

Е. С. КУНДЫШЕВА

МАТЕМАТИКА

УЧЕБНИК
ДЛЯ ЭКОНОМИСТОВ

$$Q = \sqrt{\frac{2\mu g}{s}}$$

$$k(y) = \left[\frac{1}{T} s \cdot \int_0^T y(t) dt + g \cdot n(T) \right]$$

$$Q - \frac{1}{2} \left(\frac{\mu T}{Q} \frac{Q}{\mu} \right) \left(T - \left[\frac{\mu T}{Q} \frac{Q}{\mu} \right] \right)$$



Издательско-торговая корпорация «Дашков и К°»

Е. С. Кундышева

МАТЕМАТИКА

Учебник для экономистов

4-е издание

*Рекомендовано
Учебно-методическим объединением
по образованию в области экономики
и экономической теории в качестве учебника
для студентов высших учебных заведений,
обучающихся по направлению подготовки
“Экономика”*

Москва
2015

УДК 51
ББК 22.1
К91

Рецензенты:

Б. С. Касаев — член-корреспондент РАЕН, проректор по научной работе ИНЭП, почетный работник ВПО РФ, доктор экономических наук, профессор;

Ю. Н. Павловский — член-корреспондент Российской академии наук (ВЦ РАН), доктор физико-математических наук, профессор (МФТИ).

Кундышева Е. С.

К91 Математика: Учебник для экономистов / Е. С. Кундышева. — 4-е изд. — М.: Издательско-торговая корпорация «Дашков и К^о», 2015. — 564 с.

ISBN 978-5-394-02261-6

В учебнике на простейшем уровне изложены необходимые экономистам основы высшей математики, на которых базируются экономико-математические методы и строятся математические модели рыночной экономики.

Основные положения учебного материала сопровождаются большим количеством примеров и задач из области корпоративного управления, макро- и микроэкономики с соответствующими упражнениями, контрольными вопросами для самостоятельной работы.

Для студентов бакалавриата, обучающихся по направлениям подготовки «Экономика», «Менеджмент», «Торговое дело».

Санитарно-эпидемиологическое заключение
№ 77.99.60.953.Д.007399.06.09 от 26.06.2009 г.

Подписано в печать 20.06.2014. Формат 60×84 1/16
Печать офсетная. Бумага офсетная № 1. Печ. л. 35,25.
Тираж 500 экз.

Издательско-торговая корпорация «Дашков и К^о»
129347, Москва, Ярославское шоссе, д. 142, к. 732.

Тел./факс: 8(495) 741-34-28,

E-mail: sales@dashkov.ru — отдел продаж;
<http://www.dashkov.ru>

Отпечатано в ГУП Академиздатцентр «Наука» РАН,
ОП «Производственно-издательский комбинат «ВИНИТИ»-«Наука»,
140014, Московская обл., г. Люберцы, Октябрьский пр-т, д. 403.
Тел./факс: 554-21-86, 554-25-97, 974-69-76

ISBN 978-5-394-02261-6

© Кундышева Е. С., 2007

*Светлой памяти
любимого учителя,
доктора физико-математических наук,
профессора Моргунова Бориса Ивановича
посвящается...*

Введение

Дорогие читатели! Ваши многочисленные отклики на первые три издания учебного пособия "Математика" укрепили веру автора в необходимость продолжения работы над книгой. Огромное всем спасибо за отзывы и пожелания. Сейчас вы держите в руках учебник, подготовленный с учетом высказанных замечаний и предложений.

В книге рассмотрены основные вопросы курса высшей математики в соответствии с Государственным образовательным стандартом высшего профессионального образования.

Учебник, помимо основных разделов математики (основ линейной алгебры, математического анализа и линейного программирования, введения в теорию вероятностей и математическую статистику), содержит задачи моделирования экономических процессов и основные аспекты теории игр. Эти разделы, собранные воедино, имеют большое практическое значение для подготовки будущих экономистов.

Учебник отражает практику преподавания автором предмета в Московском экономико-лингвистическом институте, в Институте экономики Российской экономической академии им. Г.В. Плеханова, на экономических факультетах Московского государственного университета путей сообщения (МИИТ) и Социально-технологического института МГУС, а также использует опыт работы автора со студентами Всероссийской государственной налоговой академии при Министерстве финансов Российской Федерации.

Для наиболее эффективного усвоения материала рассмотрены примеры решения задач экономического содержания из об-

ласти корпоративного управления, микро- и макроэкономики, имеются задачи, в которых предлагается моделировать различные экономические процессы.

Объединение всех перечисленных выше разделов в одной книге, выразившееся не только в приведении основных сведений из теории, но и в рассмотрении большого количества конкретных примеров и задач, контрольных заданий для самостоятельной работы, позволяет рекомендовать его не только для студентов бакалавриата, но и для заочного и дистанционного обучения, которое в последние годы получило широкое распространение.

Поэтому курс математики сопровождается контрольными работами, индивидуальными заданиями со ссылками на аналогичные уже решенные примеры и задачи из учебника. Также все желающие, в том числе преподаватели курса и их студенты, могут пользоваться специально созданным автором электронным лабораторным практикумом к четырехсеместровому курсу “Математика”, включающим в себя не только обучающий комплекс, но и многократно опробованную систему опроса знаний студентов. Электронный аппаратно-программный комплекс имеет широкие возможности: электронный учебник плюс тестовая база опроса и контроля знаний студентов со статистикой для преподавателя (шкала оценок, журнал, диаграмма сложности вопросов, досье на каждого студента и др.). В состав обучающего комплекса по математике входят Интернет-сайт, CD-диск, программа и, наконец, учебник, который вы держите в руках.

Множество людей великодушно оказывали автору содействие в ходе первоначального написания и последующей переработки этой книги. Экспертные советы, доброжелательная критика специалистов, практический опыт преподавателей и общие консультации коллег оказались бесценными.

Отзывы и предложения со стороны студентов, получать которые особенно приятно, укрепили веру в необходимость продолжения работы над учебником.

Автор благодарит всех, кто оказал помощь при подготовке данного издания, особенно Е.В. Безрукавую, а также В.Я. Якубова, Н.Д. Харламову, И.В. Мартынову, В.А. Каймина, В.П. Майбо-

роду, Б.А. Сулакова, А.А. Никулина, И.Е. Степанова, Л. А. Ковалкину, Я.П. Кундышева за то, что стимулировали ее мышление.

Большое спасибо всему доброжелательному творческому коллективу Издательско-торговой корпорации «Дашков и К°» за успешное взаимовыгодное сотрудничество.

Со всеми замечаниями и предложениями по содержанию учебника, а также по вопросам сотрудничества, приобретения электронного аппаратно-программного комплекса обращайтесь на WEB-сайт: <http://www.matecon.com>. Автор будет глубоко признательна вам за общение и за практические советы и рекомендации по усовершенствованию своих разработок.

Подпись

9 июня 2007 г.
Кундышева Е.С.

Глава 1

ВВЕДЕНИЕ В ЛИНЕЙНУЮ АЛГЕБРУ

§1. Векторные пространства и матрицы

1.1. Основные сведения о векторах и векторных пространствах

Очень часто для построения экономических моделей требуется простая, компактная форма записи экономических процессов. Этим объясняется необходимость введения понятия матрицы и основанного на нем раздела математики — *матричной алгебры*. Но прежде нужно вспомнить некоторые понятия и ввести ряд определений.

Многие утверждения и теоремы учебника приводятся без доказательств, чтобы упростить студентам-экономистам их и без того нелегкий путь овладения необходимым математическим инструментарием. Цифры в квадратных скобках означают ссылки на список рекомендуемой литературы, помещенный в конце книги.

Кроме того, все главы этой книги можно рассматривать как продолжение школьного курса математики. Поэтому мы не ставим своей целью изложение логических основ предмета и опираемся на многие понятия и теоремы курса элементарной математики. Например, определение вещественных (действительных) чисел, декартовой системы координат, отображения, точки, прямой, длины отрезка.

Понятие вектора также известно из школьного курса математики, но мы напомним основные положения, с ним связанные.

Если про две точки известно, какая из них первая, а какая — вторая, то эту пару точек мы назовем *упорядоченной*.

ОПРЕДЕЛЕНИЕ. Отрезок, концы которого упорядочены, называется *направленным отрезком*, или *вектором*. Первый из его

концов называется *началом*, второй — *концом* вектора. *Нулевым* будет вектор, у которого начало и конец совпадают.

В книге буквы, обозначающие вектора, набраны курсивным шрифтом, например вектор *a*. Нулевой вектор обозначается *o*.

ОПРЕДЕЛЕНИЕ. Расстояние между началом и концом вектора называется его *длиной*, или *абсолютной величиной*, и обозначается $|A|$.

ОПРЕДЕЛЕНИЕ. Векторы называются *коллинеарными*, если существует такая прямая, которой они параллельны. Векторы *компланарны*, если существует плоскость, которой они параллельны.

Нулевой вектор считается коллинеарным любому вектору, так как он не имеет определенного направления. Его длина равна нулю.

ОПРЕДЕЛЕНИЕ. Два вектора называются *равными*, если они коллинеарны, одинаково направлены и имеют равные длины.

Понятие множества также является одним из основных в математике.

ОПРЕДЕЛЕНИЕ. Семейство объектов, объединенных по определенному признаку, называется *множеством*. Объекты, составляющие множество, называются его *элементами*, или *точками*.

Обычно множества обозначаются большими буквами, а входящие в них элементы — малыми буквами. Элемент x из множества X соответствует записи $x \in X$ (x принадлежит X); если же элемент x не входит в множество X , то это соответствует записи $x \notin X$ (x не принадлежит X). Если все элементы множества X содержатся в другом множестве Y , то $X \subset Y$ и говорят, что X является *подмножеством* множества Y .

ОПРЕДЕЛЕНИЕ. *Вещественным векторным пространством* называется множество L , элементы которого являются векторами, удовлетворяющее следующим условиям (аксиомам).

1. Определена операция сложения векторов, результатом которой является вектор: $a, b \in L \Rightarrow a + b \in L$.
2. $a + b = b + a$ для всех $a, b \in L$ (коммутативность).
3. $a + (b + c) = (a + b) + c$ для всех $a, b, c \in L$ (ассоциативность).
4. Существует нулевой вектор o , такой, что $o + a = a + o = a$ для любого $a \in L$.
5. Для всякого вектора $a \in L$ и вещественного числа $\alpha \in R$ определено их произведение $a\alpha \in L$.
6. $(\alpha + \beta)a = \alpha a + \beta a$ для всех $\alpha, \beta \in R$ и $a \in L$.
7. $\alpha(a + b) = \alpha a + \alpha b$ для всех $\alpha \in R$ и $a, b \in L$.
8. $0a = o$ для всех $a \in L$.
9. $1a = a$ для всех $a \in L$.

Элементами (точками, векторами) вещественного векторного пространства R^n являются векторы-столбцы, состоящие из n ве-

щественных чисел $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$; операции сложения и умножения на

число определены следующим образом:

$$a + b = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}; a\alpha = \begin{pmatrix} \alpha a_1 \\ \vdots \\ \alpha a_n \end{pmatrix}.$$

Нулевой вектор имеет все координаты, равные нулю.

ОПРЕДЕЛЕНИЕ. Вектор $a = (a_1, a_2, \dots, a_n)$ и вектор $b = (b_1, b_2, \dots, b_n)$ **равны** в том случае, если совпадают их компоненты, стоящие на одинаковых местах, т.е. если $a_j = b_j$, при $j = 1, 2, \dots, n$.

ОПРЕДЕЛЕНИЕ. **Противоположным вектору** a называется вектор $-a = (-a_1, -a_2, \dots, -a_n)$. Очевидно, что $a + (-a) = 0$.

ОПРЕДЕЛЕНИЕ. **Разностью векторов** a и b называется вектор $a - b = a + (-b)$, или $a - b = (a_1 - b_1, a_2 - b_2, \dots, a_n - b_n)$.

Сложение n -мерных векторов возникло из геометрического сложения векторов на плоскости или в трехмерном пространстве и производится по правилу параллелограмма.

Умножение вектора a на действительное число k означает растяжение вектора в $|k|$ раз при $|k| > 1$ и сжатие вектора в $|k|$ раз при $|k| < 1$. При этом если $k < 0$, то направление вектора ka противоположно направлению вектора a .

ОПРЕДЕЛЕНИЕ. *Скалярным произведением* двух векторов a и b называется действительное число, равное сумме произведений соответствующих компонент этих векторов: $a \cdot b = a_1 b_1 + a_2 b_2 + \dots + a_n b_n$.

В векторной записи $(\bar{a}, \bar{b}) = \bar{a} \bar{b} = |\bar{a}| |\bar{b}| \cos \varphi$, т. е. скалярное произведение равно произведению длин этих векторов на косинус угла φ между ними.

ОПРЕДЕЛЕНИЕ. *Длиной вектора a , или его модулем*, называется действительное неотрицательное число $a = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$.

Рассмотрим линейное уравнение, содержащее n неизвестных: $a_1 x_1 + a_2 x_2 + \dots + a_n x_n = b$. Левая часть этого уравнения — линейная функция от n неизвестных:

$z = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$; она может быть представлена в виде скалярного произведения векторов $z = a \cdot x$, где $a = (a_1, a_2, \dots, a_n)$, $x = (x_1, x_2, \dots, x_n)$.

ОПРЕДЕЛЕНИЕ. Вектор b n -мерного векторного пространства называется *пропорциональным* вектору a , если существует такое число k , при котором выполняется соотношение $b = ka$.

В частности, нулевой вектор пропорционален любому вектору a , так как $0 \cdot a = 0$.

ОПРЕДЕЛЕНИЕ. Вектор b называется *линейной комбинацией* векторов $a_1, a_2, \dots, a_n$, если существуют такие числа $k_1, k_2, \dots, k_n$, при которых выполняется соотношение: $b = k_1 a_1 + k_2 a_2 + \dots + k_n a_n$.

Следовательно, каждая j -я компонента вектора b при $j = 1, 2, \dots, n$ равна сумме произведений j -х компонент векторов $a_1, a_2, \dots, a_n$ соответственно на числа $k_1, k_2, \dots, k_n$.

ОПРЕДЕЛЕНИЕ. Система векторов $a_1, a_2, \dots, a_r$, ($r \geq 0$) называется *линейно зависимой*, если хотя бы один из векторов системы является линейной комбинацией остальных, и линейно независимой — в противном случае.

Например, система векторов $b = (8, 5, 11)$, $a_1 = (1, 2, 3)$, $a_2 = (3, 2, 1)$, $a_3 = (3, 1, 2)$ линейно зависима, так как вектор b — линейная комбинация векторов a_1, a_2, a_3 , так как вектор b можно представить в виде $b = 2a_1 - a_2 + 3a_3$:

$$\begin{pmatrix} 8 \\ 5 \\ 11 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} + 3 \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}.$$

Можно несколько иначе определить линейно независимые и линейно зависимые векторы.

ОПРЕДЕЛЕНИЕ. Векторы $a_1, \dots, a_k$ называются *линейно независимыми*, если из того, что

$$\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_k a_k = 0, \alpha_i \in \mathbf{R},$$

следует, что все $\alpha_i = 0$.

ОПРЕДЕЛЕНИЕ. Векторы $a_1, \dots, a_k$ называются *линейно зависимыми*, если существует набор α_i , $i = 1, \dots, k$, где хотя бы одно α_i отлично от нуля.

Например, система двух векторов $a_1 = (1, 0)$ и $a_2 = (0, 2)$ линейно независима; система двух векторов $b_1 = (1, 2, 1)$ и $b_2 = (2, 4, 2)$ линейно зависима, так как $b_2 - 2b_1 = 0$ или $b_2 = 2b_1$.

Укажем свойства линейно зависимой системы векторов.

1°. Система, состоящая из одного вектора, линейно зависима.

2°. Система, содержащая нулевой вектор, всегда линейно зависима.

3°. Система, содержащая более одного вектора, линейно зависима тогда и только тогда, когда среди ее векторов содержится, по крайней мере, один вектор, который линейно выражается через остальные.

Геометрический смысл линейной зависимости векторов очевиден для случаев двумерных векторов на плоскости и трехмерных векторов в пространстве.

В случае двух векторов, когда один вектор выражается через другой,

$$b_2 = \lambda b_1,$$

т.е. эти векторы *коллинеарны* или, что то же самое, находятся на параллельных прямых. В пространственном случае линейной зависимости трех векторов они параллельны одной плоскости, т.е. *компланарны*; достаточно “подправить” соответствующими сомножителями длины этих векторов, чтобы один из них стал суммой двух других или выражался через них.

УТВЕРЖДЕНИЕ. В пространстве $\mathbf{R}^n$ любая система, содержащая m векторов, линейно зависима при $m > n$.

ОПРЕДЕЛЕНИЕ. *Линейным подпространством* линейного пространства $\mathbf{L}$ называется подмножество $\mathbf{K}$ векторов пространства $\mathbf{L}$, замкнутое относительно операций сложения и умножения на число, т.е. из того, что векторы $a, b \in \mathbf{K}$, следует, что $a + b$ и αa принадлежат $\mathbf{K}$.

Например, рассмотрим множество векторов из $\mathbf{R}^n$, состоящее из таких векторов, у которых последние $n - k$ координат равны 0. Нетрудно проверить, что это множество является линейным подпространством пространства $\mathbf{R}^n$, совпадающим с $\mathbf{R}^n$.

ОПРЕДЕЛЕНИЕ. Множество всех *линейных комбинаций* векторов $a_1, \dots, a_k \in \mathbf{L}$ и $\alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_k a_k = 0$, $\alpha_i \in \mathbf{R}$, называется *пространством, порожденным векторами* $a_1, \dots, a_k$. (Проверьте, что оно является линейным подпространством векторного пространства $\mathbf{L}$.)

Если линейное подпространство $\mathbf{K}$ векторного пространства $\mathbf{L}$ не совпадает с ним, то его часто называют *гиперплоскостью*.

ОПРЕДЕЛЕНИЕ. Набор векторов $a_1, \dots, a_n \in \mathbf{L}$ называется *базисом* пространства $\mathbf{L}$, если выполняются два условия:

- 1) векторы $a_1, \dots, a_n \in \mathbf{L}$ линейно независимы;
- 2) пространство, порожденное векторами $a_1, \dots, a_n \in \mathbf{L}$, совпадает с $\mathbf{L}$, или всякий вектор пространства $\mathbf{L}$ линейно выражается через эти векторы.

Например, набор векторов e_i , $i = 1, \dots, n$, у которых все координаты, кроме 1-й, равны нулю, а i -я координата равна 1, является базисом в пространстве $\mathbf{R}^n$. Сами векторы e_i еще называют *базисными*.

Для записи таких векторов удобно использовать *символ Кронекера*:

$$\delta_{ij} = \begin{cases} 0, i \neq j, \\ 1, i = j. \end{cases}$$

Тогда базисные векторы можно записать как
$$e_i = \begin{pmatrix} \delta_{i_1} \\ \vdots \\ \delta_{i_n} \end{pmatrix}.$$

УТВЕРЖДЕНИЕ. *Все базисы векторного пространства $\mathbf{L}$ содержат одно и то же число векторов, которое называется размерностью $\dim(\mathbf{L})$ векторного пространства $\mathbf{L}$.*

Например, размерность $\mathbf{R}^n$ равна: $\dim(\mathbf{R}^n) = n$.

Рассмотрим в трехмерном векторном пространстве прямоугольную систему координат $OXYZ$ и три единичных вектор: e_1, e_2, e_3 . Рассмотрим векторы на осях координат так, чтобы их направление совпадало с положительным направлением осей; тогда их можно записать в следующем виде:

$$\begin{cases} e_1 = (1; 0; 0); \\ e_2 = (0; 1; 0); \\ e_3 = (0; 0; 1). \end{cases}$$

Эти векторы образуют систему единичных векторов трехмерного векторного пространства $\mathbf{L}$.

Система единичных векторов n -мерного векторного пространства:

$$\begin{cases} e_1 = (1, 0, \dots, 0); \\ e_2 = (0, 1, \dots, 0); \\ \dots \\ e_n = (0, 0, \dots, 1). \end{cases}$$

УТВЕРЖДЕНИЕ. Система единичных векторов *линейно независима* и образует базис.

УТВЕРЖДЕНИЕ. Любой вектор a линейного пространства можно единственным способом разложить по базису, т.е. представить в виде линейной комбинации базисных векторов: $a = \alpha_1 a_1 + \alpha_2 a_2 + \dots + \alpha_n a_n$, $\alpha_i \in \mathbf{R}$.

Разложим вектор $a = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$ по единичному базису:

$$a = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} = 3e_1 + 2e_2 + 5e_3 = 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 5 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \text{ Точно так же мож-}$$

но разложить любой вектор.

Приведем пример из экономики. Под *товаром* понимается некоторое благо (или услуга), поступившее в продажу в определенное время и в определенном месте.

Будем считать, что имеется n различных товаров, количество i -го товара обозначается x_i , тогда некоторый набор товаров

обозначается $X = (x_1, x_2, \dots, x_n)$. Как уже выяснили, упорядоченный набор n чисел называется n -мерным вектором, так что X есть n -мерный вектор. Будем рассматривать, как правило, только неотрицательные количества товаров, так что $x_i \geq 0$ для любого $i = 1, \dots, n$, или $X \geq 0$. Множество всех наборов товаров называется *пространством товаров* $\mathbf{G}$. Набор товаров можно трактовать как корзину, в которой лежат эти товары в соответствующем количестве. Аналогично интерпретируются и операции с наборами товаров.

ОПРЕДЕЛЕНИЕ. *Линейным оператором* называется отображение векторного пространства $\mathbf{L}$ в векторное пространство $\mathbf{M}$, $A : \mathbf{L} \rightarrow \mathbf{M}$, обладающее следующими свойствами:

1. $A(a + b) = A(a) + A(b)$ для всех $a, b \in \mathbf{L}$,
2. $A(\alpha a) = \alpha A(a)$ для всех $\alpha \in \mathbf{R}$, $a \in \mathbf{L}$.

Линейный оператор полностью определяется своими значениями на базисных векторах. В самом деле, любой вектор x , принадлежащий n -мерному векторному пространству $\mathbf{L}$, можно разложить по базисным векторам e_i : $x = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$, $\alpha_i \in \mathbf{R}$. Тогда, используя определение линейного оператора, получаем: $A(x) = \alpha_1 A(e_1) + \alpha_2 A(e_2) + \dots + \alpha_n A(e_n)$.

Каждый из n векторов $A(e_j)$, $j = 1, \dots, n$ также можно разложить

по базису l_i , $i = 1, \dots, m$ (в m -мерном пространстве $\mathbf{M}$): $A(e_j) = \sum_{i=1}^m a_{ij} l_i$.

