 333
– чистая 329
Сумма событий 6

Т

Таблица моделей прогноза 164
Теорема двойственности 241
Тренд 159
Транспортная задача 200, 270
– закрытая 271
– усложненная 278
Транспортная модель 271

У

Управляемые факторы 152
Уравнение Колмогорова 49
Условия существования финальных вероятностей 52
– неопределенности 308
– неотрицательности переменных 189
– риска 309
Условные вероятности 9
Установившееся состояние системы 49

Ф

Финальные вероятности состояний 49, 51
Фильтрация исходного временно́го ряда 165
Функция распределения 11
– восстановления 132
– гиперболическая 168
– квадратическая 168
– линейная 167
– регрессии 139
– степенная 168
– экспоненциальная 167

Ц

Целевая функция 188
Ценность ресурса 230
Цена игры 329–330
Цепь Маркова 43, 48
– Маркова эргодическая 50

Э

Экономическая интерпретация регрессионного анализа 157
Экстраполяция 141
Экспоненциальные средние 172
Элемент симплекс-таблицы разрешающий 217
Эргодический случайный процесс 50
Эффективность оценок 149

Оглавление

Предисловие 3

Раздел I. Вероятностно-статистические методы моделирования экономических систем 5

Глава 1. **Основы вероятностных методов анализа и моделирования экономических систем** . . 5

1.1. Элементарные понятия о случайных событиях, величинах и функциях 5

1.2. Числовые характеристики случайных величин 14

1.3. Статистическая оценка законов распределения случайных величин 17

1.4. Основные законы распределения случайных величин 24

1.5. Выбор теоретического закона распределения случайной величины 34

Задачи 39

Глава 2. **Моделирование экономических систем с использованием марковских случайных процессов** 41

2.1. Основные понятия марковских процессов 41

2.2. Марковские цепи 43

2.3. Непрерывные цепи Маркова 48

2.4. Моделирование работы подвижного состава с использованием марковских случайных процессов 62

Задачи 73

Глава 3. **Моделирование систем массового обслуживания** 82

3.1. Компоненты и классификация моделей массового обслуживания 82

3.2. Определение характеристик систем массового обслуживания 86

Задачи 112

Глава 4. **Статистическое моделирование экономических систем** 118

4.1. Теоретические основы метода 118

4.2. Моделирование систем массового обслуживания с использованием метода Монте-Карло 128

4.3. Моделирование потоков отказов элементов сложных технических систем 131

Задачи 136

Глава 5. **Методы и модели корреляционно-регрессионного анализа** 138
5.1. Общие сведения 138
5.2. Исходные предпосылки регрессионного анализа и свойства оценок 148
5.3. Этапы построения многофакторной корреляционно-регрессионной модели 149

Глава 6. **Методы и модели прогнозирования временных рядов экономических показателей** 158
6.1. Основные положения и понятия в прогнозировании временных рядов 158
6.2. Характеристика методов и моделей прогнозирования показателей работы предприятий 160
6.3. Прогнозирование с помощью методов экстраполяции 165
6.4. Прогнозирование на основе временных рядов с использованием пакета программ для персональных ЭВМ 186

Раздел II. Оптимизационные методы и модели в управлении экономическими системами 188

Глава 7. **Линейное программирование** 188
7.1. Задачи линейного программирования 189
7.2. Построение экономико-математических моделей задач линейного программирования 191
7.3. Графическое решение задачи линейного программирования 202
7.4. Анализ моделей на чувствительность 208
7.5. Симплекс-метод 215
7.6. Методы нахождения опорного решения задачи линейного программирования 220
7.7. Экономическая интерпретация решения задачи линейного программирования 228
7.8. Двойственные задачи линейного программирования 237
7.9. Экономико-математический анализ полученных оптимальных решений 247
Задачи 253

Глава 8. **Транспортные задачи линейного программирования** 270
8.1. Постановка задачи 270
8.2. Алгоритм метода потенциалов 272
8.3. Усложненные задачи транспортного типа 278
8.4. Метод Фогеля 286
8.5. Транспортная задача в сетевой постановке 288
8.6. Доставка груза в кратчайший срок 292
Задачи 294

Глава 9. **Теория игр и принятия решений** 308
9.1. Основные понятия 308
9.2. Принятие решений в условиях полной определенности 311
9.3. Принятие решений в условиях риска 317
9.4. Принятие решений в условиях неопределенности 320
9.5. Теория игр 328
Задачи 339

Глава 10. **Нелинейное программирование** 346
10.1. Геометрическая интерпретация задач нелинейного программирования 346
10.2. Метод множителей Лангранжа 349
10.3. Градиентный метод 351
10.4. Многоцелевые задачи линейного программирования 359
Задачи 364

Глава 11. **Динамическое программирование** 366
11.1. Общие понятия о динамическом программировании 366
11.2. Задача о замене оборудования 368

Глава 12. **Параметрическое программирование** 375
12.1. Постановка задачи и геометрический метод ее решения 375
12.2. Аналитический метод решения задач параметрического программирования 378

Глава 13. **Целочисленное программирование** 386

Приложения 393

Приложение 1. Процентные точки распределения Стьюдента (t_k) 393
Приложение 2. Процентные точки распределения Фишера (при $\alpha = 0{,}05, f_1 = m, f_2 = n - m - 1$) 394

Приложение 3. Критические значения коэффициента циклической автокорреляции 395

Приложение 4. Критические значения статистики Дарбина–Уотсона при 5%-ном уровне значимости 396

Приложение 5. Таблица критических уровней RS-критерия 396

Приложение 6. Функция распределения для закона Гаусса $\Phi^*(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{\lambda^2}{2}}\cdot d\lambda$ 397

Приложение 7. Значения $P_m=\frac{a^m}{m!}\cdot e^{-a}$ (распределение Пуассона) 398

Приложение 8. Критические точки распределения χ^2 Пирсона 400

Приложение 9. Виды распределений и их основные характеристики 401

Приложение 10. Равномерно распределенные случайные числа 404

Приложение 11. Задание для расчетно-графической работы «Моделирование показателей надежности технических систем с использованием аппарата марковских случайных процессов» 406

Приложение 12. Метод Жордана–Гаусса 419

Дополнительная литература 422

Предметный указатель 424

Учебное издание

Бережная Елена Викторовна
Бережной Владимир Иванович

МАТЕМАТИЧЕСКИЕ МЕТОДЫ МОДЕЛИРОВАНИЯ ЭКОНОМИЧЕСКИХ СИСТЕМ

Зав. редакцией *Л.А. Табакова*
Редактор *Л.В. Сергеева*
Младший редактор *Н.А. Федорова*
Художественный редактор *Ю.И. Артюхов*
Технический редактор *Т.С. Маринина*
Корректоры *Н.Н. Зубенко, Г.Д. Кузнецова*
Компьютерная верстка *Н.А. Пиминовой*
Обложка художника *О.В. Толмачевой*

ИБ № 4837

Подписано в печать 29.05.2006. Формат 60×88/16
Печать офсетная. Гарнитура «Таймс»
Усл. печ. л. 26,46. Уч.-изд. л. 24,71.
Тираж 2000 экз. Заказ 1536. «С» 107

Издательство «Финансы и статистика»,
101000, Москва, ул. Покровка, 7
Телефон (495) 625-35-02. Факс (495) 625-09-57
E-mail: mail@finstat.ru http://www.finstat.ru

ООО «Великолукская городская типография»
182100, Псковская область,
г. Великие Луки, ул. Полиграфистов, 78/12
Тел./факс: (811-53) 3-62-95
E-mail: zakaz@veltip.ru