Таким образом, линейному оператору A и выбранным базисам $\{e_j\}$, $\{l_i\}$ в пространствах $\mathbf{L}$, $\mathbf{M}$ соответствует таблица $m \times n$ действительных чисел.

1.2. Основные сведения о матрицах

ОПРЕДЕЛЕНИЕ. *Прямоугольной матрицей* типа $m \times n$ называется совокупность $m \times n$ действительных чисел, расположенных в виде прямоугольной таблицы, где m — количество строк и n — количество столбцов. Совокупность целых чисел $m \times n$ называют *размерами* матрицы. Числа, составляющие матрицу, называются ее *элементами*.

Матрицы обозначаются прописными (заглавными) буквами латинского алфавита: $A, B, C, \dots$. Для обозначения элементов матрицы используются строчные буквы с двойной индексацией: a_{ij} , где i — номер строки, j — номер столбца.

Например, матрица

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}, \quad (1.1)$$

или в сокращенной записи $A = (a_{ij})$; $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

Например, $A_{2 \times 3} = \begin{pmatrix} 6 & 0 & -3 \\ 12 & -5 & 18 \end{pmatrix}$.

Матрица A имеет размеры 2×3 , у нее две строки и три столбца. Таким образом, элемент $a_{22} = -5$ стоит на пересечении 2-й строки и 2-го столбца.

Наряду с круглыми скобками используются и другие обозначения матрицы: $[]$, $\| \|$.

ОПРЕДЕЛЕНИЕ. Две матрицы A и B одинаковых размеров называются **равными** $A_{m \times n} = B_{m \times n}$, если все элементы с одинаковыми индексами обеих матриц совпадают, т.е. $a_{ij} = b_{ij}$ для любых $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

С помощью матриц удобно записывать некоторые экономические зависимости. Например, таблица распределения средств в банковской системе из трех банков A, B, B

Банк	Распределение средств, усл. ед.	
	Обязательства	Обязательные резервы
A	100 000	20 000
B	80 000	16 000
B	64 000	12 800

может быть записана в компактной форме в виде матрицы распределения средств по банкам:

$$A_{3 \times 2} = \begin{pmatrix} 100000 & 20000 \\ 80000 & 16000 \\ 64000 & 12800 \end{pmatrix}.$$

В этой записи, например, матричный элемент $a_{31} = 64\,000$ показывает, каковы вклады в банк B , а элемент $a_{22} = 16\,000$ — сколько обязательных резервов должен иметь банк B .

ОПРЕДЕЛЕНИЕ. Матрица размеров $1 \times n$, состоящая из одной строки, называется **матрицей (вектором) — строкой**, матрица размеров $m \times 1$, состоящая из одного столбца — **матрицей (вектором) — столбцом**, а матрица размеров 1×1 — **скалярной матрицей**.

Например, $A_{1 \times n} = (a_{11}, a_{12}, \dots, a_{1n})$ — матрица-строка;

$$B_{m \times 1} = \begin{pmatrix} b_{11} \\ b_{21} \\ \dots \\ b_{m1} \end{pmatrix} \text{ — матрица-столбец.}$$

Далее, если не оговорено противное, мы будем для удобства рассматривать вектор как вектор-столбец.

ОПРЕДЕЛЕНИЕ. Если число строк в матрице равно числу столбцов, то матрица называется **квадратной**, а число строк — ее **порядком**.

$$\text{Например, } A_{3 \times 3} = \begin{pmatrix} 0 & 1 & -3 \\ 6 & 9 & -1 \\ -1 & 3 & 2 \end{pmatrix} \text{ — квадратная матрица третьего}$$

порядка.

ОПРЕДЕЛЕНИЕ. Если матрица квадратная, то совокупность тех ее элементов a_{ii} , у которых номер строки равен номеру столбца, называется *главной диагональю*, или просто *диагональю* матрицы.

Таким образом, главную диагональ квадратной матрицы образуют элементы $a_{11}, a_{22}, \dots, a_{nn}$.

ОПРЕДЕЛЕНИЕ. Квадратная матрица называется *диагональной*, если все ее недиагональные элементы равны нулю.

Например, $A = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ — диагональная матрица третьего

порядка.

ОПРЕДЕЛЕНИЕ. Если у диагональной матрицы n -го порядка все диагональные элементы равны единице, то матрица называется *единичной матрицей n -го порядка* и обозначается буквой E .

Например, $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ — единичная матрица третьего

порядка.

ОПРЕДЕЛЕНИЕ. Если все элементы матрицы любых размеров равны нулю, то она называется *нулевой*, или *нуль-матрицей*:

$$O_{m \times n} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

1.3. Операции над матрицами

1. Умножение матрицы на число.

ОПРЕДЕЛЕНИЕ. *Произведением матрицы A на число λ называется матрица $B = \lambda A$, элементы которой $b_{ij} = \lambda a_{ij}$ для $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.*

Например, если $A = \begin{pmatrix} -2 & -4 \\ 1 & 0 \end{pmatrix}$, то $B = 5A = \begin{pmatrix} -10 & -20 \\ 5 & 0 \end{pmatrix}$.

СЛЕДСТВИЕ. *Общий множитель всех элементов матрицы можно выносить за знак матрицы.*

Например, $\begin{pmatrix} 6 & -2 & 4 \\ -4 & 32 & 0 \end{pmatrix} = 2 \times \begin{pmatrix} 3 & -1 & 2 \\ -2 & 16 & 0 \end{pmatrix}$.

В частности, произведение матрицы A на число 0 есть нулевая матрица, т.е. $0 \cdot A = O$.

2. Сложение матриц.

ОПРЕДЕЛЕНИЕ. *Суммой двух матриц A и B одинаковых размеров $m \times n$ называется матрица $C = A + B$ того же типа, элементы которой равны сумме соответствующих элементов матриц — слагаемых, т.е. $c_{ij} = a_{ij} + b_{ij}$ для $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.*

Например,

$$A = \begin{pmatrix} 3 & -1 & 0 \\ 1 & -1 & 1 \end{pmatrix}, B = \begin{pmatrix} -3 & 3 & 1 \\ 1 & 0 & -2 \end{pmatrix}, C = A + B = \begin{pmatrix} 0 & 2 & 1 \\ 2 & -1 & -1 \end{pmatrix}.$$

В частном случае $A + O = A$.

3. Вычитание матриц.

ОПРЕДЕЛЕНИЕ. *Разность двух матриц одинаковых размеров определяется через операции сложения и умножения матриц на число: $A - B = A + (-1) \cdot B$.*

4. Транспонирование матрицы.

Пусть дана матрица $A_{m \times n}$ (см. формулу (1.1)), поменяем в ней местами строки со столбцами: на место 1-й строки поставим 1-й

ОПРЕДЕЛЕНИЕ. Произведением матриц $A_{m \times k} \cdot B_{m \times k}$ называется такая матрица $C_{m \times n}$, каждый элемент которой c_{ij} равен сумме произведений элементов i -й строки матрицы A на соответствующие элементы j -го столбца матрицы B .

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ik}b_{kj} = \sum_{s=1}^k a_{is}b_{sj}, i = 1, 2, \dots, m;$$

$$j = 1, 2, \dots, n.$$

▷ ПРИМЕР 1.1. Вычислить произведение $A \cdot B$, где

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \end{pmatrix}; B = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Решение. Найдем размер матрицы-произведения (если умножение матриц возможно): $A_{2 \times 3} \cdot B_{3 \times 3} = C_{2 \times 3}$.

Вычислим элементы матрицы-произведения C , умножая элементы каждой строки матрицы A на соответствующие элементы столбцов матрицы B следующим образом:

$$C = \begin{pmatrix} 1 \cdot 1 + 2(-1) + 2 \cdot 1 & 1 \cdot 1 + 1(-1) + 2(-1) & 1(-1) + 0(-1) + 2 \cdot 1 \\ 0 \cdot 1 + 1 \cdot 2 + 1 \cdot 1 & 0 \cdot 1 + 1 \cdot 1 + 1 \cdot (-1) & 0(-1) + 1 \cdot 0 + 1 \cdot 1 \end{pmatrix}.$$

$$\text{Получаем } C = \begin{pmatrix} 1 & -2 & 1 \\ 3 & 0 & 1 \end{pmatrix}. \blacktriangleright$$

Многие свойства, присущие операциям над числами, справедливы и для операций над матрицами (что следует из определений этих операций).

Свойства операций над матрицами:

- 1) $A + B = B + A$; 2) $(A + B) + C = A + (B + C)$;
- 3) $\lambda(A + B) = \lambda A + \lambda B$; 4) $A \cdot (B + C) = A \cdot B + A \cdot C$;
- 5) $(A + B) \cdot C = A \cdot C + B \cdot C$; 6) $\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$.
- 7) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

Однако имеются и специфические свойства матриц. Так, операция умножения матриц имеет некоторые отличия от умножения чисел:

I. Если произведение матриц $A \cdot B$ существует, то после перестановки сомножителей местами произведения матриц $B \cdot A$ может и не существовать. Действительно, в примере 1.1 получили произведение матриц $A_{2 \times 3} \cdot B_{3 \times 3}$, а произведения $B_{3 \times 3} \cdot A_{2 \times 3}$ не существует, так как число столбцов первой матрицы не совпадает с числом строк второй матрицы.

II. Если даже произведения $A \cdot B$ и $B \cdot A$ существуют, то они могут быть матрицами разных размеров.

▷ ПРИМЕР 1.2. Найти произведения матриц $A \cdot B$ и $B \cdot A$:

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 1 & -1 \end{pmatrix}.$$

Р е ш е н и е. $A_{2 \times 3} \cdot B_{3 \times 2} = C_{2 \times 2} = \begin{pmatrix} 1 & -2 \\ 3 & 0 \end{pmatrix};$

$$B_{3 \times 2} \cdot A_{2 \times 3} = D_{3 \times 3} = \begin{pmatrix} 1 & 0 & 3 \\ 2 & -1 & 5 \\ 1 & -2 & 1 \end{pmatrix}, \text{ т.е. } A \cdot B \neq B \cdot A. \blacktriangleright$$

III. Даже в случае, когда оба произведения $A \cdot B$ и $B \cdot A$ существуют и оба — матрицы одинаковых размеров (это возможно только при умножении квадратных матриц A и B одного порядка), коммутативный (переместительный) закон умножения, вообще говоря, не выполняется, т.е. $A \cdot B \neq B \cdot A$.

▷ ПРИМЕР 1.3. Найти произведения матриц $A \cdot B$ и $B \cdot A$, где

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Р е ш е н и е. $A \cdot B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; B \cdot A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \text{ т.е. } A \cdot B \neq B \cdot A. \blacktriangleright$

В частном случае коммутативным законом обладает произведение любой квадратной матрицы A n -го порядка на единичную матрицу E того же порядка, причем это произведение равно A :

$$A \cdot E = E \cdot A = A.$$

Это легко проверить:

$$A_{n \times n} \cdot E_{n \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = A.$$

$$E_{n \times n} \cdot A_{n \times n} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = A.$$

Таким образом, единичная матрица играет при умножении матриц ту же роль, что и число 1 при умножении чисел.

ОПРЕДЕЛЕНИЕ. Если $A \cdot B = B \cdot A$, то матрицы A и B называются *перестановочными* между собой, или *коммутирующими* между собой.

Так, единичная матрица перестановочна с любой квадратной матрицей того же порядка.

IV. Произведение двух ненулевых матриц может равняться нулевой матрице, т.е. из того, что $A \cdot B = O$, не следует, что $A = O$ или $B = O$.

Например,

$$A = \begin{pmatrix} 2 & -3 \\ 4 & -6 \end{pmatrix} \neq O; B = \begin{pmatrix} 9 & -6 \\ 6 & -4 \end{pmatrix} \neq O, \text{ но } A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = O.$$

V. Важным частным случаем произведения матриц является произведение квадратной $n \times n$ матрицы A на вектор b .

Например,

$$A \cdot b = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 2 & 4 & 6 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 1 \cdot b_1 + 2 \cdot b_2 + 3 \cdot b_3 \\ 1 \cdot b_1 + 0 \cdot b_2 + 1 \cdot b_3 \\ 2 \cdot b_1 + 4 \cdot b_2 + 6 \cdot b_3 \end{pmatrix} = b_1 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + b_2 \begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix} + b_3 \begin{pmatrix} 3 \\ 1 \\ 6 \end{pmatrix}.$$

Как видно из примера, вектор Ab является *линейной комбинацией* столбцов матрицы A с коэффициентами b_i .

Аналогично при умножении матрицы A на вектор-строку (слева) $b'A$ мы получаем вектор-строку, являющийся линейной комбинацией *строк* матрицы A с коэффициентами b_i .

ОПРЕДЕЛЕНИЕ. *Скалярным произведением* двух векторов a, b размерности n называется число, равное $a'b = b'a = a_1b_1 + a_2b_2 + \dots + a_nb_n$.

Например,

$$a'b = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = 1 \cdot 0 + 1(-2) + 3 \cdot 2 = 4.$$

З а м е ч а н и е. Элемент с координатами i, j в произведении матриц $A \cdot B$ равен скалярному произведению i -го вектора-строки матрицы A на j -й вектор-столбец матрицы B .

В начале параграфа мы начали разговор о пространстве товаров. Предполагается, что каждый товар имеет цену. Все цены строго положительны. Пусть цена единицы i -го товара есть p_i , тогда $P = (p_1, p_2, \dots, p_n)$ есть вектор-строка цен.

Для набора товаров X и вектора цен P их скалярное произведение $PX = p_1x_1 + p_2x_2 + \dots + p_nx_n$ есть число, называемое *ценой набора* X . Цена набора X будет обозначаться $C(X)$.

К разговору о пространстве товаров и цене корзины товаров мы снова вернемся в главе 4 (см. § 1).

6. Возведение в степень.

С введением операции умножения матриц появилась возможность рассматривать возведение матрицы в степень.

ОПРЕДЕЛЕНИЕ. *Целой положительной степенью* A^m ($m > 1$) квадратной матрицы A называется произведение m матриц, равных A , т.е. $A^m = \underbrace{A \cdot A \cdot \dots \cdot A}_{m \text{ раз}}$.

Заметим, что операция возведения в степень определяется только для квадратных матриц.

Полагают, что $A^0 = E$, $A^1 = A$. Нетрудно показать, что $A^m \cdot A^k = A^{m+k}$, $(A^m)^k = A^{mk}$.

▷ ПРИМЕР 1.4. Найти A^2 , где $A = \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix}$.

Р е ш е н и е. $A^2 = \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 7 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 51 & 16 \\ 8 & 3 \end{pmatrix}$. ►

Заметим, что из равенства $A^m = O$ еще не следует, что матрица $A = O$.

1.4. Собственные значения и собственные векторы матрицы

Рассмотрим квадратные матрицы порядка n . При умножении матрицы порядка n на n -мерный вектор в произведении получается n -мерный вектор:

$$Ax = b.$$

Однако для любой матрицы существует набор особых векторов таких, что произведение матрицы на вектор из такого набора равносильно умножению этого вектора на определенное число.

ОПРЕДЕЛЕНИЕ. Число λ называется *собственным значением матрицы* A порядка n , если существует такой ненулевой вектор $x \in R^n$, что выполняется равенство:

$$Ax = \lambda x. \quad (1.2)$$

При этом вектор x называется *собственным вектором матрицы* A .

Уравнение (1.2) представлено в матричной форме. Группируя все слагаемые этого уравнения в левой части, его можно переписать в более удобном виде:

$$(A - \lambda E)x = o,$$

где E и o — соответственно единичная матрица и нулевой вектор.

Если $(A - \lambda E)x = o$, и a_{ij} — элементы матрицы A , то **характеристическая** матрица $A - \lambda E$, согласно определениям умножения матрицы на число и суммы матриц, имеет вид:

$$A - \lambda E = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

В дальнейшем мы вернемся к вопросу отыскания собственных значений и собственных векторов матриц.

Завершая первоначальное знакомство с матрицами, нельзя не сказать о той роли, которую играет алгебра матриц. Американский математик Ричард Беллман называл теорию матриц “арифметикой высшей математики”. Это сравнительно “молодой” раздел математики. Упоминание о матрицах впервые встречается в середине XIX века в работах ирландского астронома и математика У. Гамильтона (1805—1865) и у английских математиков Дж. Силвера (1814—1897) и А. Кэли (1821—1895). Основы теории матриц были заложены во второй половине XIX века немецкими математиками К. Вейерштрассом (1815—1897) и Фробениусом (1849—1917). Теория матриц продолжает развиваться до сих пор. Этому способствуют многочисленные и разнообразные приложения матриц.

Рассмотренные выше операции над матрицами позволяют упростить решения некоторых экономических задач.

§2. Определители квадратных матриц

Очень многие задачи макро- и микроэкономики сводятся к решению систем линейных алгебраических уравнений.

Уже в школьном курсе приходится встречаться с системой двух или трех уравнений 1-й степени.

Например,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3, \end{cases}$$

где числа $a_{11}, a_{12}, \dots, a_{33}$ и b_1, b_2, b_3 заданы, а x_1, x_2, x_3 — неизвестные, которые подлежат определению.

Мы в дальнейшем будем подробно рассматривать, как следует находить эти неизвестные x_1, x_2, x_3 , т.е. как решать систему уравнений (см. §5 главы 1). Сейчас же заметим, что такая система определяется заданием матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \text{ заданием вектора-столбца } B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},$$

неизвестные x_1, x_2, x_3 можно считать неизвестными координатами

вектора-столбца $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$

Введя такие обозначения, легко проверить, что наша система коротко запишется в матрично-векторном виде: $AX = B$.

В реальной жизни приходится иметь дело с системами уравнений, где порядок матрицы A определяется несколькими десятками. Понятно, насколько удобно применять компактную матричную запись вместо громоздкой системы. Мы увидим, что и при решении таких систем уравнений используется матричный аппарат.

Для этого, прежде чем перейти к решению систем линейных алгебраических уравнений, необходимо ввести несколько поня-

тий, в том числе — *опредетителя квадратной матрицы*¹ A — числа, которое ставится ей в соответствие и является важной характеристикой матрицы порядка n .

2.1. Основные определения

Определитель (детерминант) матрицы A обозначается $|A|$, $\det(A)$ или Δ .

ОПРЕДЕЛЕНИЕ. *Определителем матрицы $A = (a_{ij})$ первого порядка, или определителем первого порядка, называется ее единственный элемент a_{11} :*

$$\Delta_1 = |A| = a_{11}.$$

Например, пусть $A = (-3)$, тогда $\Delta_1 = |A| = -3$.

ОПРЕДЕЛЕНИЕ. *Определителем матрицы $A = (a_{ij})$; $i = 1, 2$; $j = 1, 2$ второго порядка, или определителем второго порядка, называется число, вычисляемое по формуле*

$$\Delta_2 = |A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (1.3)$$

Произведения $a_{11}a_{22}$ и $a_{12}a_{21}$ называются членами определителя второго порядка. Например, пусть $A = \begin{pmatrix} -1 & 2 \\ 5 & 1 \end{pmatrix}$, тогда

$$\Delta_2 = |A| = \begin{vmatrix} -1 & 2 \\ 5 & 1 \end{vmatrix} = 1(-1) - 2 \cdot 5 = -11.$$

Рассмотрим квадратную матрицу третьего порядка:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

¹ Везде далее, где будет идти речь об определителе, квадратную матрицу будем называть просто матрицей.

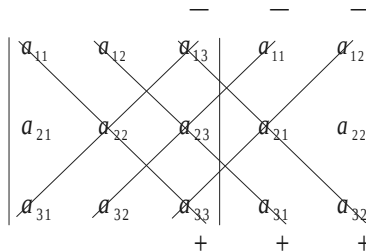
ОПРЕДЕЛЕНИЕ. *Определителем матрицы $A = (a_{ij})$; $i = 1, 2, 3$; $j = 1, 2, 3$ третьего порядка, или определителем третьего порядка, называется число, вычисляемое по формуле*

$$\Delta_3 = |A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} - a_{31}a_{22}a_{13} - a_{12}a_{21}a_{33} - a_{32}a_{23}a_{11}. \quad (1.4)$$

Это число состоит из шести слагаемых или шести членов определителя. Каждое слагаемое содержит ровно по одному элементу из каждой строки и каждого столбца матрицы. Если рассмотреть следующую схему, называемую *правилом треугольников* или *правилом Сарруса*, легко запомнить знаки, с которыми члены определителя входят в формулу (1.4):



Приведем еще одно правило, которое здесь назовем *правилом "полосочек"*. К определителю, состоящему из девяти элементов, справа припишем первые два столбца:



Параллельно главной диагонали подисуем две линии, а из элементов, стоящих вдоль полученных полос, составим произведения $a_{11}a_{22}a_{33}$; $a_{12}a_{23}a_{31}$; $a_{13}a_{21}a_{32}$. Теперь сложим их в сумму $S_1 = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$.

Аналогично предыдущему параллельно побочной диагонали подисуем две линии; из элементов, стоящих вдоль полос, составим произведения и сложим их в сумму $S_2 = a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12}$. Теперь вычтем из суммы S_1 сумму S_2 . Получим иско-
мый определитель $\Delta_3 = S_1 - S_2$, соответствующий формуле (1.4).

З а м е ч а н и е. Важно помнить, что по правилу Сарруса и по правилу “полосочек” можно вычислять лишь определители *тре-
тьего* порядка. Определители более высоких порядков считаются другими способами, о которых речь пойдет далее.

▷ ПРИМЕР 1.5. Вычислить определитель третьего порядка

$$A = \begin{vmatrix} 1 & -1 & 3 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{vmatrix}.$$

Р е ш е н и е. $\Delta = 1 \cdot 1 \cdot 2 + (-1) \cdot 1 \cdot (-1) + 0 \cdot 0 \cdot 3 - 3 \cdot 1 \cdot (-1) - 1 \times$
 $\times 0 \cdot 1 - 0 \cdot (-1) \cdot 2 = 6.$ ►

Пусть дана квадратная матрица A n -го порядка.

ОПРЕДЕЛЕНИЕ. *Минором* M_{ij} элемента a_{ij} матрицы n -го по-
рядка называется определитель матрицы $(n - 1)$ -го порядка, полу-
ченной из матрицы A вычеркиванием i -й строки и j -го столбца.

Например, минором элемента a_{12} матрицы A третьего порядка
будет:

$$M_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = a_{21}a_{33} - a_{31}a_{23}.$$

Каждая матрица n -го порядка имеет n^2 миноров $(n - 1)$ -го по-
рядка.

ОПРЕДЕЛЕНИЕ. *Алгебраическим дополнением* A_{ij} элемента a_{ij}
матрицы n -го порядка называется его минор, взятый со знаком
 $(-1)^{i+j}$:

$$A_{ij} = (-1)^{i+j} M_{ij}, \quad (1.5)$$

т.е. алгебраическое дополнение совпадает с минором, когда сумма номеров строки и столбца $(i + j)$ — четное число, и отличается от минора знаком, когда $(i + j)$ — нечетное число.

Например, $A_{23} = (-1)^{2+3} M_{23} = -M_{23}$; $A_{31} = (-1)^{3+1} M_{31} = M_{31}$.

▷ ПРИМЕР 1.6. Найти алгебраические дополнения всех элементов матрицы

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

Р е ш е н и е. По формуле (1.5) получаем

$$\begin{aligned} A_{11} &= (-1)^{1+1} \cdot \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1; \quad A_{12} = (-1)^{1+2} \cdot \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = -3; \quad A_{13} = (-1)^{1+3} \cdot \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} = -1; \\ A_{21} &= (-1)^{2+1} \cdot \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} = 1; \quad A_{22} = (-1)^{2+2} \cdot \begin{vmatrix} 0 & 0 \\ 1 & 1 \end{vmatrix} = 0; \quad A_{23} = (-1)^{2+3} \cdot \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = -1; \\ A_{31} &= (-1)^{3+1} \cdot \begin{vmatrix} -1 & 0 \\ 1 & -1 \end{vmatrix} = 1; \quad A_{32} = (-1)^{3+2} \cdot \begin{vmatrix} 0 & 0 \\ 2 & -1 \end{vmatrix} = 0; \quad A_{33} = (-1)^{3+3} \cdot \begin{vmatrix} 0 & -1 \\ 2 & 1 \end{vmatrix} = 2. \quad \blacktriangleright \end{aligned}$$

Большое значение для вычисления определителей имеет следующее определение.

ОПРЕДЕЛЕНИЕ¹. *Определитель квадратной матрицы порядка n* есть сумма произведений всех элементов любой строки (столбца) на их алгебраические дополнения:

$$\Delta = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} = \sum_{s=1}^n a_{is}A_{is} \quad (1.6)$$

(разложение по элементам i -й строки; $i = 1, 2, \dots, n$);

$$\Delta = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj} = \sum_{s=1}^n a_{sj}A_{sj} \quad (1.7)$$

(разложение по элементам j -го столбца; $j = 1, 2, \dots, n$).

¹ Точнее, данное определение является частным случаем теоремы Лапласа.

Убедиться в справедливости формул (1.6), (1.7) можно на примере определителя матрицы третьего порядка, разложив его вначале по элементам первой строки:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(-1)^{1+1} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)^{1+2} \cdot \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} +$$

$$+ a_{13}(-1)^{1+3} \cdot \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

После преобразований (предлагаем читателю выполнить их самостоятельно) нетрудно убедиться в том, что полученное выражение совпадает с формулой (1.4). Аналогичный результат получаем разложением определителя матрицы по любой строке или столбцу.

ОПРЕДЕЛЕНИЕ. Квадратная матрица называется *нижнетреугольной (верхнетреугольной)*, если все ее элементы, расположенные ниже (или выше) главной диагонали, равны нулю.

▷ **ПРИМЕР 1.7.** Вычислить определитель треугольной матрицы:

$$\begin{pmatrix} 2 & 1 & 4 & 0 \\ 0 & -1 & 5 & 9 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Решение. Раскладывая по первому столбцу, получаем

$$\begin{vmatrix} 2 & 1 & 4 & 0 \\ 0 & -1 & 5 & 9 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2 \cdot (-1)^{1+1} \cdot \begin{vmatrix} -1 & 5 & 9 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{vmatrix} + 0 + 0 + 0 =$$

$$= 2 \cdot (-1) \cdot (-1)^{1+1} \cdot \begin{vmatrix} -2 & 1 \\ 0 & 1 \end{vmatrix} + 0 + 0 = 2 \cdot (-1) \cdot (-2) = 4. \blacktriangleright$$

На частном примере мы убедились в том, что *определитель верхнетреугольной (и, очевидно, нижнетреугольной) матрицы равен произведению элементов, стоящих на главной диагонали.*

Это определение позволяет свести вычисление определителей n -го порядка к вычислению более простых определителей $(n - 1)$ -го порядка.

2.2. Свойства определителей

1. При умножении всех элементов некоторой строки матрицы на число определитель исходной матрицы умножается на это число.

Пусть определитель исходной матрицы равен Δ . Для определенности первую строку матрицы умножим на λ , получим новый определитель Δ' , который разложим по элементам первой строки:

$$\Delta' = \begin{vmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \lambda a_{11} A_{11} + \lambda a_{12} A_{12} + \dots + \lambda a_{1n} A_{1n} =$$

$$= \lambda (a_{11} A_{11} + a_{12} A_{12} + \dots + a_{1n} A_{1n}) = \lambda \Delta.$$

З а м е ч а н и е. За знак определителя можно выносить общий множитель любой отдельной строки или отдельного столбца в отличие от матрицы, за знак которой можно выносить общий множитель лишь всех элементов.

Например, для *определителя* верно:

$$\begin{vmatrix} 2 & 4 & -2 \\ 0 & 2 & -8 \\ 0 & 0 & -4 \end{vmatrix} = 2 \cdot \begin{vmatrix} 1 & 2 & -1 \\ 0 & 2 & -8 \\ 0 & 0 & -4 \end{vmatrix} = 2 \cdot 1 \cdot 2 \cdot (-4) =$$

$$= -16, \text{ но для матрицы верно: } \begin{pmatrix} 2 & 4 & -2 \\ 0 & 2 & -8 \\ 0 & 0 & -4 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -4 \\ 0 & 0 & -2 \end{pmatrix}.$$

2. *Определитель матрицы с нулевой строкой (столбцом) равен 0.*

3. *При транспонировании матрицы ее определитель не изменяется:*

$$|A'| = |A|.$$

4. *При перестановке двух строк (столбцов) матрицы ее определитель меняет знак на противоположный.*

5. *Если квадратная матрица содержит две одинаковые строки (столбца), то ее определитель равен 0.*

Действительно, переставим эти строки (столбцы). С одной стороны, определитель не изменится, но, с другой стороны, по свойству 4 поменяет знак, т.е. $\Delta = -\Delta$, откуда $\Delta = 0$.

6. *Если элементы двух строк (столбцов) матрицы пропорциональны, то ее определитель равен 0.*

Пусть для определенности пропорциональны первая и вторая строки. Тогда, вынося коэффициент пропорциональности λ , получаем по свойству 1: $\Delta' = \lambda\Delta$, где определитель Δ имеет две одинаковые строки и по свойству 5 равен 0.

7. *Сумма произведений элементов какой-либо строки (столбца) матрицы на алгебраические дополнения элементов другой строки (столбца) этой матрицы равна 0, т.е.*

$$\sum_{s=1}^n a_{is} A_{js} = 0 \left(\sum_{s=1}^n a_{si} A_{sj} = 0 \right) \text{ при } i \neq j. \quad (1.8)$$

З а м е ч а н и е. Объединяя определение определителя любого порядка и свойство 7, получаем:

$$\sum_{s=1}^n a_{is} A_{js} = \begin{cases} |A| & \text{при } i = j, \\ 0 & \text{при } i \neq j. \end{cases} \quad (1.9)$$

8. *Определитель матрицы не изменится, если к элементам какой-либо строки (столбца) матрицы прибавить элементы другой строки (столбца), предварительно умноженные на одно и то же число.*

Пусть для определенности к элементам i -й строки матрицы прибавили элементы j -й строки, умноженные на число λ ($i \neq j$). Тогда i -я строка матрицы примет вид $[(a_{i1} + \lambda a_{j1})(a_{i2} + \lambda a_{j2}) \dots (a_{in} + \lambda a_{jn})]$. Определитель полученной матрицы вычислим разложением по элементам i -й строки:

$$\Delta' = (a_{i1} + \lambda a_{j1})A_{i1} + (a_{i2} + \lambda a_{j2})A_{i2} + \dots + (a_{in} + \lambda a_{jn})A_{in},$$

где A_{is} — алгебраические дополнения элементов i -й строки исходной матрицы ($s = 1, 2, \dots, n; i \neq j$). Раскроем скобки и получим после преобразования:

$$\Delta' = \sum_{s=1}^n a_{is} A_{is} + \lambda \sum_{s=1}^n a_{js} A_{is} \quad (i \neq j).$$

Используя формулу (1.9), получаем, что первая сумма равна определителю исходной матрицы, а вторая — нулю, т.е. $\Delta' = \Delta$.

9. *Определитель не меняется, если к одной строке (столбцу) матрицы добавить линейную комбинацию (см. §4 главы 1) других строк (столбцов).*

Это свойство следует из свойства 8.

10. *Определитель равен 0 тогда и только тогда, когда строки (столбцы) линейно зависимы (см. §4 главы 1).*

11. *Определитель произведения двух квадратных матриц равен произведению их определителей: $|C| = |A| \cdot |B|$, где $C = A \cdot B$; A и B — матрицы n -го порядка.*

З а м е ч а н и е. Из свойства 11 следует, что даже если $AB \neq BA$, то $|AB| = |BA|$.

Все эти свойства определителей помогают значительно упрощать их вычисления. Если определитель имеет высокий порядок, то с помощью свойств 1—8 можно так преобразовать исходную матрицу, что ее определитель вычислится мгновенно. Для этого можно предложить два метода.

I. Для упрощения вычисления определителей целесообразно изменить определитель таким образом, чтобы его преобразованная матрица имела столбец (или строку), содержащий как можно больше нулей. Разложением по этому столбцу (или строке) найти определитель.

▷ ПРИМЕР 1.8. Вычислить определитель четвертого порядка

$$|A| = \begin{vmatrix} 1 & 2 & 5 & 4 \\ 0 & 1 & 2 & -2 \\ 0 & 4 & 3 & 2 \\ 4 & 0 & 2 & 1 \end{vmatrix}.$$

Решение. Преобразуем матрицу так, чтобы в первом столбце все элементы, кроме одного (например, $a_{11} = 1$), обращались в 0. Для этого умножим элементы первой строки на (-4) и прибавим их соответственно к элементам четвертой строки. Раскладывая полученный определитель по элементам первого столбца, найдем

$$\begin{aligned} |A| &= \begin{vmatrix} 1 & 2 & 5 & 4 \\ 0 & 1 & 2 & -2 \\ 0 & 4 & 3 & 2 \\ 4 & 0 & 2 & 1 \end{vmatrix} = |A| = \begin{vmatrix} 1 & 2 & 5 & 4 \\ 0 & 1 & 2 & -2 \\ 0 & 4 & 3 & 2 \\ 0 & -8 & -18 & -15 \end{vmatrix} = \\ &= 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} 1 & 2 & -2 \\ 4 & 3 & 2 \\ -8 & -18 & -15 \end{vmatrix}. \end{aligned}$$

Полученный определитель третьего порядка можно вычислить по правилу треугольников или разложением по любой строке (столбцу), однако можно продолжить упрощение матрицы. “Обнулим” в матрице третьего порядка элементы первого столбца (кроме одного). Для этого умножим вторую строку на 2, одновременно вынося множитель $\frac{1}{2}$ за знак определителя (свойство 1). Затем к элементам третьей строки прибавим элементы второй строки, после чего ко второй строке прибавим первую, умноженную на (-8) :

$$|A| = \begin{vmatrix} 1 & 2 & -2 \\ 4 & 3 & 2 \\ -8 & -18 & -15 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 1 & 2 & -2 \\ 8 & 6 & 4 \\ -8 & -18 & -15 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 1 & 2 & -2 \\ 8 & 6 & 4 \\ 0 & -12 & -11 \end{vmatrix} =$$

$$= \frac{1}{2} \begin{vmatrix} 1 & 2 & -2 \\ 0 & -10 & 20 \\ 0 & -12 & -11 \end{vmatrix}.$$

Внесем общий множитель $\frac{1}{2}$ внутрь определителя, помножив на второй столбец, затем разложим его по элементам первого столбца:

$$|A| = \begin{vmatrix} 1 & 1 & -2 \\ 0 & -5 & 20 \\ 0 & -6 & -11 \end{vmatrix} = 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} -5 & 20 \\ -6 & -11 \end{vmatrix} = 1 \cdot 1 \cdot (55 + 120) = 175. \blacktriangleright$$

II. Зная, что определитель треугольной матрицы равен произведению элементов, стоящих вдоль главной диагонали, и используя свойства определителей, легко свести определитель любого порядка к определителю верхнетреугольной (нижнетреугольной) матрицы.

Алгоритм сведения любого определителя к определителю верхнетреугольной матрицы:

1. Найти в определителе элемент, равный единице ($a_{ij} = 1$). Если такого элемента в определителе нет, то, используя свойства определителя, добиться этого самим.

2. Перестановкой строк и столбцов местами, не забывая при этом менять знак определителя, перегнать этот элемент в левый верхний угол (т.е. теперь $a_{11} = 1$).

3. Используя свойство 8 определителя, добиться того, чтобы весь первый столбец, кроме первого единичного элемента, содержал нулевые элементы (т.е. $a_{21} = a_{31} = \dots = a_{n1} = 0$).

4. Не выполняя более никаких действий с первой строкой и первым столбцом (чтобы не испортить весь первый столбец), повторить пункты 1—3 для элемента a_{22} , добиваясь получения определителя вида

$$\begin{vmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 1 & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & a_{n3} & \dots & a_{nn} \end{vmatrix}.$$

5. Прделаав все то же самое с остальными элементами, стоящими вдоль главной диагонали определителя, после элементарных преобразований должен получиться определитель следующего вида:

$$\Delta = C \cdot \begin{vmatrix} 1 & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & 1 & b_{23} & \dots & b_{2n} \\ 0 & 0 & 1 & \dots & b_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix} = C \cdot 1 = C.$$

▷ ПРИМЕР 1.9. Вычислить определитель матрицы

$$A = \begin{pmatrix} 2 & -4 & 2 \\ 4 & 6 & -2 \\ 3 & -3 & 6 \end{pmatrix}.$$

Р е ш е н и е. Определитель имеет вид

$$\Delta = \begin{vmatrix} 2 & -4 & 2 \\ 4 & 6 & -2 \\ 3 & -3 & 6 \end{vmatrix}. \quad (1.10)$$

Вынесем из определителя (1.10) для его упрощения из первой строки 2, из второй строки 2, из третьей строки 3. Тогда согласно свойству 1 определителя получим

$$\Delta = \begin{vmatrix} 2 & -4 & 2 \\ 4 & 6 & -2 \\ 3 & -3 & 6 \end{vmatrix} = 2 \cdot 2 \cdot 3 \cdot \begin{vmatrix} 1 & -2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{vmatrix}.$$

Левый верхний элемент $a_{11} = 1$. “Обнулим” элементы, стоящие в первом столбце под ним. Для этого из второй строки поэлементно вычтем первую, предварительно умножив ее на число 2, т.е. $(2 \ 3 \ -1) - 2 \cdot (1 \ -2 \ 1) = (0 \ 7 \ -3)$ (как известно, эта операция определителя не изменит).

Из третьей строки вычтем первую $(1 \ -1 \ 2) - (1 \ -2 \ 1) = (0 \ 1 \ 1)$.

Получим определитель

$$\Delta = 12 \cdot \begin{vmatrix} 1 & -2 & 1 \\ 0 & 7 & -3 \\ 0 & 1 & 1 \end{vmatrix}. \quad (1.11)$$

В определителе (1.11) поменяем вторую строку с третьей местами в целях нахождения второго диагонального элемента, равного единице. Знак определителя при этом поменяется (согласно свойству 4). Затем из третьей строки поэлементно вычтем вторую, умноженную на число 7. Получим

$$\begin{aligned} \Delta &= -12 \cdot \begin{vmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 7 & -3 \end{vmatrix} = -12 \cdot \begin{vmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -10 \end{vmatrix} = \\ &= -12 \cdot 1 \cdot 1 \cdot (-10) = 120. \blacktriangleright \end{aligned}$$

§3. Обратная матрица

Теперь, когда мы познакомились с определителем квадратной матрицы, мы можем ввести еще одно важное понятие *обратной матрицы*. Известно, что для каждого числа $a \neq 0$ существует обратное число a^{-1} такое, что произведение $a \cdot a^{-1} = 1$. Для квадратных матриц тоже существует аналогичное понятие.

ОПРЕДЕЛЕНИЕ. Пусть A — квадратная матрица порядка n . Матрица A^{-1} называется *обратной* для матрицы A , если выполняется условие:

$$A \cdot A^{-1} = A^{-1} \cdot A = E, \quad (1.12)$$

т.е. при умножении обратной матрицы на исходную матрицу как справа, так и слева получается единичная матрица.

Из определения следует, что лишь квадратная матрица имеет обратную; в этом случае и обратная матрица является квадратной того же порядка.

Однако не каждая квадратная матрица имеет обратную. Необходимым и достаточным условием существования матрицы A^{-1} является требование $|A| \neq 0$, как и условие $a \neq 0$ для существования числа a^{-1} .

ОПРЕДЕЛЕНИЕ. Квадратная матрица, определитель которой равен нулю ($|A| = 0$), называется *вырожденной*, или *особенной*, в противном случае (при $|A| \neq 0$) она называется *невырожденной*, или *неособенной*.

Из определения следует, что произведение любых матриц, хотя бы одна из которых вырождена, — вырожденная матрица; результат произведения — невырожденная матрица, если заданы матрицы невырожденные.

□ **ТЕОРЕМА 1.1** (необходимое и достаточное условие существования обратной матрицы). *Обратная матрица A^{-1} существует (и единственна) тогда и только тогда, когда исходная матрица невырожденная.* ■

Алгоритм вычисления обратной матрицы:

1. Находим определитель исходной матрицы. Если $|A| = 0$, то матрица A — вырожденная и обратной матрицы A^{-1} не существует. Если $|A| \neq 0$, то матрица A — невырожденная и обратная матрица существует.

2. Находим алгебраические дополнения элементов исходной матрицы A_{ij} и составляем из них матрицу $\bar{A}$.

3. Находим матрицу, транспонированную к матрице $\bar{A}$, учитывая, что $A'_{ij} = A_{ji}$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, n$). Полученная таким образом матрица называется *присоединенной матрицей* $\tilde{A}$, т.е. $\tilde{A} = (\bar{A})'$.

4. Вычисляем обратную матрицу по формуле

$$A^{-1} = \frac{1}{|A|} \tilde{A} \quad (|A| \neq 0). \quad (1.13)$$

5. Проверить правильность вычисления обратной матрицы A^{-1} , исходя из ее определения $A \cdot A^{-1} = A^{-1} \cdot A = E$. (Пункт 5 не обязателен).

▷ ПРИМЕР 1.10. Найти матрицу, обратную к данной:

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix}.$$

Р е ш е н и е. 1. Определитель матрицы $|A| = 10 \neq 0$, т.е. матрица A — невырожденная и обратная матрица A^{-1} существует.

2. Находим алгебраические дополнения элементов матрицы A и составляем из них матрицу $\bar{A}$:

$$\bar{A} = \begin{pmatrix} 5 & -5 & -5 \\ 3 & 1 & -1 \\ -1 & 3 & 7 \end{pmatrix}.$$

3. Находим присоединенную матрицу $\tilde{A}$, транспонированную к $\bar{A}$, учитывая, что $A'_{ij} = A_{ji}$:

$$\tilde{A} = \begin{pmatrix} 5 & 3 & -1 \\ -5 & 1 & 3 \\ -5 & -1 & 7 \end{pmatrix}$$

4. Вычисляем обратную матрицу $A^{-1} = \frac{1}{|A|} \tilde{A}$:

$$A^{-1} = \frac{1}{10} \cdot \begin{pmatrix} 5 & 3 & -1 \\ -5 & 1 & 3 \\ -5 & -1 & 7 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{3}{10} & -\frac{1}{10} \\ -\frac{1}{2} & \frac{1}{10} & \frac{3}{10} \\ -\frac{1}{2} & -\frac{1}{10} & \frac{7}{10} \end{pmatrix}$$

5. Проверяем правильность вычисления обратной матрицы по формулам (1.12) (рекомендуем в этом убедиться самому читателю). ►

Свойства невырожденных матриц:

1. $|A^{-1}| = \frac{1}{|A|}$;
2. $(A^{-1})^{-1} = A$;
3. $(A^m)^{-1} = (A^{-1})^m$;
4. $(AB)^{-1} = B^{-1} \cdot A^{-1}$;
5. $(A^{-1})' = (A')^{-1}$.

Существует более простой способ вычисления обратной матрицы. С ним мы познакомимся чуть позже (см. подпараграф 5.3).

§4. Ранг матрицы

Пусть A — матрица типа $m \times n$ (не обязательно квадратная).

ОПРЕДЕЛЕНИЕ. *Рангом по строкам* матрицы A называется размерность линейного подпространства в R^n , порожденного m векторами-строками матрицы A .

ОПРЕДЕЛЕНИЕ. *Рангом по столбцам* матрицы A называется размерность линейного подпространства в R^m , порожденного n векторами-столбцами матрицы A .

Итак, матрицы размеров $m \times n$ можно рассматривать как системы, состоящие из m n -мерных векторов (или из n m -мерных векторов).

Так как здесь имеют место две совокупности векторов — векторы-строки и векторы-столбцы, то у матрицы, вообще говоря, имеется два ранга: строчный и столбцовый.

□ **ТЕОРЕМА 1.2.** *Строчный и столбцовый ранги любой матрицы равны.* ■

Значит, ранг любой матрицы размеров $m \times n$ можно искать как ранг одной из двух систем векторов: либо m векторов-строк, либо n векторов-столбцов.

ОПРЕДЕЛЕНИЕ. *Рангом по минорам* матрицы A называется наибольший порядок ненулевого минора этой матрицы.

В матрице A размеров $m \times n$ вычеркиванием каких-либо строк и столбцов можно вычленить квадратные подматрицы k -го порядка, где $k \leq \min\{m; n\}$. Определители таких подматриц называются *минорами k -го порядка* матрицы A .

Например, из матрицы $A_{3 \times 4}$ можно получить подматрицы первого, второго и третьего порядков. Максимальный ее ранг может быть равен или меньше 3.

З а м е ч а н и е. Все три приведенных выше определения дают одно и то же число, называемое *рангом матрицы A* : $\text{rang } A$ или $r = r(A)$. Для прямоугольной матрицы максимальный ранг

$r = \min\{m; n\}$. Для квадратной матрицы размеров $n \times n$ ее максимальный ранг не может превышать порядка матрицы, т.е. $r \leq n$.

Из определений следует:

а) ранг матрицы $A_{m \times n}$ не превосходит меньшего из ее размеров, т.е. $r(A) \leq \min\{m; n\}$.

б) $r(A) = 0$ тогда и только тогда, когда все элементы матрицы равны нулю, т.е. матрица A является нулевой матрицей: $A = 0$;

в) для квадратной матрицы n -го порядка $r(A) = n$ тогда и только тогда, когда матрица A — невырожденная.

▷ ПРИМЕР 1.11. Вычислить ранг матрицы

$$A_{4 \times 4} = \begin{pmatrix} 1 & -1 & 4 & 2 \\ 0 & 0 & 0 & 0 \\ 2 & -2 & 8 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Решение. Матрица A имеет четвертый порядок, поэтому $r(A) \leq 4$. Матрица A содержит нулевую строку, поэтому $|A| = 0$, значит $r(A) \leq 3$. Все подматрицы третьего порядка тоже содержат нулевую строку и поэтому имеют нулевые определители, значит $r(A) \leq 2$. Все подматрицы второго порядка имеют либо нулевую строку (вторую и четвертую), либо пропорциональные строки (первую и третью), поэтому тоже имеют нулевые определители; таким образом $r(A) \leq 1$. Поскольку матрица A содержит ненулевые элементы, т.е. невырожденные подматрицы первого порядка, то $r(A) = 1$. ►

Вычисления ранга матрицы. При вычислении ранга матрицы следует переходить от миноров меньших порядков к минорам больших порядков. Если уже найден минор k -го порядка M , отличный от нуля, то требуют вычисления лишь миноры $(k + 1)$ -го порядка, окаймляющие минор M ; если все они равны нулю, то ранг матрицы равен k .

Продemonстрируем метод окаймления миноров на примере.

▷ ПРИМЕР 1.12. Методом окаймления миноров найти ранг матрицы A :

$$A_{4 \times 5} = \begin{pmatrix} 1 & -1 & \boxed{\begin{vmatrix} 2 & 0 \\ 4 & -1 \end{vmatrix}} & 2 \\ 2 & -2 & \boxed{\begin{vmatrix} 4 & -1 \\ -2 & -3 \end{vmatrix}} & 1 \\ -1 & 1 & \boxed{\begin{vmatrix} -2 & -3 \\ 1 & -1 \end{vmatrix}} & 1 \\ 1 & -1 & 2 & -1 & -1 \end{pmatrix}.$$

Решение. Возьмем в этой матрице любой ненулевой элемент, например, $a_{13} = 2$. Его определитель *первого* порядка $\Delta_1 = 2 \neq 0$. Так как он не равен нулю, ранг исходной матрицы больше (или равен) *единицы*.

Окаймляем этот минор соседними элементами и строим из них

определитель $\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} = \begin{vmatrix} -1 & 2 \\ -2 & 4 \end{vmatrix} = 0$. Этот минор не годится, так как он нулевой.

Находим хотя бы один ненулевой минор *второго* порядка, окаймляющий элемент $a_{13} = 2$. Если переставить первый и второй столбцы местами, наш элемент будет окаймлен минором

$$\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0. \text{ Поскольку он тоже нулевой, находим другой.}$$

Определитель $\Delta_2 = \begin{vmatrix} 2 & 0 \\ 4 & -1 \end{vmatrix} = -2 \neq 0$. Так как он второго порядка, ранг матрицы равен как минимум *два*, т.е. $r(A) \geq 2$.

Попробуем найти ненулевой определитель *третьего* порядка, в который входит уже найденный нами ненулевой опреде-

литель $\Delta_2 = \begin{vmatrix} 2 & 0 \\ 4 & -1 \end{vmatrix} = -2 \neq 0$. Таких миноров много (например, M_2 получен заменой третьей строки на четвертую, а минор M_3 — перестановкой первого столбца со вторым):

$$1) \quad M_1 = \begin{vmatrix} -1 & \boxed{\begin{vmatrix} 2 & 0 \\ 4 & -1 \end{vmatrix}} \\ -2 & \boxed{\begin{vmatrix} 4 & -1 \\ -2 & -3 \end{vmatrix}} \\ 1 & \boxed{\begin{vmatrix} -2 & -3 \\ 1 & -1 \end{vmatrix}} \end{vmatrix} = 12 - 2 + 0 - (0 + 12 - 2) = 0;$$

$$2) M_2 = \begin{vmatrix} -1 & 2 & 0 \\ -2 & 4 & -1 \\ -1 & 2 & -1 \end{vmatrix} = 4 + 2 + 0 - (0 + 4 + 2) = 0;$$

$$3) M_3 = \begin{vmatrix} 1 & 2 & 0 \\ 2 & 4 & -1 \\ -1 & -2 & -3 \end{vmatrix} = -12 + 2 + 0 - (0 + 2 - 12) = 0;$$

$$4) M_4 = \begin{vmatrix} 2 & 0 & 2 \\ 4 & -1 & 1 \\ -2 & -3 & 1 \end{vmatrix} = -2 + 0 - 24 - (4 + 0 - 6) =$$

$$= -26 + 2 = -24 \neq 0.$$

Последний минор содержит внутри себя найденный нами ненулевой минор второго порядка, а значит, он также окаймляет

$$\Delta_2 = \begin{vmatrix} 2 & 0 \\ 4 & -1 \end{vmatrix} = -2 \neq 0. \text{ Поскольку } M_4 \neq 0, \text{ то мы нашли ненулевой}$$

минор третьего порядка. Можно найти и M_5 , если поменять местами третью и четвертые строки. Но нас его значение не интересует, поскольку ранг матрицы уже не меньше *трех*. Если же нам удастся найти хотя бы один минор четвертого порядка, содержащий внутри себя найденный нами ненулевой M_4 , значит, ранг будет равен четырем.

У нашей матрицы таких миноров два. Отметим один из них:

$$A = \begin{pmatrix} 1 & -1 & 2 & 0 & 2 \\ 2 & -2 & 4 & -1 & 1 \\ -1 & 1 & -2 & -3 & 1 \\ 1 & -1 & 2 & -1 & -1 \end{pmatrix}.$$

Второй минор получается при замене второго столбца первым. Можно проверить, что оба они — нулевые. Так как в матрице A нет ни одного ненулевого минора четвертого порядка, следовательно, ранг матрицы равен трем. ►

Метод окаймления миноров очень громоздкий и требует больших вычислительных затрат. Существует более экономный способ вычисления ранга, и связан он с элементарными преобразованиями матрицы.

Назовем **элементарными преобразованиями матрицы** следующие:

- 1) отбрасывание нулевой строки (столбца);
- 2) умножение всех элементов строки (столбца) матрицы на число, не равное нулю;
- 3) изменение порядка строк (столбцов) матрицы;
- 4) прибавление к каждому элементу одной строки (столбца) соответствующих элементов другой строки (столбца), умноженных на любое число;
- 5) транспонирование матрицы.

□ ТЕОРЕМА 1.3. *Ранг матрицы не изменяется при элементарных преобразованиях матрицы.* ■

При изучении свойств определителей было показано, что при преобразованиях квадратных матриц их определители либо сохраняются, либо умножаются на число, не равное нулю. В результате сохраняется наивысший порядок отличных от нуля миноров исходной матрицы, т.е. ее ранг не изменяется.

С помощью элементарных преобразований можно привести матрицу к так называемому ступенчатому виду, когда вычисление ее ранга не представляет труда.

ОПРЕДЕЛЕНИЕ. Матрица A называется **ступенчатой**, если она имеет вид:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} & \dots & a_{1k} \\ 0 & a_{22} & \dots & a_{2r} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr} & \dots & a_{rk} \end{pmatrix}, \quad (1.14)$$

где a_{ii} — числа, отличные от нуля; $r \leq k$.

З а м е ч а н и е. Условие $r \leq k$ всегда может быть достигнуто транспонированием матрицы.

Очевидно, что ранг ступенчатой матрицы равен r , так как имеется минор r -го порядка, не равный нулю:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ 0 & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr} \end{vmatrix} = a_{11} \cdot a_{22} \cdots a_{rr} \neq 0.$$

□ **ТЕОРЕМА 1.4.** Ранг ступенчатой матрицы равен числу ступенек этой матрицы. ■

Покажем на примере алгоритм вычисления ранга матрицы с помощью элементарных преобразований.

▷ **ПРИМЕР 1.13.** Найти ранг матрицы

$$A = \begin{pmatrix} 0 & 2 & -3 & 4 & 5 \\ 2 & 4 & 2 & -1 & 1 \\ 0 & -4 & 6 & -8 & -10 \\ 2 & 6 & -1 & 3 & 6 \end{pmatrix} \quad (1.15)$$

Р е ш е н и е. 1. Если $a_{11} = 0$, то при перестановке строк или столбцов добиваемся того, чтобы $a_{11} \neq 0$. В данном примере поменяем местами, например, первую и вторую строки матрицы:

$$\tilde{A} = \begin{pmatrix} 2 & 4 & 2 & -1 & 1 \\ 0 & 2 & -3 & 4 & 5 \\ 0 & -4 & 6 & -8 & -10 \\ 2 & 6 & -1 & 3 & 6 \end{pmatrix} \quad (1.16)$$

2. Так как в матрице (1.16) $a_{11} \neq 0$, избавимся от ненулевых элементов в первом столбце, кроме a_{11} . Используя элементарные преобразования, поэлементно к четвертой строке прибавим первую,

умноженную на число (-1) : $(2 \ 6 \ -1 \ 3 \ 6) - (2 \ 4 \ 2 \ -1 \ 1) = (0 \ 2 \ -3 \ 4 \ 5)$. Полученный первый ноль и есть тот результат, к которому мы стремились. Получим:

$$\tilde{A} = \begin{pmatrix} 2 & 4 & 2 & -1 & 1 \\ 0 & 2 & -3 & 4 & 5 \\ 0 & -4 & 6 & -8 & -10 \\ 2 & 6 & -1 & 3 & 6 \end{pmatrix} \sim \begin{pmatrix} 2 & 4 & 2 & -1 & 1 \\ 0 & 2 & -3 & 4 & 5 \\ 0 & -4 & 6 & -8 & -10 \\ 0 & 2 & -3 & 4 & 5 \end{pmatrix} \quad (1.17)$$

(знак $\sim$ означает равенство рангов матриц, но не самих матриц).

3. Мы добились нулевых элементов под первым элементом в первом столбце. У нас есть одна ступенька. Чтобы довести операцию преобразования исходной матрицы к ступенчатому виду до конца, нужно проделать все то же самое со вторым столбцом, уже не трогая первую строку и первый столбец, а спускаясь как бы по главной диагонали вниз. Так как в полученной матрице $a_{22} \neq 0$ (в нашем случае $a_{22} = 2 \neq 0$), то, умножая элементы третьей и четвертой строк на подходящие числа, добьемся того, чтобы все элементы второго столбца (кроме a_{12}, a_{22}) равнялись нулю. В матрице (1.17) к третьей и к четвертой строкам прибавим поочередно вторую, умноженную на число 2, затем — на (-1) :

$$\begin{pmatrix} 2 & 4 & 2 & -1 & 1 \\ 0 & 2 & -3 & 4 & 5 \\ 0 & -4 & 6 & -8 & -10 \\ 0 & 2 & -3 & 4 & 5 \end{pmatrix} \sim \begin{pmatrix} 2 & 4 & 2 & -1 & 1 \\ 0 & 2 & -3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (1.18)$$

Если в процессе преобразований получаются строки (или столбцы), целиком состоящие из нулей (как в данном примере), то отбрасываем эти строки (или столбцы):

$$\begin{pmatrix} 2 & 4 & 2 & -1 & 1 \\ 0 & 2 & -3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & 4 & 1 & -1 & 1 \\ 0 & 2 & -3 & 4 & 5 \end{pmatrix}. \quad (1.19)$$

Последняя матрица имеет ступенчатый вид и содержит *две* ступеньки, т.е. миноры второго порядка, не равные нулю, например:

$$\begin{vmatrix} 4 & 2 \\ 2 & -3 \end{vmatrix} = -16 \neq 0.$$

Поэтому ранг полученной ступенчатой матрицы, а следовательно, и данной матрицы равен 2. ►

Отметим, что при подсчете ранга в примере 1.13 мы обращаемся с матрицей не так, как раньше. Мы действуем с каждой отдельной строкой или столбцом, как с отдельной матрицей, поэтому не ставим знака равенства между теми матрицами, которые преобразовываем.

Подытоживая сказанное, заметим, что для подсчета ранга матрицы необходимо элементарными преобразованиями над строками или столбцами свести матрицу к ступенчатому виду и подсчитать число ступенек. Ранг будет равен этому числу.

На самом деле эта процедура с использованием элементарных преобразований нам уже хорошо известна. Аналогичные преобразования мы проделывали с определителем при сведении его к определителю верхнетреугольной матрицы (см. §2 главы 1) с той лишь разницей, что теперь не нужно выносить множитель и менять знак на противоположный, так как для подсчета числа ступенек это значения не имеет. Сравните пример 1.9 с нижеприведенным.

▷ ПРИМЕР 1.14. С использованием элементарных преобразований вычислить ранг матрицы $A = \begin{pmatrix} 2 & -4 & 2 \\ 4 & 6 & -2 \\ 3 & -3 & 6 \end{pmatrix}$.

Решение. Выполняем те же шаги, что и в примере 1.9:

$$\begin{aligned}
 A = \begin{pmatrix} 2 & -4 & 2 \\ 4 & 6 & -2 \\ 3 & -3 & 6 \end{pmatrix} &\sim \begin{pmatrix} 1 & -2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 7 & -3 \\ 0 & 1 & 1 \end{pmatrix} \sim \\
 &\sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 7 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -10 \end{pmatrix}. \quad (1.20)
 \end{aligned}$$

Ранг исходной матрицы равен количеству ступенек полученной матрицы (1.20) и, значит, он равен трем, т.е. $r(A) = 3$. ►

Свойства рангов матриц:

- 1) $r(A + B) \leq r(A) + r(B)$; 2) $r(AB) \leq \min \{r(A), r(B)\}$;
- 3) $r(A'A) = r(A) = r(AA')$, где $A'A$ — $n \times n$ — матрица, AA' — $m \times m$ — матрица;
- 4) $r(AB) = r(A)$, если B — квадратная матрица ранга n ;
- 5) $r(BA) = r(A)$, если B — квадратная матрица ранга m ;
- 6) $r(AB) \geq r(A) + r(B) - n$, где n — число столбцов матрицы A или строк матрицы B .

Линейная зависимость строк (столбцов) матрицы. Понятие ранга матрицы тесно связано с понятием линейной зависимости (независимости) ее строк или столбцов. (Далее материал излагается для строк матрицы; для столбцов матрицы изложение аналогично.)

Пусть дана матрица

$$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Обозначим ее строки следующим образом:

$$I_1 = (a_{11} \ a_{12} \ \dots \ a_{1n}), I_2 = (a_{21} \ a_{22} \ \dots \ a_{2n}), \dots, I_m = (a_{m1} \ a_{m2} \ \dots \ a_{mn}).$$

ОПРЕДЕЛЕНИЕ. Две строки матрицы называются **равными**, если равны их соответствующие элементы: $I_k = I_s$, если $a_{kj} = a_{sj}$, $j = 1, 2, \dots, n$.

Арифметические операции над строками матрицы (умножение строки на число, сложение строк) вводятся как операции, проводимые поэлементно:

$$\begin{aligned} \lambda I_k &= (\lambda a_{k1} \ \lambda a_{k2} \ \dots \ \lambda a_{kn}); \\ I_k + I_s &= [(a_{k1} + a_{s1})(a_{k2} + a_{s2}) \ \dots \ (a_{kn} + a_{sn})]. \end{aligned}$$

ОПРЕДЕЛЕНИЕ. Строка I называется **линейной комбинацией** строк $I_1, I_2, \dots, I_s$ матрицы, если она равна сумме произведений этих строк на произвольные действительные числа:

$$I = \lambda_1 I_1 + \lambda_2 I_2 + \dots + \lambda_s I_s, \quad (1.21)$$

где $\lambda_1, \lambda_2, \dots, \lambda_s$ — любые числа.

ОПРЕДЕЛЕНИЕ. Строки матрицы $I_1, I_2, \dots, I_s$ называются **линейно зависимыми**, если существуют такие числа $\lambda_1, \lambda_2, \dots, \lambda_s$, не равные одновременно нулю, что линейная комбинация строк матрицы равна нулевой строке:

$$\lambda_1 I_1 + \lambda_2 I_2 + \dots + \lambda_s I_s = O, \quad (1.22)$$

где $O = (0 \ 0 \ \dots \ 0)$.

Линейная зависимость строк матрицы означает, что хотя бы одна строка матрицы является линейной комбинацией остальных.

ОПРЕДЕЛЕНИЕ. Если некоторые из строк $I_1, I_2, \dots, I_s$ составляют сами по себе линейно зависимую систему, то вся система строк $I_1, I_2, \dots, I_s$ — **линейно зависима**.

ОПРЕДЕЛЕНИЕ. Если линейная комбинация строк (1.21) равна нулю тогда и только тогда, когда $\lambda_1 = \lambda_2 = \dots = \lambda_s = 0$, то строки $I_1, I_2, \dots, I_s$ называются **линейно независимыми**.

□ ТЕОРЕМА 1.5 (о ранге матрицы). *Ранг матрицы равен максимальному числу ее линейно независимых строк или столбцов, через которые линейно выражаются все остальные ее строки (столбцы).* ■

Рассмотрим ранг матрицы A размеров $m \times n$: $r \leq \min \{m; n\}$. Всякий ненулевой минор r -го порядка будем называть *базисным*. Пусть для определенности это минор

$$M = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \dots & \dots & \dots & \dots \\ a_{r1} & a_{r2} & \dots & a_{rr} \end{vmatrix} \neq 0.$$

Тогда строки матрицы $I_1, I_2, \dots, I_r$ линейно независимы. Действительно, предположим противное, т.е. — одна из этих строк, например I_r , является линейной комбинацией остальных:

$$I_r = I_1\lambda_1 + I_2\lambda_2 + \dots + \lambda_{r-1}I_{r-1}.$$

Вычтем из элементов r -й строки элементы первой строки, умноженные на λ_1 , элементы второй строки, умноженные на λ_2 , и так далее вплоть до элементов $(r-1)$ -й строки, умноженных на λ_{r-1} . На основании свойства 8 (см. подпараграф 2.2 главы 1) при таких преобразованиях матрицы ее определитель Δ не изменится, но так как теперь r -я строка будет состоять из одних нулей, то $\Delta = 0$. Однако это противоречит условию $\Delta \neq 0$, следовательно, наше предположение о том, что строки $I_1, I_2, \dots, I_r$ матрицы линейно зависимы, неверно.

Строки $I_1, I_2, \dots, I_r$, участвующие в образовании базисного минора, назовем *базисными*.

Тогда любые $r+1$ строк матрицы линейно зависимы, т.е. любая строка выражается через базисные.

□ ТЕОРЕМА 1.6. *Всякая строка (столбец) матрицы является линейной комбинацией ее базисных строк (столбцов).* ■

▷ ПРИМЕР 1.15. Найти число линейно независимых строк матрицы A :

$$A = \begin{pmatrix} 1 & -1 & 2 & 0 & 2 \\ 2 & -2 & 4 & -1 & 1 \\ -1 & 1 & -2 & -3 & 1 \\ 1 & -1 & 2 & -1 & -1 \end{pmatrix}. \quad (1.23)$$

Р е ш е н и е. Поиск числа линейно независимых базисных строк матрицы сводится к операции вычисления ранга матрицы (см. пример 1.13).

Сведем матрицу A к ступенчатой матрице с использованием элементарных преобразований над строками. Тогда

$$\begin{aligned} A &= \begin{pmatrix} 1 & -1 & 2 & 0 & 2 \\ 2 & -2 & 4 & -1 & 1 \\ -1 & 1 & -2 & -3 & 1 \\ 1 & -1 & 2 & -1 & -1 \end{pmatrix} \sim \begin{matrix} \tilde{I}_2 = I_2 - 2I_1 \\ \tilde{I}_3 = I_1 + I_3 \\ \tilde{I}_4 = I_4 - I_1 \end{matrix} \sim \\ &\sim \begin{pmatrix} 1 & -1 & 2 & 0 & 2 \\ 0 & 0 & 0 & -1 & -3 \\ 0 & 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & -1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 12 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (1.24)$$

У результирующей матрицы (1.24) всего три ступеньки, значит, $r(A) = 3$.

Так как число ненулевых строк в исходной матрице A (1.23) равно четырем, а в результирующей матрице (1.24) — трем, значит, только три строки из четырех, имеющихся в исходной матрице, линейно независимы, а четвертая линейно выражается через другие, т.е. линейно зависима.

Чтобы убедиться в справедливости этого, внимательно посмотрите на вторую и четвертую строки исходной матрицы. Если поэлементно из второй строки отнять первую, то получим четвертую:

$$\begin{aligned} I_4 &= I_2 - I_1 = (2 \quad -2 \quad 4 \quad -1 \quad 1) - (1 \quad -1 \quad 2 \quad 0 \quad 2) = \\ &= (1 \quad -1 \quad 2 \quad -1 \quad -1). \end{aligned}$$

Запишем это так: $I_4 - I_2 + I_1 = O$; $1 \cdot I_4 + 0 \cdot I_3 - 1 \cdot I_2 + 1 \cdot I_1 = O$ и сравним с формулой (1.22). ►

Понятие ранга матрицы очень важно для решения и исследования целого ряда математических и прикладных экономических задач, в частности, сводящихся к решению систем линейных алгебраических уравнений.

Так как число столбцов матрицы $A_{m \times n}$ равно числу строк матрицы $X_{n \times 1}$, то их произведение

$$AX = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

есть матрица-столбец. Элементами полученной матрицы являются левые части системы (1.25). На основании определения равенства матриц систему (1.25) можно записать в виде:

$$AX = B. \quad (1.27)$$

ОПРЕДЕЛЕНИЕ. Совокупность таких n чисел ($x_1 = k_1, x_2 = k_2, \dots, x_n = k_n$), при подстановке которых в систему каждое уравнение системы обращается в верное равенство, называется **решением системы** линейных алгебраических уравнений (1.25).

ОПРЕДЕЛЕНИЕ. Если система уравнений имеет хотя бы одно решение, то она называется *совместной*, и — *несовместной*, если решений она не имеет.

ОПРЕДЕЛЕНИЕ. Совместная система уравнений называется *определенной*, если она обладает одним единственным решением, и *неопределенной*, если она имеет бесконечно много решений.

Например, система уравнений $\begin{cases} x_1 + x_2 = 2, \\ 2x_1 - 3x_2 = -6 \end{cases}$ — совместная

и определенная, так как имеет единственное решение $(0, 2)$; система

$$\begin{cases} 3x_1 - x_2 = 1, \\ 3x_1 - x_2 = 5 \end{cases} \text{ — несовместная; а система уравнений}$$

$$\begin{cases} x_1 + x_2 = 1, \\ -2x_1 - 2x_2 = -2 \end{cases} \text{ — совместная и неопределенная, так как имеет}$$

бесконечное множество решений ($x_1 = c$; $x_2 = 1 - c$, где c — любое число).

ОПРЕДЕЛЕНИЕ. Если две (или несколько) системы уравнений имеют одно и то же множество решений, то они называются *равносильными*, или *эквивалентными*.

5.2. Решение системы n линейных уравнений с n неизвестными. Теорема Крамера и метод обратной матрицы

ОПРЕДЕЛЕНИЕ. Рассмотрим систему (1.25), у которой число уравнений равно числу неизвестных, т.е. $m = n$. Тогда матрица системы является квадратной, а ее определитель $\Delta = |A|$ называется *определителем системы*.

Рассмотрим решение системы двух уравнений с двумя неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2, \end{cases} \quad (1.28)$$

в которой хотя бы один из коэффициентов при неизвестных отличен от нуля.

Для решения этой системы исключим переменную x_2 , умножив первое уравнение на a_{22} , второе — на $(-a_{12})$ и сложив их. Затем исключим переменную x_1 , умножив первое уравнение на $(-a_{21})$, второе — на a_{11} и также сложив их. В результате получим систему

$$\begin{cases} (a_{11}a_{22} - a_{21}a_{12})x_1 = b_1a_{22} - b_2a_{12}, \\ (a_{11}a_{22} - a_{21}a_{12})x_2 = a_{11}b_2 - a_{21}b_1. \end{cases} \quad (1.29)$$

Выражение в скобках есть определитель системы

$$\Delta = a_{11}a_{22} - a_{21}a_{12} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Введем следующие обозначения:

$$\Delta_1 = b_1 a_{22} - b_2 a_{12} = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix},$$

$$\Delta_2 = a_{11} b_2 - a_{21} b_1 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix},$$

система (1.29) при этом примет вид:

$$\begin{cases} \Delta \cdot x_1 = \Delta_1, \\ \Delta \cdot x_2 = \Delta_2. \end{cases} \quad (1.30)$$

Из полученной системы следует, что если определитель системы $\Delta \neq 0$, то система (1.28) имеет единственное решение, определяемое по формулам:

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}.$$

Если $\Delta = 0$, а $\Delta_1 \neq 0$ (или $\Delta_2 \neq 0$), то система (1.28) несовместная,

так как в этом случае приводится к виду:
$$\begin{cases} 0 \cdot x_1 = \Delta_1, \\ 0 \cdot x_2 = \Delta_2. \end{cases}$$

Если $\Delta = \Delta_1 = \Delta_2 = 0$, то система (1.28) неопределенная и имеет бесконечное множество решений, так как в этом случае приво-

дится к виду
$$\begin{cases} 0 \cdot x_1 = 0, \\ 0 \cdot x_2 = 0. \end{cases}$$

□ ТЕОРЕМА 1.7 (Крамера). Пусть $\Delta \neq 0$ — определитель матрицы системы A , а Δ_j — определитель матрицы, получаемой из матрицы A заменой j -го столбца столбцом свободных членов. Тогда система имеет единственное решение, определяемое по формулам

$$x_j = \frac{\Delta_j}{\Delta}, \quad j = 1, 2, \dots, n. \quad \blacksquare \quad (1.31)$$

Формулы (1.31) получили название **формул Крамера**.

Метод обратной матрицы. Чтобы получить решения системы (1.25) при $m = n$, предположим, что квадратная матрица системы $A_{n \times n}$ невырожденная, т.е. ее определитель $|A| \neq 0$. В этом случае существует обратная матрица A^{-1} .

Умножая слева обе части матричного равенства (1.27) на матрицу A^{-1} , получим $A^{-1}(AX) = A^{-1}B$. Так как $A^{-1}(AX) = (A^{-1}A)X = EX = X$, то решением системы методом обратной матрицы будет матрица-столбец

$$X = A^{-1}B. \quad (1.32)$$

Таким образом, чтобы найти решение системы уравнений, нужно обратную матрицу умножить на столбец свободных членов. Полученный в результате столбец и есть решение.

▷ ПРИМЕР 1.16. Решить систему уравнений

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ 2x_1 + 3x_2 - x_3 = 0, \\ x_1 - x_2 + 2x_3 = 2. \end{cases}$$

а) методом обратной матрицы; б) по формулам Крамера.

Решение. а) Обозначим

$$A = \begin{pmatrix} 1 & -2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix}; \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}; \quad B = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Тогда в матричной форме данная система имеет вид: $AX = B$. Найдем определитель $|A| = 10$. Так как $|A| \neq 0$, то матрица A невырожденная и существует обратная матрица A^{-1} . Матрицу A^{-1} найдем по алгоритму, приведенному в § 3 главы 1. Получим

$$A^{-1} = \frac{1}{10} \cdot \begin{pmatrix} 5 & 3 & -1 \\ -5 & 1 & 3 \\ -5 & -1 & 7 \end{pmatrix}$$

(см. пример 1.10).

Теперь по формуле (1.32)

$$X = A^{-1}B = \frac{1}{10} \cdot \begin{pmatrix} 5 & 3 & -1 \\ -5 & 1 & 3 \\ -5 & -1 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 3 \\ 1 \\ 9 \end{pmatrix} = \begin{pmatrix} 3/10 \\ 1/10 \\ 9/10 \end{pmatrix}.$$

Таким образом, $x_1 = \frac{3}{10}$, $x_2 = \frac{1}{10}$, $x_3 = \frac{9}{10}$.

б) Найдем определитель системы $\Delta = |A| = 10$. Так как $\Delta \neq 0$, то по теореме Крамера система имеет единственное решение.

Вычислим определители матриц $\Delta_1, \Delta_2, \Delta_3$, полученных из матрицы A заменой соответственно первого, второго и третьего столбцов столбцом свободных членов:

$$\Delta_1 = \begin{vmatrix} 1 & -2 & 1 \\ 0 & 3 & -1 \\ 2 & -1 & 2 \end{vmatrix} = 3; \Delta_2 = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & -1 \\ 1 & 2 & 2 \end{vmatrix} = 1; \Delta_3 = \begin{vmatrix} 1 & -2 & 1 \\ 2 & 3 & 0 \\ 1 & -1 & 2 \end{vmatrix} = 9.$$

(рекомендуем читателю вычислить самостоятельно). Теперь по формулам Крамера (1.31)

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{3}{10}; \quad x_2 = \frac{\Delta_2}{\Delta} = \frac{1}{10}; \quad x_3 = \frac{\Delta_3}{\Delta} = \frac{9}{10}.$$

После решения системы (любым способом) рекомендуем сделать проверку, подставив найденные значения неизвестных в уравнения системы, и убедиться в том, что они обращаются в верные равенства. ►

Иногда приходится сталкиваться с применением метода обратной матрицы для решения матричных уравнений.

▷ ПРИМЕР 1.17. Даны матрицы

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}; B = \begin{pmatrix} 1 & 2 & 7 \\ 0 & 4 & 8 \end{pmatrix}; C = \begin{pmatrix} 4 & 7 \\ 0 & 1 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}.$$

Решить матричные уравнения: а) $AX = B$; б) $XA = C$.

Решение. а) Так как матрица A – невырожденная ($|A| \neq 0$), то решение уравнения $AX = B$ находится по формуле $X = A^{-1}B$.

Найдем обратную матрицу A^{-1} согласно алгоритму, приведенному в § 3 главы 1: $|A| = 1$. Матрица A' , транспонированная к A ,

имеет вид $A' = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$, а матрица $\tilde{A}$ из алгебраических дополнений

элементов матрицы A' есть $\tilde{A} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$

Следовательно, $A^{-1} = \frac{1}{|A|} \tilde{A} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$ и искомая матрица

$$X = A^{-1}B = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 7 \\ 0 & 4 & 8 \end{pmatrix} = \begin{pmatrix} 1 & -6 & -9 \\ -1 & 10 & 17 \end{pmatrix}.$$

б) Так как матрица A — невырожденная, то существует матрица A^{-1} и $X = CA^{-1}$.

Следовательно,

$$X = \begin{pmatrix} 4 & 7 \\ 0 & 1 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} -3 & 13 \\ -1 & 3 \\ -1 & 5 \\ -1 & 6 \end{pmatrix} \blacktriangleright$$

При решении систем n линейных уравнений с n неизвестными по формулам Крамера и методом обратной матрицы нам приходится сталкиваться с вычислением определителей и нахождением обратной матрицы, что является существенным недостатком этих методов. Особенно, если матрицы системы имеют достаточно высокий порядок. Поэтому эти методы не всегда могут быть использованы для решения реальных экономических задач, сводящихся к системам с большим числом уравнений и неизвестных.

5.3. Методы Гаусса и Жордана-Гаусса

Рассмотрим решение системы (1.25) m линейных уравнений с n неизвестными. Заметим, что матрица коэффициентов системы не обязательно должна быть квадратной. Предлагаемые методы решения систем линейных алгебраических уравнений сводятся к элементарным преобразованиям над уравнениями системы.

Преобразования Гаусса удобно проводить, осуществляя преобразования не самих уравнений, а *строк* матрицы их коэффициентов. Рассмотрим матрицу

$$A_1 = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right), \quad (1.33)$$

называемую *расширенной матрицей системы* (1.25), так как в нее, кроме коэффициентов матрицы A системы (1.25), дополнительно включен столбец свободных членов B .

Метод Гаусса (метод последовательного исключения неизвестных. Рассмотрим систему (1.25) в случае $m = n$. Суть метода Гаусса заключается в том, что с помощью элементарных преобразований расширенная матрица системы (1.33) приводится к равносильной матрице ступенчатого вида по алгоритму поиска ранга матрицы (см. пример 1.13). Это и есть *прямой ход метода Гаусса*.

На основании полученной ступенчатой матрицы составляется новая система уравнений, равносильная исходной, из которой последовательно, начиная с последнего уравнения, находятся все неизвестные; это суть *обратного хода метода Гаусса*.

▷ ПРИМЕР 1.18. Решить систему уравнений методом Гаусса:

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ 2x_1 + 3x_2 - x_3 = -1, \\ x_1 - x_2 + 2x_3 = 0. \end{cases}$$

Р е ш е н и е. Расширенная матрица системы имеет вид

$$A_1 = \left(\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 2 & 3 & -1 & -1 \\ 1 & -1 & 2 & 0 \end{array} \right)$$

Используя элементарные преобразования (см. пример 1.14), сведем эту матрицу к ступенчатой, не забывая при этом, что теперь в строке уже не три, а четыре элемента.

$$\begin{aligned} A_1 &= \left(\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 2 & 3 & -1 & -1 \\ 1 & -1 & 2 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 7 & -3 & -3 \\ 0 & 1 & 1 & -1 \end{array} \right), \\ &\sim \left(\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 7 & -3 & -3 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & -2 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -10 & 4 \end{array} \right). \end{aligned}$$

Вспомнив, что до черты стоят коэффициенты системы уравнений, а после нее — столбец свободных членов, выпишем получившуюся систему уравнений:

$$\begin{cases} x_1 - 2x_2 + x_3 = 1, \\ x_2 + x_3 = -1, \\ -10x_3 = 4. \end{cases}$$

Теперь применим обратный ход метода Гаусса. Из последнего уравнения полученной системы найдем $x_3 = \frac{4}{-10} = -\frac{2}{5}$. Из второго

уравнения найдем $x_2 = -1 - x_3 = -1 + \frac{2}{5} = -\frac{3}{5}$. Аналогично найдем x_1

из первого уравнения, подставив в него уже найденные два неиз-

вестных: $x_1 = 1 + 2x_2 - x_3 = 1 + 2\left(-\frac{3}{5}\right) + \frac{2}{5} = \frac{1}{5}.$

Получим решение системы $\left(\frac{1}{5}; -\frac{3}{5}; -\frac{2}{5}\right)$

Предлагаем читателю убедиться, что найденные числа образуют решение данной системы. ►

Метод Жордана-Гаусса. Суть метода Жордана-Гаусса заключается в построении такой ступенчатой матрицы, вдоль главной диагонали которой будут стоять лишь одни *единицы*. Затем, не производя обратного хода, как это было сделано в методе Гаусса, нужно продолжать элементарными преобразованиями *снизу вверх* обращать в нули элементы, стоящие над главной диагональю, до тех пор, пока слева до черты в расширенной матрице не будет стоять единичная матрица. Тогда справа получим решение системы уравнений. Это один из самых простых и изящных способов решения систем линейных уравнений.

▷ ПРИМЕР 1.19. Решить систему уравнений методом Жордана-Гаусса:

$$\begin{cases} x_1 + 2x_2 + 3x_3 - 2x_4 = 6, \\ 2x_1 + 4x_2 - 2x_3 - 3x_4 = 18, \\ 3x_1 + 2x_2 - x_3 + 2x_4 = 4, \\ 2x_1 - 3x_2 + 2x_3 + x_4 = -8. \end{cases}$$

Решение. Расширенная матрица системы имеет вид:

$$A_1 = \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 2 & 4 & -2 & -3 & 18 \\ 3 & 2 & -1 & 2 & 4 \\ 2 & -3 & 2 & 1 & -8 \end{array} \right).$$

Преобразуем первый столбец:

$$\tilde{I}_2 = I_2 - 2I_1; \quad \tilde{I}_3 = I_3 - 3I_1; \quad \tilde{I}_4 = I_4 - 2I_1, \quad \tilde{I}_2 = I_3,$$

в результате получим

$$A_1 = \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 2 & 4 & -2 & -3 & 18 \\ 3 & 2 & -1 & 2 & 4 \\ 2 & -3 & 2 & 1 & -8 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & -4 & -10 & 8 & -14 \\ 0 & 0 & -8 & 1 & 6 \\ 0 & -7 & -4 & 5 & -20 \end{array} \right).$$

Преобразовываем далее:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & -4 & -10 & 8 & -14 \\ 0 & 0 & -8 & 1 & 6 \\ 0 & -7 & -4 & 5 & -20 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 4 & 10 & -8 & 14 \\ 0 & 0 & -8 & 1 & 6 \\ 0 & -7 & -4 & 5 & -20 \end{array} \right) \sim$$

$\tilde{I}_2 = I_2 \cdot (-1) \qquad \tilde{I}_4 = I_4 + I_2$

$$\sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 4 & 10 & -8 & 14 \\ 0 & 0 & -8 & 1 & 6 \\ 0 & -3 & 6 & -3 & -6 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 1 & 16 & -11 & 8 \\ 0 & 0 & -8 & 1 & 6 \\ 0 & -1 & 2 & -1 & -2 \end{array} \right) \sim$$

$\tilde{I}_2 = \tilde{I}_2 + I_4, \quad \tilde{I}_4 = I_4 / 3 \qquad \tilde{I}_4 = I_4 + I_2$

$$\sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 1 & 16 & -11 & 8 \\ 0 & 0 & -8 & 1 & 6 \\ 0 & 0 & 18 & -12 & 6 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 1 & 16 & -11 & 8 \\ 0 & 0 & -8 & 1 & 6 \\ 0 & 0 & 2 & -10 & 18 \end{array} \right) \sim$$

$\tilde{I}_4 = I_4 + 2I_3 \qquad I_4 \leftrightarrow I_3$

$$\begin{aligned}
& \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 1 & 16 & -11 & 8 \\ 0 & 0 & 2 & -10 & 18 \\ 0 & 0 & -8 & 1 & 6 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 1 & 16 & -11 & 8 \\ 0 & 0 & 2 & -10 & 18 \\ 0 & 0 & 0 & -39 & 78 \end{array} \right) \sim \\
& \quad \tilde{I}_4 = I_4 + 4I_3 \quad \tilde{I}_3 = \frac{1}{2}I_3, \quad \tilde{I}_4 = I_4 / (-39) \\
& \sim \left(\begin{array}{cccc|c} 1 & 2 & 3 & -2 & 6 \\ 0 & 1 & 16 & -11 & 8 \\ 0 & 0 & 1 & -5 & 9 \\ 0 & 0 & 0 & 1 & -2 \end{array} \right) \\
& \quad \tilde{I}_3 = I_3 + 5I_4, \quad \tilde{I}_2 = I_2 + 11I_4, \quad \tilde{I}_1 = I_1 + 2I_4
\end{aligned}$$

Теперь работаем с нижней строкой, содержащей единицу в четвертом столбце, и с ее помощью “обнуляем” весь столбец. Не забываем выполнять действия со строчкой, содержащей не четыре, а пять элементов. Далее получаем

$$\begin{aligned}
& \left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 2 \\ 0 & 1 & 16 & 0 & -14 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -2 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 2 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -2 \end{array} \right) \sim \\
& \quad \tilde{I}_2 = I_2 - 16I_3, \quad \tilde{I}_1 = I_1 - 3I_3 \quad \tilde{I}_1 = I_1 - 2I_2
\end{aligned}$$

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -2 \end{array} \right)$$

Справа получили столбец решений. Таким образом:
 $x_1 = 1, x_2 = 2, x_3 = -1, x_4 = -2$. ►

Метод Жордана-Гаусса вычисления обратной матрицы. Используя метод Жордана-Гаусса, можно вычислять обратные матрицы менее трудоемким способом, чем через алгебраические дополнения.

Возьмем нашу обычную квадратную матрицу $A_{n \times n}$ и припишем к ней справа единичную матрицу той же размерности:

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 & 1 \end{array} \right).$$

Элементарными преобразованиями над строками, используя алгоритм метода Жордана-Гаусса, приведем левую часть к единичной матрице:

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & 1 & 0 & \dots & b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{array} \right).$$

Матрица $B_{n \times n}$, полученная справа, и будет обратной к $A_{n \times n}$.

ПРИМЕР 1.20. Найти обратную матрицу к матрице $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Решение. Припишем справа к матрице единичную матрицу той же размерности и применим к полученной двойной матрице преобразования над строками:

$$\begin{aligned} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array} \right) &\sim \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array} \right) \sim \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right) \sim \\ &\sim \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array} \right). \end{aligned}$$

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \text{ Сделав проверку, получим } A^{-1}A = A A^{-1} = E.$$

5.4. Решение систем общего вида

Мы уже выяснили, что методы Крамера и обратной матрицы применимы лишь к квадратной невырожденной системе, методы Гаусса и Жордана-Гаусса применимы к любой системе линейных уравнений (m уравнений, n неизвестных).

Пусть задана система линейных алгебраических уравнений общего вида (1.25), где $m \leq n$, т.е. число неизвестных не меньше числа уравнений. Представим общий порядок решения этой системы.

Необходимо сначала определить совместность системы, т.е. сравнить ранги матрицы системы и расширенной матрицы.

Вопрос о разрешимости системы (1.25) в общем виде рассматривается в следующей теореме.

□ ТЕОРЕМА 1.8. (Кронекера-Капелли). Система линейных уравнений совместна тогда и только тогда, когда ранг матрицы системы равен рангу расширенной матрицы этой системы ($r(A) = r(A_1)$). ■

Например, системы $\begin{cases} x_1 + 5x_2 = 1, \\ x_1 + 5x_2 = 7 \end{cases}$ и $\begin{cases} x_1 + 3x_2 = 2, \\ 0 \cdot x_2 = 1 \end{cases}$ — несовмест-

ные, так как $r(A)=1$, $r(A_1)=2$, т.е. $r(A) \neq r(A_1)$.

Для совместных систем линейных уравнений верны следующие теоремы.

□ ТЕОРЕМА 1.9. Если ранг матрицы совместной системы равен числу неизвестных, т.е. $r = n$, то система (1.25) определенная и имеет единственное решение. ■

Система $\begin{cases} x_1 + 2x_2 = 7, \\ x_1 + x_2 = 4 \end{cases}$ — определенная, так как $r(A) = r(A_1) = 2$.

Она имеет одно решение: $x_1 = 1$, $x_2 = 3$.

□ ТЕОРЕМА 1.10. Если ранг матрицы совместной системы меньше числа неизвестных, т.е. $r < n$, то система (1.25) неопределенная и имеет бесконечное множество решений. ■

$$\text{Система } \begin{cases} 3x_1 - x_2 = 1, \\ 6x_1 - 2x_2 = 2 \end{cases} \text{ — неопределенная, так как } r(A) = r(A_1) = 1,$$

а $n = 2$, т.е. $n > r$; она имеет бесконечно много решений: $x_1 = k$, $x_2 = 3k - 1$, k — любое число.

ОПРЕДЕЛЕНИЕ. Решение системы (1.25), в котором все $n - r$ неосновных переменных равны нулю, называется **базисным**.

З а м е ч а н и е. Система будет несовместной, если в процессе преобразований мы получим уравнение, в котором коэффициенты при всех неизвестных равны нулю, а свободный член отличен от нуля; если же такое уравнение отсутствует, система будет совместной. Совместная система уравнений будет определенной, если она приводится к треугольному виду, и неопределенной, если она приводится к трапецеидальному виду.

Это связано с проведенным нами ранее (см. §4 главы 1) исследованием того, что ранг матрицы равен максимальному числу ее линейно независимых строк. Поэтому, если строки расширенной матрицы A_1 , т.е. уравнения системы (1.25), линейно независимы, то ранг матрицы A_1 равен числу ее уравнений, т.е. $r(A_1) = m$; если строки линейно зависимы, то $r(A_1) < m$.

▷ ПРИМЕР 1.21. Методом Гаусса решить систему уравнений

$$\begin{cases} 2x_1 - 3x_2 - x_3 + x_4 = 2, \\ x_1 - 2x_2 + x_3 + x_4 = 1, \\ 3x_1 - 5x_2 + 2x_4 = 3 \end{cases}$$

и сделать выводы: является ли она совместной и определенной или нет.

Р е ш е н и е. Преобразуем расширенную матрицу системы (для удобства вычислений берем в качестве первой строки коэффици-

енты второго уравнения, у которого коэффициент при x_1 равен 1). Из второй строки поэлементно отнимем первую, умноженную на число 2, из третьей отнимем первую, умноженную на число 3. Далее отнимем из третьей строки вторую. Полученную нулевую строку отбросим.

В результате получим

$$\begin{pmatrix} 1 & -2 & 1 & 1 & | & 1 \\ 2 & -3 & -1 & 1 & | & 2 \\ 3 & -5 & 0 & 2 & | & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 & 1 & | & 1 \\ 0 & 1 & -3 & -1 & | & 0 \\ 0 & 1 & -3 & -1 & | & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & -2 & 1 & 1 & | & 1 \\ 0 & 1 & -3 & -1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 1 & 1 & | & 1 \\ 0 & 1 & -3 & -1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix},$$

т.е. ранг матрицы системы равен числу ступенек, а значит, $r = 2$.

Эта система имеет решения, так как столбец свободных членов, стоящий после черты, ранга не увеличил, т. е. $r(A) = r(A_1) = 2$, и система *с о в м е с т н а*.

З а м е ч а н и е. Сравните, например, две расширенные матрицы:

$$\begin{pmatrix} 1 & 2 & -2 & 3 & | & -6 \\ 0 & -5 & 5 & -7 & | & 17 \end{pmatrix} \text{ и } \begin{pmatrix} 1 & 2 & -2 & 3 & | & -6 \\ 0 & -5 & 5 & -7 & | & 17 \\ 0 & 0 & 0 & 0 & | & 2 \end{pmatrix}.$$

Во втором случае ранг исходной матрицы, стоящей до черты, равен двум, а ранг расширенной матрицы (включая столбец после черты) равен трем, т.е. система несовместна.

Число неизвестных у исходной системы равно четырём:

$$n = 4 (x_1, x_2, x_3, x_4).$$

Так как $r < n$ ($2 < 4$), т.е. ранг меньше числа неизвестных, значит, система неопределённая и имеет бесконечное множество решений.

Составим систему, соответствующую полученной матрице.

Оставляем в левой части переменные x_1, x_2 , которые берем за *основные*, т.е. *базисные* (определитель из коэффициентов при них

(базисный минор) отличен от нуля: $\begin{vmatrix} 1 & -2 \\ 0 & 1 \end{vmatrix} \neq 0$). Остальные

неосновные переменные x_3, x_4 , называемые еще *параметрическими*, или *свободными*, перенесем в правые части уравнений. В результате получим систему

$$\begin{cases} x_1 - 2x_2 = 1 - x_3 - x_4, \\ x_2 = 3x_3 + x_4, \end{cases}$$

откуда

$$x_2 = 3x_3 + x_4,$$

$$x_1 = 1 - x_3 - x_4 + 2(3x_3 + x_4) = 1 + 5x_3 + x_4.$$

Задавая переменным произвольные значения $x_3 = c_1, x_4 = c_2$, найдем бесконечное множество решений системы: $x_1 = 1 + 5c_1 + c_2, x_2 = 3c_1 + c_2, x_3 = c_1, x_4 = c_2$; решение $x_1 = 1, x_2 = 0, x_3 = 0, x_4 = 0$ — базисное. ►

▷ ПРИМЕР 1.22. Решить систему уравнений и сделать выводы о ее совместности:

$$\begin{cases} x_1 + 2x_2 - x_3 = 7, \\ 2x_1 - 3x_2 + x_3 = 3, \\ 4x_1 + x_2 - x_3 = 16. \end{cases}$$

Р е ш е н и е. Преобразуем расширенную матрицу системы:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 7 \\ 2 & -3 & 1 & 3 \\ 4 & 1 & -1 & 16 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -1 & 7 \\ 0 & -7 & 3 & -11 \\ 0 & -7 & 3 & -12 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -1 & 7 \\ 0 & -7 & 3 & -11 \\ 0 & 0 & 0 & -1 \end{array} \right).$$

Итак, ранг исходной матрицы равен двум, а ранг расширенной — трем, значит, уравнение, соответствующее третьей строке последней матрицы, противоречиво — оно свелось к неверному равенству $0 \cdot x_3 = -1$. Следовательно, данная система несовместна. ►

5.5. Разложение вектора по базису

Всякий вектор a , расположенный на числовой оси, определяется координатой $a = (x)$, т.е. одним действительным числом. Прямая — одномерное векторное пространство.

Всякий вектор, расположенный на плоскости с заданной системой координат, определяется двумя координатами $a(x, y)$, т.е. упорядоченной системой двух действительных чисел. Плоскость — двумерное векторное пространство и т.д.

ОПРЕДЕЛЕНИЕ. Упорядоченная система из n действительных чисел $a_1, a_2, \dots, a_n$ называется **n -мерным вектором** и обозначается так: $a = (a_1, a_2, \dots, a_n)$. Числа a_j ($j = 1, 2, \dots, n$) называются *компонентами вектора A* .

Если внимательно посмотреть на систему (1.25), то можно заметить, что решение системы линейных алгебраических уравне-

ний $\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$ есть n -мерный вектор. Сама матрица коэффициентов

$A_{m \times n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$ состоит из m строк-векторов раз-

мерности n .

В подпараграфе 1.1 мы говорили о базисе, о разложении вектора размерности n по базису в n -мерном пространстве. Ввели по-

нятие линейно зависимой и линейно независимой системы векторов. Поскольку матрица $A_{m \times n}$ представляет собой не что иное, как систему векторов, то с помощью элементарных преобразований методом Жордана-Гаусса легко можно выяснить, является эта система линейно независимой или нет.

Понятие линейной независимости тесно связано с понятием ранга (см. § 4 главы 1), что, в свою очередь, влияет на количество решений системы. Мы выяснили, что ранг матрицы равен максимальному числу ее линейно независимых строк. Если ранг расширенной матрицы системы меньше числа уравнений ($r < m$), то уравнения системы зависимые, если же число уравнений совпадает с рангом ($r = m$), то уравнения системы независимые, и наоборот.

ПРИМЕР 1.23. С помощью элементарных преобразований определить, является ли линейно зависимой система векторов $a_1 = (3, 2, 4, 3)$, $a_2 = (2, 3, 5, 4)$, $a_3 = (5, 5, 9, 7)$.

Решение. Составим из векторов-строк матрицу A :

$$A = \begin{pmatrix} 3 & 2 & 4 & 3 \\ 2 & 3 & 5 & 4 \\ 5 & 5 & 9 & 7 \end{pmatrix} \begin{matrix} I_1 \\ I_2 \\ I_3 \end{matrix} \sim \begin{pmatrix} 3 & 2 & 4 & 3 \\ 6 & 9 & 15 & 12 \\ 15 & 15 & 27 & 21 \end{pmatrix} \sim \begin{pmatrix} 3 & 2 & 4 & 3 \\ 0 & 5 & 7 & 6 \\ 0 & 5 & 7 & 6 \end{pmatrix}$$

С помощью элементарных преобразований строк получаем:

$$\begin{aligned} &\sim \begin{pmatrix} 15 & 0 & 6 & 3 \\ 0 & 5 & 7 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 15 & 0 & 0 & 0 \\ 0 & 5 & 105 & 90 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 15 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \\ &\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Мы видим, что ранг системы $r = 2$, система линейно зависима (так как $r < 3$) и только 2 вектора a_1 и a_2 составляют линейно независимую подсистему векторов.

ОПРЕДЕЛЕНИЕ. Линейное пространство R называется ***n*-мерным**, если в нем существует “ n ” линейно независимых векторов, а любые из $(n + 1)$ векторов уже являются зависимыми.

Следовательно, размерность пространства — это максимальное число содержащихся в нем линейно независимых векторов. Здесь n — размерность пространства и обозначается $n = \dim(R)$.

ОПРЕДЕЛЕНИЕ. Два вектора называются ***ортогональными***, если их скалярное произведение равно нулю.

Очевидно, что нулевой вектор ортогонален любому другому вектору. Из определения следует, что если два ненулевых вектора

ортогональны, то угол между ними равен $\frac{\pi}{2}$.

ОПРЕДЕЛЕНИЕ. Векторы $e_1, e_2, \dots, e_n$ n -мерного пространства образуют ***ортонормированный базис***, если эти векторы попарно ортогональны и норма каждого из них равна единице, т.е. если $(e_i, e_j) = 0$ при $i \neq j$ и $|e_i| = 1$ при $i = 1, 2, \dots, n$.

ПРИМЕР 1.24. Заданы векторы $a = (1, 0, 2)$, $b = (-2, 1, 1)$, $c = (2, 5, -1)$, $d = (1, -1, 2)$. Вычислить: 1) $a - 2b + 3c = f$, 2) скалярное произведение (a, d) , 3) $\cos \varphi$ — угол между векторами b и d , 4) выяснить, образуют ли векторы a, b, c базис в пространстве R^3 : если да, то а) является ли этот базис ортогональным, б) разложить вектор d по этому базису.

Р е ш е н и е.

$$1) f = a - 2b + 3c = (11, 13, 3).$$

Из определения скалярного произведения векторов (см. подпараграф 1.1) найдем:

$$2) (a, d) = 5; \text{ откуда}$$

$$3) \cos \varphi = -\frac{1}{6},$$

4) составим систему из трех векторов в виде матрицы, найдем ее ранг:

$$\begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & 1 \\ 2 & 5 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ 0 & 5 & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & -30 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$$

Ранг равен числу ступенек и равен трем ($r = 3$), значит, векторы линейно независимы и образуют базис.

А. Определим, ортогональный он или нет. Вычислим скалярные произведения: $(a, b) = 0$, $(b, c) = 0$, $(a, c) = 0$. Значит, базис ортогональный.

Б. Чтобы разложить вектор d по этому базису, нужно решить систему алгебраических уравнений, которую в краткой форме можем записать:

$$d = \lambda_1 a + \lambda_2 b + \lambda_3 c.$$

Составим расширенную матрицу системы из компонент векторов:

$$A_1 = \left(\begin{array}{ccc|c} 1 & -2 & 2 & 1 \\ 0 & 1 & 5 & -1 \\ 2 & 1 & -1 & 2 \end{array} \right), \text{ методом Гаусса находим решения } \lambda_1, \lambda_2,$$

λ_3 :

$$\left(\begin{array}{ccc|c} 1 & -2 & 2 & 1 \\ 0 & 1 & 5 & -1 \\ 0 & 5 & -5 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & -2 & 2 & 1 \\ 0 & 1 & 5 & -1 \\ 0 & 0 & -30 & 5 \end{array} \right).$$

Обратным ходом метода Гаусса находим: $-30 \lambda_3 = 5$, $\lambda_3 = -\frac{1}{6}$,

$$\lambda_2 - \frac{5}{6} = -1, \lambda_2 = -\frac{1}{6}, \lambda_1 = 1 + 2\left(-\frac{1}{6}\right) - 2\left(-\frac{1}{6}\right) = 1 - \frac{1}{3} + \frac{1}{3} = 1.$$

Итак, вектор d можно разложить по указанному базису в следующем виде:

$$d = a - \frac{1}{6}b - \frac{1}{6}c.$$

5.6. Приведенная система. Фундаментальная система решений

ОПРЕДЕЛЕНИЕ. Система m линейных уравнений с n переменными называется *однородной*, если во всех ее уравнениях свободные члены равны нулю.

В общем случае однородная система (или система однородных уравнений) имеет вид:

[illegible]

Система линейных однородных уравнений всегда совместна, так как она всегда имеет, по крайней мере, нулевое (тривиальное) решение $(0; 0; \dots; 0)$. Действительно, набор значений неизвестных $x_i = 0$ ($i = 1, 2, \dots, n$) удовлетворяет всем уравнениям системы.

ОПРЕДЕЛЕНИЕ. По отношению к системе (1.25) система (1.34) называется *приведенной*.

Если в системе (1.34) $m = n$ и $\Delta \neq 0$, то она имеет только одно нулевое решение (см. теорему 1.7).

□ ТЕОРЕМА 1.11. Система линейных однородных уравнений имеет ненулевое решение тогда и только тогда, когда ранг этой системы меньше числа ее неизвестных, т.е. при $\text{r}(A) < n$. ■

С л е д с т в и е 1. Если число уравнений однородной системы меньше числа ее неизвестных, то эта система имеет *ненулевое решение*.

С л е д с т в и е 2. Если в однородной системе число уравнений равно числу неизвестных, то она имеет *ненулевое решение* тогда и только тогда, когда определитель матрицы системы равен нулю.

Обозначим решение системы (1.34) $x_1 = k_1, x_2 = k_2, \dots, x_n = k_n$ в виде строки $I_1 = (k_1, k_2, \dots, k_n)$.

Решения системы линейных однородных уравнений обладают следующими **свойствами**:

1. Если строка $I_1 = (k_1, k_2, \dots, k_n)$ — решение системы (1.34), то и строка $\lambda I_1 = (\lambda k_1, \lambda k_2, \dots, \lambda k_n)$ — также решение этой системы.

2. Если строки $I_1 = (k_1, k_2, \dots, k_n)$ и $I_2 = (l_1, l_2, \dots, l_n)$ — решения системы (1.34), то при любых c_1 и c_2 их линейная комбинация $c_1 I_1 + c_2 I_2 = (c_1 k_1 + c_2 l_1, c_1 k_2 + c_2 l_2, \dots, c_1 k_n + c_2 l_n)$ — также решение данной системы.

Убедиться в справедливости указанных свойств решений системы линейных однородных уравнений можно непосредственной подстановкой их в уравнения системы.

Из сформулированных свойств следует, что *всякая линейная комбинация решений системы линейных однородных уравнений также является решением этой системы*. Поэтому целесообразно найти такие линейно независимые решения системы (1.34), через которые линейно выражались бы все остальные ее решения.

ОПРЕДЕЛЕНИЕ. Система линейно независимых решений $I_1, I_2, \dots, I_k$ называется **фундаментальной**, если каждое решение системы (1.34) является линейной комбинацией решений $I_1, I_2, \dots, I_k$.

□ **ТЕОРЕМА 1.12.** Если ранг r матрицы однородной системы линейных уравнений (1.34) меньше числа неизвестных n , то всякая ее фундаментальная система решений состоит из $n - r$ решений (или матрица фундаментальной системы имеет $n - r$ столбцов). ■

Поэтому *общее решение* системы (1.34) линейных однородных уравнений имеет вид:

$$c_1 I_1 + c_2 I_2 + \dots + c_k I_k, \quad (1.35)$$

где $I_1, I_2, \dots, I_k$ — любая фундаментальная система решений; $c_1, c_2, \dots, c_k$ — произвольные числа и $k = n - r$.

З а м е ч а н и е. Общее решение системы m линейных уравнений с n неизвестными (1.25) равно сумме общего решения соответствующей ей приведенной системы линейных уравнений (1.34) и произвольного частного решения этой системы (1.25).

Для нахождения фундаментальной системы решений предположим, что ранг $r < n$. Тогда базисные неизвестные этой системы $x_1, x_2, \dots, x_r$ линейно выражаются через свободные переменные $x_{r+1}, x_{r+2}, \dots, x_n$. Положим значения свободных переменных $x_{r+1} = 1, x_{r+2} = x_{r+3} = \dots = x_n = 0$. Затем находим второе решение, принимая $x_{r+2} = 1, x_{r+1} = x_{r+3} = x_{r+4} = \dots = x_n = 0$. Иными словами, мы последовательно присваиваем каждой свободной переменной единичное значение, положив остальные нулями.

▷ **ПРИМЕР 1.25.** Найти решение и фундаментальную систему решения системы линейных однородных уравнений:

$$\begin{cases} x_1 + 3x_2 + 4x_3 + x_5 = 0, \\ x_1 + 2x_2 + 3x_3 + x_4 = 0, \\ \quad \quad \quad x_2 + 8x_4 - x_5 = 0, \\ x_2 + x_3 + x_4 + 2x_5 = 0. \end{cases}$$

Р е ш е н и е. Составим матрицу системы, и прямым ходом метода Гаусса приведем ее к ступенчатому виду:

$$\begin{pmatrix} 1 & 3 & 4 & 0 & 1 \\ 1 & 2 & 3 & 1 & 0 \\ 0 & 1 & 0 & 8 & -1 \\ 0 & 1 & 1 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & 4 & 0 & 1 \\ 0 & -1 & -1 & 1 & -1 \\ 0 & 1 & 0 & 8 & -1 \\ 0 & 1 & 1 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} \underline{1} & 3 & 4 & 0 & 1 \\ 0 & \underline{-1} & -1 & 1 & -1 \\ 0 & 0 & \underline{-1} & 9 & -2 \\ 0 & 0 & 0 & \underline{2} & 1 \end{pmatrix}$$

Выпишем систему уравнений:

$$\begin{cases} x_1 + 3x_2 + 4x_3 + x_5 = 0, \\ -x_2 - x_3 + x_4 - x_5 = 0, \\ \quad -x_3 + 9x_4 - 2x_5 = 0, \\ \quad \quad \quad 2x_4 + x_5 = 0. \end{cases}$$

Обратный ход метода Гаусса дает значения базисных неизвестных x_1, x_2, x_3, x_4 , выраженные через свободную переменную x_5 . Обозначим ее $c = x_5$.

Из последнего уравнения находим $x_4 = -\frac{1}{2}x_5 = -\frac{1}{2}c$. Затем, поднимаясь вверх по системе, определяем все неизвестные $x_1 = 10c$; $x_2 = 5c$; $x_3 = -\frac{13}{2}c$; $x_4 = -\frac{1}{2}c$.

Эти последние выражения представляют запись общего решения нашей однородной системы. Если теперь давать переменной c числовые значения, можно получить фундаментальное решение системы.

Поскольку ранг однородной системы равен четырём, то фундаментальная система решений для нее состоит из $n - r = 5 - 4 = 1$ решения.

Положив значение свободной переменной $x_5 = 1$ (других свободных переменных у нас нет), получим фундаментальное решение системы:

$$x_1 = 10, x_2 = 5, x_3 = -\frac{13}{2}, x_4 = -\frac{1}{2}.$$

Заметим, что если $x_5 = c = 0$, то $x_1 = x_2 = x_3 = x_4 = 0$, и решением будет нулевой вектор o ; его называют тривиальным решением; этот вектор всегда есть среди решений однородной системы. ►

5.7. Характеристическое уравнение

В подпараграфе 1.4 были введены определения собственного значения и собственного вектора матрицы A порядка n . Тогда имеет место матричное уравнение

$$Ax = \lambda x,$$

или

$$(A - \lambda E)x = o, \tag{1.36}$$

где λ — собственное значение матрицы A , а E и o — единичная матрица и нулевой вектор-столбец соответственно. Уравнение (1.36) эквивалентно системе однородных уравнений

$$(1.37)$$

В уравнениях (1.37) a_{ij} — элементы матрицы A , x_j — координаты собственного вектора x . Поскольку собственный вектор не является нулевым, то однородная система (1.37) должна иметь ненулевое решение, т.е. в силу следствия 2 теоремы 1.11 определитель этой системы равен нулю:

$$(1.38)$$

ОПРЕДЕЛЕНИЕ. Определитель системы однородных уравнений (1.37) называется *характеристическим многочленом*, а уравнение (1.38) — *характеристическим уравнением* матрицы A .

Уравнение (1.38) имеет степень n относительно неизвестной λ . Его корни являются собственными числами матрицы A . Определив набор этих чисел, для каждого из них можно найти соответствующий собственный вектор как решение однородной системы (1.37).

▷ П Р И М Е Р 1.26. Найти собственные числа и собственные векторы матрицы A :

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}.$$

Р е ш е н и е. Характеристическое уравнение для этой матрицы имеет вид

$$\begin{vmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{vmatrix} = 0,$$

откуда, раскрывая определитель, получаем $\lambda^2 - 7\lambda + 10 = 0$. Корни этого уравнения $\lambda_1 = 2$, $\lambda_2 = 5$. Для нахождения собственных векторов подставим найденные собственные значения в систему однородных уравнений (1.37) при $n = 2$, соответствующую заданной матрице A . Собственный вектор, соответствующий собственному значению $\lambda_1 = 2$, является решением системы

$$\begin{cases} x_1 + 2x_2 = 0, \\ x_1 + 2x_2 = 0. \end{cases}$$

По сути дела, это одно уравнение. Полагая $x_2 = b$ свободной переменной, получаем первый собственный вектор $x_1 = (-2b; b) = b(-2; 1)$. Подстановка второго собственного значения $\lambda_2 = 5$ приводит к системе уравнений

$$\begin{cases} -2x_1 + 2x_2 = 0, \\ x_1 - x_2 = 0, \end{cases}$$

которая через свободную переменную $x_2 = c$ определяет второй собственный вектор матрицы A : $x_2 = (c; c) = c(1; 1)$.

Поскольку b и c — произвольные числа, то одному собственному значению может соответствовать несколько собственных векторов разной длины. Например, собственные векторы, соответствующие фундаментальным решениям однородных систем (в данном случае их будет по одному на каждое собственное значение), имеют вид: $x_1 = (-2, 1)$; $x_2 = (1, 1)$. ►

§6. Модель Леонтьева межотраслевого баланса

Заканчивая изложение матричной алгебры и методов решения систем линейных уравнений, подчеркнем, что именно эти методы и их различные модификации являются самыми распространенными и “экономными” методами решения систем линейных уравнений (см. [18]).

В математическом обеспечении каждого вычислительного центра обязательно есть программы для решения систем. Эти стандартные программы реализуют метод Гаусса; именно метод Гаусса (или Жордана-Гаусса) требует минимальных затрат машинного времени. Мы изложили метод Гаусса применительно к ручному счету. Чтобы уменьшить ошибки округления при решении систем линейных уравнений высокого порядка, на компьютере используют метод Гаусса с выбором “ведущего” элемента. Как правило, таким элементом является максимальный по модулю коэффициент в матрице системы.

Мы уже упоминали об использовании линейной алгебры, в частности методов исследования систем линейных уравнений, в экономике. Одной из задач при математическом моделировании экономических процессов является задача создания модели межотраслевого баланса. Модель эта называется моделью Леонтьева (по имени ее создателя) и активно используется для управления народным хозяйством.

Пусть в национальной экономике взаимодействует n отраслей, каждая из которых производит свою продукцию. Часть продукции идет на внутрипроизводственное потребление данной отраслью и другими отраслями, а другая часть предназначена для целей конечного (вне сферы материального производства) личного и общественного потребления.

Рассмотрим процесс производства за некоторый период времени (например, за год).

Введем следующие обозначения:

x_i — общий (валовой) объем продукции i -й отрасли ($i = 1, 2, \dots, n$);

x_{ij} — объем продукции i -й отрасли, потребляемой j -й отраслью в процессе производства ($i, j = 1, 2, \dots, n$);

y_i — объем конечного продукта i -й отрасли для непроеизводственного потребления.

Так как валовой объем продукции любой i -й отрасли равен суммарному объему продукции, потребляемой n отраслями и конечного продукта, то

$$x_i = \sum_{j=1}^n x_{ij} + y_i, \quad (i=1, 2, \dots, n). \quad (1.39)$$

Уравнения (1.39) называются *соотношениями баланса*. Будем рассматривать *стоимостный межотраслевой баланс*, когда все величины, входящие в (1.39), имеют стоимостное выражение.

Введем *коэффициенты прямых затрат*

$$a_{ij} = \frac{x_{ij}}{x_j}, \quad (i, j=1, 2, \dots, n), \quad (1.40)$$

показывающие затраты продукции i -й отрасли на производство единицы продукции j -й отрасли.

Коэффициенты a_{ij} зависят от сложившейся технологии производства. Можно полагать, что в некотором промежутке времени они будут постоянными. Это означает *линейную* зависимость материальных затрат от валового выпуска, т.е.

$$x_{ij} = a_{ij} x_j, \quad (i, j=1, 2, \dots, n), \quad (1.41)$$

вследствие чего построенная на этом основании модель межотраслевого баланса получила название *линейной*. Теперь соотношения баланса (1.39) примут вид

$$x_i = \sum_{j=1}^n a_{ij} x_j + y_i, \quad (i=1, 2, \dots, n). \quad (1.42)$$

Обозначим

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix},$$

где X — вектор валового выпуска; A — матрица прямых затрат (технологическая, или структурная, матрица); Y — вектор конечного продукта.

Тогда систему (1.42) можно записать в матричном виде:

$$X = AX + Y. \quad (1.43)$$

Основной задачей межотраслевого баланса является проблема отыскания такого вектора валового выпуска X , который при известной матрице прямых затрат A обеспечивает заданный вектор конечного продукта Y .

Если переписать уравнение (1.43) в виде

$$(E - A)X = Y, \quad (1.44)$$

где E — единичная матрица порядка n , то при условии невырожденности матрицы $E - A$ можно найти решение, помножив как левую, так и правую части равенства на $(E - A)^{-1}$. Получим

$$X = (E - A)^{-1}Y. \quad (1.45)$$

Таким образом, решение задачи межотраслевого баланса свелось к решению системы линейных алгебраических уравнений.

Подробно решение этой задачи, как и некоторых других, сводящихся к матричному исследованию, вы найдете в работе [18], где проблемам математического моделирования экономических процессов уделено центральное место.

На этом мы заканчиваем изучение основ линейной алгебры, хотя, следует отметить, что в рамках нашей главы мы рассмотрели не все вопросы курса линейной алгебры, лишь самые важные из них, для того чтобы строить линейные экономические модели. В качестве дополнения к курсу вам предлагаются работы [1], [9], [10], [14], [16].

Контрольные вопросы

1. Как определяются и выполняются элементарные операции над матрицами: умножение матрицы на число, сложение матриц и их произведение?

2. Как вычисляются определители первого, второго, третьего, n -го порядков?
3. Какими способами можно вычислить определитель шестого порядка, какому вы отдадите предпочтение и почему?
4. Каким образом считается обратная матрица?
5. Чем различаются и чем схожи процедуры подсчета определителя и вычисления ранга матрицы?
6. Почему не всякую систему линейных уравнений можно решить по формулам Крамера или методом обратной матрицы?
7. Какие системы уравнений являются неопределенными и как их решать?
8. Как, по вашему мнению, связана линейная независимость (или зависимость) уравнений системы с поиском ее решений?
9. Что называют базисным, фундаментальным решением системы?

Упражнения

1. Предприятие производит продукцию трех видов и использует сырье двух типов. Нормы затрат сырья на единицу продукции каждого вида заданы матрицей $A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 4 \end{pmatrix}$. Стоимость единицы сырья каждого типа задана матрицей $B = \begin{pmatrix} 10 & 15 \end{pmatrix}$. Каковы общие затраты предприятия на производство 100 единиц продукции первого вида, 200 единиц продукции второго вида и 150 единиц продукции третьего вида?

2. Вычислить матрицу $D = (AB) - C^2$, где

$$A = \begin{pmatrix} 3 & 4 & 2 \\ 1 & 0 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 1 & 3 \\ 0 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 \\ 0 & 4 \end{pmatrix}.$$

3. Найти произведение матриц ABC , где

$$A = \begin{pmatrix} 4 & 3 \\ 7 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} -28 & 93 \\ 38 & -126 \end{pmatrix}, \quad C = \begin{pmatrix} 7 & 3 \\ 2 & 1 \end{pmatrix}.$$

4. Вычислить матрицу $D = ABC - 3E$, где

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 1 & 0 & 3 \\ 4 & 5 & 3 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, C = (2 \ 0 \ 5), E \text{ — единичная матрица}$$

соответствующего порядка.

5. Вычислить $A^3, A^n, n = 0, 1, 2, \dots$, если $A = \begin{pmatrix} 1 & 1 & -1 \\ 3 & -1 & 2 \\ 2 & -1 & 0 \end{pmatrix}$.

Вычислить определитель:

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

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

8. $\begin{vmatrix} 0 & 6 & 3 & 5 & 1 \\ -3 & 2 & 4 & 1 & 0 \\ 5 & 1 & 4 & 3 & 2 \\ -3 & 8 & 7 & 6 & 1 \\ 1 & 0 & 3 & 4 & 0 \end{vmatrix}$.

9. Определить, имеет ли матрица A обратную, и, если имеет, вычислить ее:

$$A = \begin{pmatrix} 4 & -8 & -5 \\ -4 & 7 & -1 \\ -3 & 5 & 1 \end{pmatrix}$$

10. Вычислить матрицу $B = 11 \cdot (A^{-1})' + A'$, где $A = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 1 & 0 & 4 \end{pmatrix}$

11. Найти значения λ , при которых матрица A не имеет обратной:

$$A = \begin{pmatrix} \lambda & 4 & 1 \\ 2 & 5 & -1 \\ 0 & \lambda & 1 \end{pmatrix}$$

Найти ранги матриц:

12. $\begin{pmatrix} 2 & 5 & 6 \\ 4 & -1 & 5 \\ 2 & -6 & -1 \end{pmatrix}$

13. $\begin{pmatrix} 1 & 3 & 7 & 2 & 5 \\ -1 & 0 & 4 & 8 & 3 \\ 3 & 6 & 10 & -4 & 7 \end{pmatrix}$

14. $\begin{pmatrix} 1 & 2 & 1 & 4 \\ 0 & 5 & -1 & 4 \\ -1 & 3 & 4 & 6 \end{pmatrix}$

15. $\begin{pmatrix} 0 & 5 & -1 & 1 & 5 \\ 2 & 3 & 0 & 1 & 6 \\ -1 & -3 & 1 & 3 & 0 \\ 3 & -1 & 0 & 4 & 6 \end{pmatrix}$

16. Определить максимальное число линейно независимых строк матрицы

$$A = \begin{pmatrix} 2 & 0 & 3 & 5 & 1 \\ 4 & 3 & 1 & 7 & 5 \\ 0 & 3 & -5 & -3 & 3 \\ 2 & 3 & -2 & 2 & 4 \end{pmatrix}$$

Ответы к упражнениям

1. 28 000. 2. $D = \begin{pmatrix} 9 & -13 \\ 22 & 9 \end{pmatrix}$ 3. $ABC = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$
4. $D \equiv \begin{pmatrix} 1 & 0 & 10 \\ 6 & -3 & 15 \\ 34 & 0 & 82 \end{pmatrix}$ 5. $A^3 = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}$, $A^{3k} = 7^k$, $A^{3k+1} =$
 $= 7^k \cdot A$; $A^{3k+2} = 7^k \cdot A^2$, $k = 0, 1, 2, \dots$
6. 40. 7. 160. 8. 0. 9. $\frac{1}{13} \cdot \begin{pmatrix} -12 & 17 & -43 \\ -7 & 11 & -24 \\ -1 & -4 & 4 \end{pmatrix}$
10. $B = \begin{pmatrix} 5 & 2 & 0 \\ -6 & 8 & 2 \\ 4 & 0 & 5 \end{pmatrix}$ 11. $\{-8, 1\}$. 12. 2. 13. 2. 14. 3. 15. 4. 16. 2.
-
-

Г л а в а 2

ОСНОВЫ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

§ 1. Прямые линии и плоскости

1.1. Основные понятия об уравнениях

В качестве экономического примера рассмотрим вначале, как происходит начисление процентов.

Обладатель денег может распорядиться ими простым способом — поместить их на банковский счет ровно PV (сокращение от “present value”, дословно — современная стоимость денег, т. е. сумма, которой владелец обладает сегодня). Тогда через год у владельца этих денег на счете будет сумма

$$FV = PV(1+i),$$

где i — процентная ставка банка; FV — “future value”, дословно — будущая стоимость денег, т. е. сумма, которую получит владелец спустя определенное время.

К примеру, если банковская процентная ставка равна 20%, то через год сумма на счете вырастет в 1,2 раза. Доход обладателя счета составит $i PV = 0,2 PV$.

Существуют два способа начисления процентов: по *простой процентной ставке* и по *сложной*.

При начислении дохода по *простой процентной ставке* i_{np} доход каждый раз начисляется на первоначально вложенную сумму. То есть, через год доход составит $i_{np} PV$, через два года $2i_{np} PV$, через 5 лет — $5i_{np} PV$, через n лет — $ni_{np} PV$. Таким образом, при начислении дохода по простой процентной ставке через n лет на счете у владельца будет сумма

$$FV = PV(1+i_{np}n). \quad (2.1)$$

Это соотношение описывает линейную зависимость будущей стоимости денег FV от времени n . Таким образом, при начислении дохода по простой процентной ставке деньги на счете растут по линейному закону, и наше равенство (2.1) представляет собой уравнение, связывающее будущую стоимость денег со временем.

Пусть в пространстве задана декартова прямоугольная система координат. Рассмотрим сферу радиуса r , центр которой находится в точке P с координатами (a, b, c) . Сфера — множество всех точек, отстоящих от центра на одно и то же расстояние r . Обозначим через (x, y, z) координаты некоторой точки M и выразим через них равенство $|PM| = r$:

$$\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} = r.$$

Возводя в квадрат обе части равенства, получим более удобную форму

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2. \quad (2.2)$$

Очевидно, что это равенство выполнено для всех точек сферы, и только для них, следовательно, его можно рассматривать как запись определения сферы при помощи координат. Равенство (2.2) называется *уравнением сферы* в рассматриваемой системе координат.

Если взять пример из геометрии на плоскости, то *графиком* функции f называется линия L , состоящая из точек, координаты которых связаны соотношением $y = f(x)$. Так как нас сейчас в первую очередь интересует линия, а не функция, то мы можем встать

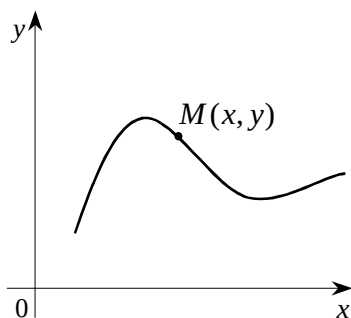


Рис. 2.1

на другую точку зрения и считать, что соотношение $y = f(x)$, или $F(x, y) = 0$, есть уравнение линии L .

Уравнение линии является важнейшим понятием аналитической геометрии.

Пусть мы имеем на плоскости некоторую линию (кривую) (рис. 2.1). Координаты x и y точки, лежащей на этой линии, не могут быть произвольными, они должны быть свя-

заны. Такая связь аналитически записывается в виде некоторого уравнения.

ОПРЕДЕЛЕНИЕ. Уравнение вида $F(x, y) = 0$ или $y = f(x)$, которому удовлетворяют координаты x и y каждой точки данной линии и не удовлетворяют координаты любой точки, не лежащей на этой линии, называется **уравнением линии (кривой) на плоскости** Oxy .

Если точка $M(x, y)$ передвигается по линии, то ее координаты, изменяясь, удовлетворяют уравнению этой линии. Поэтому координаты $M(x, y)$ называются *текущими* координатами (от слова “текут”, меняются).

▷ **ПРИМЕР 2.1.** Найти уравнение множества точек, равноудаленных от точек $A(-4; 2)$ и $B(-2; -6)$.

Решение. Расстояние между двумя точками $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ определяется по формуле

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Если $M(x, y)$ — произвольная точка искомой линии, то, согласно условию, имеем $AM = BM$ (рис. 2.2) или

$$\sqrt{(x + 4)^2 + (y - 2)^2} = \sqrt{(x + 2)^2 + (y + 6)^2}.$$

Возведя обе части уравнения в квадрат, получим после преобразования уравнение $x - 4y - 5 = 0$, или $y = \frac{1}{4}x - \frac{5}{4}$.

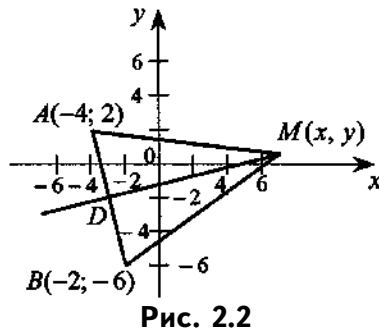


Рис. 2.2

Очевидно, это уравнение прямой MD — перпендикуляра, проведенного из середины отрезка AB (см. рис. 2.2). ►

Любую линию можно выразить соответствующим уравнением (хотя на практике это не всегда просто сделать). Однако не всякое уравнение определяет на плоскости некоторую линию.

Например, уравнение $x^2 + y^2 = 0$ определяет только одну точку $(0; 0)$, а уравнение $x^2 + y^2 + 14 = 0$ не определяет никакого множества точек, ибо левая часть уравнения не может равняться нулю.

Чаще всего уравнение представляет собой равенство, записанное математическими символами, но это вовсе не обязательно: оно может быть словесным описанием, перечислением и т. д. Например, высказывание “обе координаты точки — целые числа” мы будем считать уравнением соответствующего множества в какой-либо заранее выбранной системе координат.

Чтобы убедиться, лежит ли точка $M(a, b)$ на данной линии $F(x, y) = 0$, надо проверить, удовлетворяют ли координаты этой точки уравнению $F(x, y) = 0$.

Часто уравнению множества точек в планиметрии можно придать вид $F(x, y) = 0$, а в стереометрии — вид $F(x, y, z) = 0$, где F — функция соответственно двух или трех переменных. Уравнение сферы (2.2) имеет такой вид, если не замечать то обстоятельство, что член r^2 написан в другой части равенства.

З а м е ч а н и е. При изменении системы координат меняются координаты точки, а потому уравнения одного и того же множества в разных системах координат, вообще говоря, различны.

1.2. Уравнение прямой на плоскости

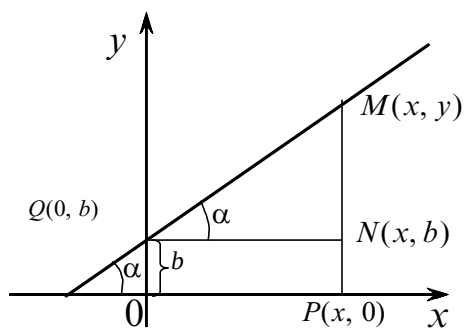


Рис. 2.3

Уравнение с двумя переменными вида

$$Ax + By + C = 0$$

описывает на координатной плоскости Oxy прямую.

Пусть прямая пересекает ось Oy в точке $Q(0; b)$ и образует с осью Ox угол

$$\alpha \left(0 < \alpha < \frac{\pi}{2} \right) \text{ (рис. 2.3).}$$

Возьмем на прямой произвольную точку $M(x, y)$. Тогда тангенс угла α наклона прямой найдем из прямоугольного треугольника MQN :

$$\operatorname{tg} \alpha = \frac{MN}{NQ} = \frac{y - b}{x}.$$

Введем угловой коэффициент прямой $k = \operatorname{tg} \alpha$, получим

$$k = \frac{y - b}{x},$$

или

$$y = kx + b. \quad (2.3)$$

Можно показать, что формула (2.3) справедлива и для случая

$$\frac{\pi}{2} < \alpha < \pi.$$

Итак, мы показали, что координаты каждой точки прямой удовлетворяют уравнению (2.3). Нетрудно доказать, что координаты любой точки, не лежащей на прямой, не удовлетворяют уравнению (2.3).

Уравнение (2.3) называется *уравнением прямой с угловым коэффициентом*.

Рассмотрим частные случаи уравнения (2.3).

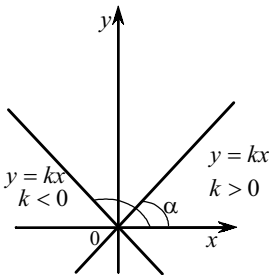


Рис. 2.4

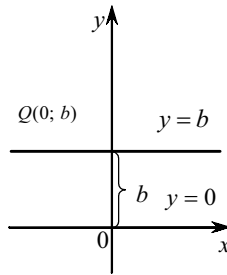


Рис. 2.5

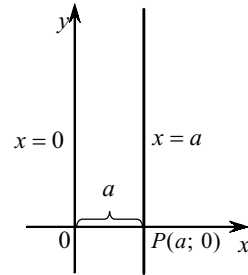


Рис. 2.6

1. Если $b = 0$, то получаем $y = kx$ — уравнение прямой, проходящей через начало координат и образующей при $k = \operatorname{tg} \alpha > 0$ острый угол α с осью Ox , а при $k = \operatorname{tg} \alpha < 0$ тупой угол (рис. 2.4). В частности, уравнение биссектрисы I и III координатных углов

имеет вид $y = x$ (так как $k = \operatorname{tg} \frac{\pi}{4} = 1$), а уравнение биссектрисы II и

IV координатных углов — вид $y = -x$ ($k = \operatorname{tg} \frac{\pi}{4} = -1$)

2. Если $\alpha = 0$, то $k = \operatorname{tg} 0 = 0$, и уравнение прямой, параллельной оси Ox , имеет вид $y = b$, а самой оси Ox — вид $y = 0$ (рис. 2.5).

3. Если $\alpha = \frac{\pi}{2}$, то прямая перпендикулярна оси Ox (рис. 2.6) и

$k = \operatorname{tg} \frac{\pi}{2}$ не существует, т.е. вертикальная прямая не имеет углово-

го коэффициента. Предположим, что эта прямая отсекает на оси Ox отрезок, равный a . Очевидно, что уравнение такой прямой $x = a$ (так как абсцисса любой точки прямой равна a), а уравнение оси Oy есть $x = 0$.

Уравнение прямой, проходящей через данную точку в данном направлении. Пусть прямая проходит через точку $M_1(x_1, y_1)$ и об-

разует с осью Ox угол $\alpha \neq \frac{\pi}{2}$ (рис. 2.7).

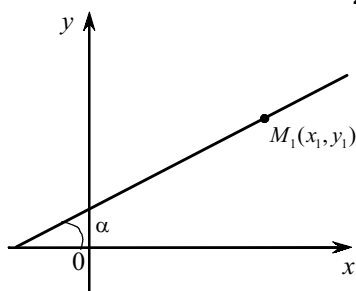


Рис. 2.7

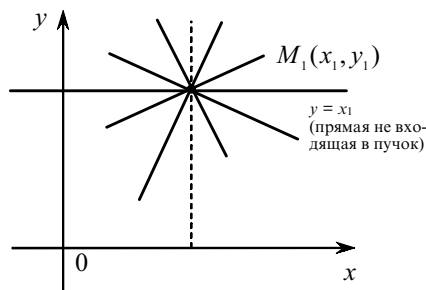


Рис. 2.8

Так как точка $M_1(x_1, y_1)$ лежит на прямой, то ее координаты удовлетворяют уравнению (2.3), т.е.

$$y_1 = kx_1 + b. \quad (2.4)$$

Вычитая равенство (2.4) из равенства (2.3), получим уравнение искомой прямой

$$y - y_1 = k(x - x_1). \quad (2.5)$$

Уравнение пучка прямых. Если в уравнении (2.5) коэффициент k — произвольное число, то это уравнение определяет пучок прямых, проходящих через точку $M_1(x_1, y_1)$, кроме прямой, параллельной оси Oy и не имеющей углового коэффициента (рис. 2.8).

▷ П Р И М Е Р 2.2. Составить уравнение прямой, проходящей через точку $P(3; -2)$: а) под углом 135° к оси Ox ; б) параллельно оси Oy .

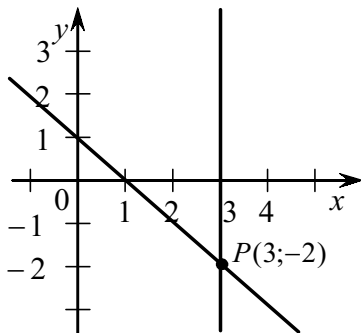


Рис. 2.9

Найти уравнение пучка прямых, проходящих через точку P .

Р е ш е н и е. а) Угловой коэффициент прямой $k = \operatorname{tg} 135^\circ = -1$.

Уравнение прямой, проходящей через точку $P(3; -2)$ (рис. 2.9), по формуле (2.5) имеет вид $y + 2 = -1(x - 3)$, или $y = -x + 1$.

б) Уравнение прямой, параллельной оси Oy , $x = 3$.

Уравнение пучка прямых, проходящих через точку $P(3; -2)$, имеет вид $y + 2 = k(x - 3)$.▶

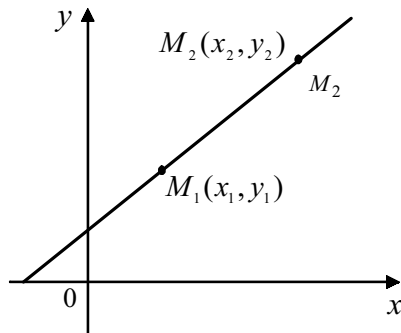


Рис. 2.10

Уравнение прямой, проходящей через две данные точки. Пусть даны две точки $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ и $x_1 \neq x_2$, $y_1 \neq y_2$.

Для составления уравнения прямой M_1M_2 (рис.2.10) запишем уравнение пучка прямых, проходящих через точку M_1 :

$$y - y_1 = k(x - x_1).$$

Чтобы выделить точку $M_2(x_2, y_2)$ из пучка прямых, подставим координаты точки M_2 (так как точка лежит на искомой прямой) в уравнение пучка $y_2 - y_1 = k(x_2 - x_1)$ и найдем угловой коэффициент прямой

$$k = \frac{y_2 - y_1}{x_2 - x_1}. \quad (2.6)$$

Теперь уравнение прямой M_1M_2 примет вид

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1),$$

или

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \quad (2.7)$$

▷ П Р И М Е Р 2.3. Составить уравнение прямой, проходящей через точки $P(-3; 2)$ и $Q(6; -3)$.

Р е ш е н и е. По уравнению (2.7): $\frac{y - 2}{-3 - 2} = \frac{x + 3}{6 + 3}$, откуда после

преобразований $y = -\frac{5}{9}x + \frac{1}{3}$. ►

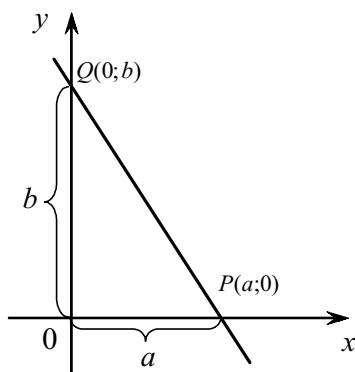


Рис. 2.11

Уравнение прямой в отрезках.

Найдем уравнение прямой по заданным отрезкам $a \neq 0$ и $b \neq 0$, отсекаемым на осях координат. Используя (2.7), получим уравнение прямой, проходящей через точки $P(a; 0)$ и $Q(0; b)$ (рис. 2.11):

$$\frac{y - 0}{b - 0} = \frac{x - a}{0 - a}$$

или после преобразований

$$\frac{x}{a} + \frac{y}{b} = 1. \quad (2.8)$$

Уравнение (2.8) называется *уравнением прямой в отрезках*.

▷ **ПРИМЕР 2.4.** Составить уравнение прямой, проходящей через точку $P(2; -1)$, если эта прямая отсекает от положительной полуоси Oy отрезок, вдвое больший, чем на положительной полуоси Ox (рис. 2.12).

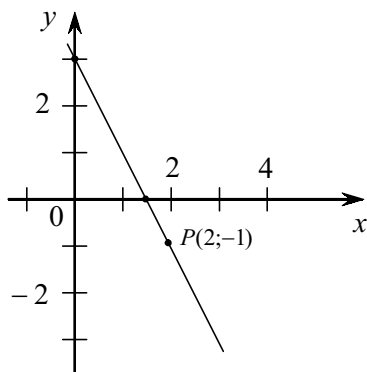


Рис. 2.12

Решение. По условию $b = 2a$ ($a > 0$, $b > 0$). Подставляя это выражение в уравнение (2.8), получим

$$\frac{x}{a} + \frac{y}{2a} = 1.$$

Так как точка $P(2; -1)$ лежит на прямой, то ее координаты удовлетворяют этому уравнению, т.е.

$$\frac{2}{a} - \frac{1}{2a} = 1, \text{ откуда } a = 1,5.$$

Итак, уравнение искомой прямой имеет вид $\frac{x}{1,5} + \frac{y}{3} = 1$, или $y = -2x + 3$. ►

Геометрическая интерпретация решения систем линейных уравнений. Мы уже упоминали ранее, что прямая на плоскости описывается уравнением первой степени с двумя переменными:

$$Ax + By + C = 0. \quad (2.9)$$

Решению систем из уравнений такого вида в главе 1 мы уделили особое внимание. Добавим только, что система из двух уравнений такого вида означает, что ее решение, как и точки на координатной плоскости, должны принадлежать одновременно двум прямым, соответствующим уравнениям этой системы.

Отсюда возможны следующие варианты: а) прямые пересекаются, и тогда система имеет единственное решение; б) прямые параллельны, и система не имеет решений (несовместна); в) прямые совпадают, т.е. ранг системы равен единице, и система имеет бесчисленное множество решений.

Итак, вернемся к уравнению (2.9), в котором коэффициенты A и B не равны одновременно нулю, т.е.

$$A^2 + B^2 \neq 0.$$

1. Пусть $B \neq 0$. Тогда уравнение (2.9) можно записать в виде

$$y = -\frac{A}{B}x - \frac{C}{B}.$$

Обозначим $k = -A/B$, $b = -C/B$. Если $A \neq 0$, $C \neq 0$, то получим $y = kx + b$ (уравнение прямой с угловым коэффициентом); если $A \neq 0$, $C = 0$, то $y = kx$ (уравнение прямой, проходящей через начало координат); если $A = 0$, $C \neq 0$, то $y = b$ (уравнение прямой, параллельной оси Oy); если $A = 0$, $C = 0$, то $y = 0$ (уравнение оси Ox).

2. Пусть $B = 0$, $A \neq 0$. Тогда уравнение (2.9) примет вид $x = -\frac{C}{A}$.

Обозначим $a = -C/A$. Если $C \neq 0$, то получим $x = a$ (уравнение прямой, параллельной оси Ox); если $C = 0$, то $x = 0$ (уравнение оси Oy).

Таким образом, при любых значениях коэффициентов A , B (не равных одновременно нулю) и C уравнение (2.9) есть уравнение некоторой прямой линии на плоскости Oxy . Можно показать, что уравнение любой прямой на плоскости можно получить из (2.9) при некоторых значениях A , B , C .

Поэтому уравнение (2.9) называется общим уравнением прямой. Заметим, что в отличие от уравнения пучка прямых (2.5) общее уравнение (2.9) включает и уравнение любой вертикальной прямой, параллельной оси Oy .

1.3. Взаимное расположение прямых на плоскости.

Расстояние от точки до прямой

Угол между двумя прямыми. Пусть заданы две прямые l_1 и l_2 :

$$y = k_1x + b_1,$$

$$y = k_2x + b_2,$$

и требуется определить угол φ между ними.

Из рис. 2.13 видно, что $\varphi = \alpha_2 - \alpha_1$, причем $k_1 = \operatorname{tg} \alpha_1$, $k_2 = \operatorname{tg} \alpha_2$,

$$\alpha_1 \neq \frac{\pi}{2}, \alpha_2 \neq \frac{\pi}{2}, \alpha_2 - \alpha_1 \neq \frac{\pi}{2}.$$

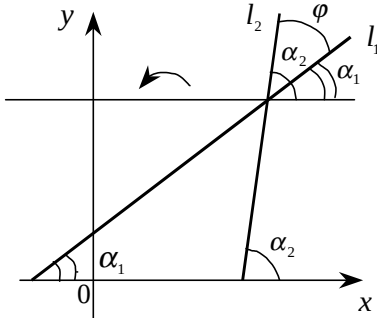


Рис. 2.13

Тогда

$$\operatorname{tg} \varphi = \operatorname{tg}(\alpha_2 - \alpha_1) = \frac{\operatorname{tg} \alpha_2 - \operatorname{tg} \alpha_1}{1 + \operatorname{tg} \alpha_1 \operatorname{tg} \alpha_2},$$

или

$$\operatorname{tg} \varphi = \frac{k_2 - k_1}{1 + k_1 k_2}, \quad (2.10)$$

где стрелка означает, что угол φ получается поворотом прямой l_1 к прямой l_2 против часовой стрелки.

Условия параллельности и перпендикулярности прямых. Если прямые $y = k_1 x + b_1$ (l_1) и $y = k_2 x + b_2$ (l_2) параллельны, то угол $\varphi = 0$, и по формуле (2.10) получаем $k_1 = k_2$. И наоборот, если $k_1 = k_2$, то по формуле (2.10) следует, что $\operatorname{tg} \varphi = 0$ и $\varphi = 0$. Таким образом, равенство угловых коэффициентов является необходимым и достаточным условием параллельности двух прямых.

Аналогично можно показать, что для перпендикулярности прямых необходимо и достаточно, чтобы их угловые коэффициенты были обратными по величине и противоположными по знаку, т.е.

$$k_1 = -\frac{1}{k_2}, \text{ или } k_1 k_2 = -1.$$

Если прямые заданы общими уравнениями $A_1 x + B_1 y + C_1 = 0$ (l_1) и $A_2 x + B_2 y + C_2 = 0$ (l_2), то, учитывая, что их угловые коэффициенты $k_1 = -A_1/B_1$ и $k_2 = -A_2/B_2$, условие параллельности прямых $k_1 = k_2$

примет вид $\frac{A_1}{A_2} = \frac{B_1}{B_2}$, $A_2 \neq 0$, $B_2 \neq 0$. Следовательно, условием параллельности прямых, заданных общими уравнениями, является пропорциональность коэффициентов при переменных.

Условие перпендикулярности прямых $k_1 k_2 = -1$ в этом случае примет вид

$$\left(-\frac{A_1}{B_1}\right)\left(-\frac{A_2}{B_2}\right) = -1, B_1 \neq 0, B_2 \neq 0, \text{ или } A_1 A_2 + B_1 B_2 = 0.$$

Следовательно, условием перпендикулярности двух прямых, заданных общими уравнениями, является равенство нулю суммы произведений коэффициентов при переменных x и y .

▷ПРИМЕР 2.5. Составить уравнения двух прямых, проходящих через точку $P(2; 1)$, одна из которых параллельна прямой $3x - 2y + 2 = 0$, а другая перпендикулярна той же прямой.

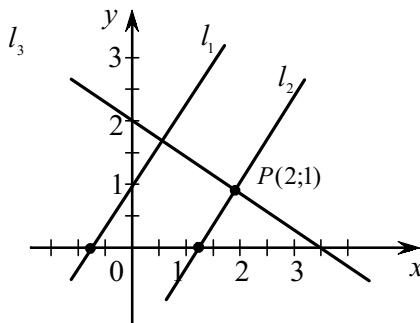


Рис. 2.14

Решение. Уравнение пучка прямых, проходящих через точку $P(2; 1)$, имеет вид $y - 1 = k(x - 2)$. Из этого пучка надо выделить две прямые l_2 и l_3 — параллельную и перпендикулярную данной прямой l_1 (рис. 2.14). Для прямой l_1 угловой коэффициент $k_1 = \frac{3}{2}$ (так как уравнение прямой l_1 можно предста-

вить в виде $y = \frac{3}{2}x + 1$). Угло-

вой коэффициент прямой l_2 (по условию параллельности) $k_2 = k_1 =$

$= \frac{3}{2}$, и ее уравнение имеет вид $y - 1 = \frac{3}{2}(x - 2)$, или $3x - 2y - 4 = 0$.

Угловой коэффициент прямой l_3 (по условию перпендикулярнос-

ти) $k = -\frac{1}{k_1} = -\frac{2}{3}$, и уравнение этой прямой $y - 1 = -\frac{2}{3}(x - 2)$, или

$2x + 3y - 7 = 0$.

Задачу можно решить и другим способом. Прямая $Ax + By + C = 0$ будет параллельна прямой $3x - 2y + 2 = 0$, если ее коэффи-

циенты при x и y пропорциональны, т.е. $\frac{A}{3} = \frac{B}{-2}$. Взяв $A = 3$,

$B = -2$ (при коэффициенте пропорциональности, равном 1), получим уравнение $3x - 2y + C = 0$. Коэффициент C найдем с учетом того, что координаты точки $A(2; 1)$, лежащей на прямой, должны

удовлетворять ее уравнению, т.е. $3 \cdot 2 - 2 \cdot 1 + C = 0$, откуда $C = -4$ и уравнение прямой l_2 : $3x - 2y - 4 = 0$.

Уравнение прямой, перпендикулярной данной: $3x - 2y + 2 = 0$, будет иметь вид $2x + 3y + C = 0$ (так как в этом случае сумма произведений коэффициентов при переменных x и y равна нулю, т.е. $3 \cdot 2 + (-2) \cdot 3 = 0$). Теперь, подставляя координаты точки $P(2; 1)$ в уравнение прямой, получим $2 \cdot 2 + 3 \cdot 1 + C = 0$, откуда $C = -7$ и уравнение прямой l_3 : $2x - 3y - 7 = 0$. ►

Точка пересечения прямых. Пусть на плоскости даны две прямые $A_1x + B_1y + C_1 = 0$ и $A_2x + B_2y + C_2 = 0$. Очевидно, что координаты их точки пересечения должны удовлетворять уравнению каждой прямой, т.е. они могут быть найдены из системы

$$\begin{cases} A_1x + B_1y + C_1 = 0, \\ A_2x + B_2y + C_2 = 0. \end{cases}$$

Если прямые не параллельны, т.е. $\frac{A_1}{A_2} \neq \frac{B_1}{B_2}$, то решение систе-

мы дает единственную точку пересечения. В этом случае определитель матрицы коэффициентов не равен нулю.

Если же прямые параллельны, то определитель матрицы коэффициентов равен нулю и система уравнений несовместна, т.е. решений

не имеет. В случае пропорциональных коэффициентов у уравнений прямые совпадают, тогда $n = 2$, $r = 1$, и система имеет бесчисленное множество решений.

Расстояние от точки до прямой. Пусть на плоскости заданы точка $M(x_0; y_0)$ и прямая FK : $Ax + By + C = 0$. Под расстоянием от точки M до прямой FK понимается длина перпендикуляра $d = MN$, опущенного из точки M на прямую FK (рис. 2.15).

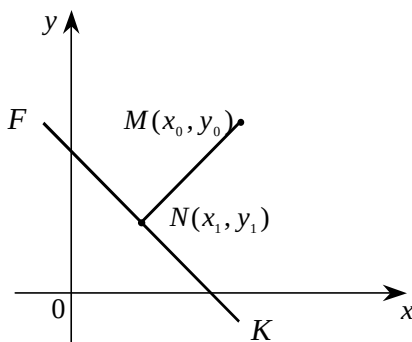


Рис. 2.15

Для определения расстояния d необходимо: а) составить уравнение прямой MN , перпендикулярной заданной прямой и проходящей через точку $M(x_0, y_0)$; б) найти точку $N(x_1, y_1)$ пересечения прямых, решив систему уравнений этих прямых; в) по формуле примера (2.1) определить расстояние между двумя точками, т.е. найти $d = MN$.

В результате преобразований получим

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}} \quad (2.11)$$

(доказательство формулы (2.11) опускаем).

▷ П Р И М Е Р 2.6. Найти расстояние между параллельными прямыми $3x + 4y - 24 = 0$ и $3x + 4y + 6 = 0$.

Р е ш е н и е. Возьмем на одной из прямых, например, $3x + 4y - 24 = 0$, произвольную точку $P(0; 6)$ (рис. 2.16).

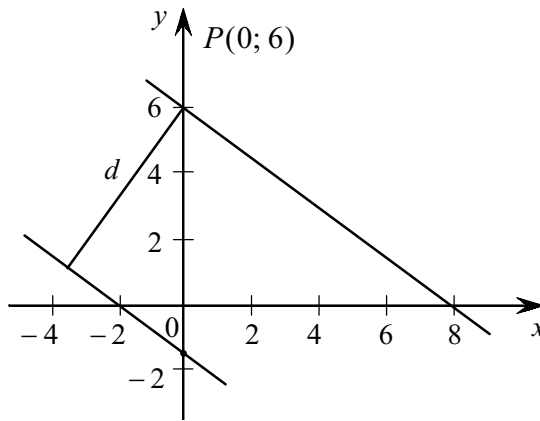


Рис. 2.16

Тогда искомое расстояние равно расстоянию от точки P до прямой $3x + 4y + 6 = 0$:

$$d = \frac{|3 \cdot 0 + 4 \cdot 6 + 6|}{\sqrt{3^2 + 4^2}} = 6. \blacktriangleright$$

Определение координат точки, делящей отрезок в заданном соотношении. Если точка $M(x, y)$ делит отрезок, определяемый точками $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$, в отношении $\frac{M_1M}{MM_2} = \lambda$, то координаты

точки M определяются по формулам

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}; \quad y = \frac{y_1 + \lambda y_2}{1 + \lambda}.$$

Если M — середина отрезка M_1M_2 , то $\lambda = 1$ и, следовательно,

$$x = \frac{x_1 + x_2}{2}; \quad y = \frac{y_1 + y_2}{2}.$$

1.4. Уравнение плоскости

ОПРЕДЕЛЕНИЕ. Уравнение, записанное в виде

$$Ax + By + Cz + D = 0, \tag{2.12}$$

называется общим уравнением плоскости и описывает плоскость в трехмерном пространстве.

Рассмотрим уравнение (2.12) как систему из одного матричного уравнения (см. параграф 5 главы 1). Пусть $|A| \neq 0$, и потому является базисным минором матрицы системы. Ранг расширенной матрицы равен единице, значит, система совместна. Одно ее решение можно найти, положив параметрические неизвестные равными нулю: $y = z = 0$. Тогда получим $x = -D/A$. Так как $n = 3, r = 1$, фундаментальная матрица имеет два столбца. Мы найдем их, придав параметрическим неизвестным два набора значений: $y = 1, z = 0$ и $y = 0, z = 1$. Соответствующие значения базисной неизвестной x , найденные из приведенной системы, будут $-B/A$ и $-C/A$. Итак, общее решение системы (2.12) есть

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -D/A \\ 0 \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} -B/A \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -C/A \\ 0 \\ 1 \end{pmatrix} \tag{2.13}$$

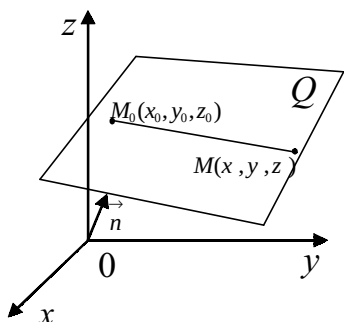


Рис. 2.17

Формула (2.13) описывает не что иное, как параметрические уравнения плоскости.

Пусть плоскость Q проходит через точку $M_0(x_0, y_0, z_0)$ перпендикулярно вектору $n = (A, B, C)$ (рис. 2.17).

Этими условиями определяется *единственная* плоскость в пространстве $Oxyz$. Вектор называется *нормальным вектором* плоскости Q . Возьмем в плоскости Q произвольную точку $M(x, y, z)$. Тогда вектор $M_0M = (x - x_0, y - y_0, z - z_0)$ будет перпендикулярен вектору $n = (A, B, C)$.

Следовательно, скалярное произведение этих векторов равно нулю, т.е.

$$(n, M_0M) = 0.$$

Полученное уравнение представим в координатной форме:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \quad (2.14)$$

Уравнение (2.14) представляет собой *уравнение плоскости*, перпендикулярной данному вектору $n = (A, B, C)$ и проходящей через данную точку $M_0(x_0, y_0, z_0)$.

В уравнении (2.12) тогда $D = -Ax_0 - By_0 - Cz_0$.

Можно доказать, что *всякое уравнение первой степени с тремя неизвестными есть уравнение плоскости*.

Если $D = 0$, то уравнение $Ax + By + Cz = 0$ определяет плоскость, проходящую через начало координат. Другие частные случаи зависят от расположения нормального вектора $n = (A, B, C)$.

Так, например, если $A = 0$, то уравнение (2.12) определяет плоскость, проходящую через ось Ox , если $A = B = 0$, то уравнение $Cz + D = 0$ определяет плоскость, параллельную оси Oxy ; если $A = B = D = 0$, то уравнение $Cz = 0$ (или $z = 0$) определяет координатную плоскость Oxy .

Рассмотрим несколько моментов, связанных с решением задач.

1. Предположим, что надо составить уравнение плоскости, которая проходит через точку P и перпендикулярна прямой PQ ,

если $P(p_1, p_2, p_3)$, $Q(q_1, q_2, q_3)$. Вектор $PQ = (q_1 - p_1, q_2 - p_2, q_3 - p_3) = (A_1, B_2, C_3)$.

Поэтому уравнение плоскости можно записать так:

$$A_1x + B_2y + C_3z + D = 0. \quad (2.15)$$

Так как точка P лежит в плоскости, то ее координаты должны удовлетворять уравнению: $Ap_1 + Bp_2 + Cp_3 + D = 0$. Отсюда находим D и подставляем в уравнение (2.15).

2. Допустим уравнением (2.12) задана плоскость, и требуется указать какой-нибудь вектор, параллельный плоскости. Ясно, что им будет вектор, перпендикулярный вектору $n = (A, B, C)$, удовлетворяющий уравнению (2.12).

3. Плоскость, пересекающая ось Ox в точке $(a; 0; 0)$, ось Oy в точке $(0; b; 0)$ и ось Oz — в точке $(0; 0; c)$, имеет уравнение

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1, \quad (2.16)$$

называемое *уравнением плоскости в отрезках*.

4. Предположим, что нам известны три точки в трехмерном пространстве $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, $P_3(x_3, y_3, z_3)$, не лежащие на одной прямой. Чтобы записать уравнение плоскости, проходящей через эти три заданные точки, нужно раскрыть определитель:

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0.$$

5. Для четырех точек $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, $P_3(x_3, y_3, z_3)$, $P_4(x_4, y_4, z_4)$, лежащих в одной плоскости, должно быть выполнено условие

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0. \quad (2.17)$$

6. Чтобы найти угол φ между двумя заданными плоскостями $a_1x + b_1y + c_1z + d_1 = 0$ и $a_2x + b_2y + c_2z + d_2 = 0$, нужно воспользоваться формулой:

$$\cos \varphi = \frac{n_1(a_1; b_1; c_1) \cdot n_2(a_2; b_2; c_2)}{|n_1| \cdot |n_2|} = \frac{a_1a_2 + b_1b_2 + c_1c_2}{\sqrt{(a_1^2 + b_1^2 + c_1^2)(a_2^2 + b_2^2 + c_2^2)}}.$$

Особый интерес представляют условия параллельности и перпендикулярности плоскостей, которые определяются соответственно условиями коллинеарности и перпендикулярности нормальных векторов $n_1 = (A_1, B_1, C_1)$ и $n_2 = (A_2, B_2, C_2)$.

Рассмотрим систему уравнений, каждое из которых — плоскость в трехмерном пространстве:

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \quad (2.18)$$

□ ТЕОРЕМА 2.1. *Плоскости, задаваемые в общей декартовой системе координат уравнениями (2.18), параллельны тогда и только тогда, когда соответствующие коэффициенты в их уравнениях пропорциональны, т.е. существует такое число λ , что*

$$A_2 = \lambda A_1, \quad B_2 = \lambda B_1, \quad C_2 = \lambda C_1. \quad (2.19)$$

Плоскости совпадают в том и только в том случае, когда их уравнения (2.18) пропорциональны, т.е. помимо уравнений (2.18) выполнено (с тем же λ) равенство

$$D_2 = \lambda D_1. \quad \blacksquare$$

Таким образом, условием параллельности двух плоскостей (2.18) является условие пропорциональности коэффициентов при

одноименных переменных: $\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} = \lambda$,

которое можно записать в виде

$$\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} = \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix} = \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = 0. \quad (2.20)$$

Условием *перпендикулярности* двух плоскостей является условие перпендикулярности их нормальных векторов $n_1 = (A_1, B_1, C_1)$ и $n_2 = (A_2, B_2, C_2)$.

Это возможно, когда скалярное произведение этих векторов равно нулю:

$$A_1A_2 + B_1B_2 + C_1C_2 = 0. \quad (2.21)$$

Решение системы из трех уравнений типа (2.12) с тремя неизвестными — это точки пространства, которые должны принадлежать одновременно трем плоскостям, которые описываются уравнениями системы. В этом случае возможны следующие варианты:

а) три плоскости пересекаются в одной точке, и система имеет единственное решение;

б) три плоскости пересекаются по одной прямой, т.е. система имеет бесчисленное множество решений (все точки на этой прямой);

в) две плоскости совпадают, а третья пересекает их - бесчисленное множество решений (все точки прямой пересечения трех плоскостей), ранг системы равен двум;

г) все три плоскости совпадают, т.е. все точки общей плоскости являются решениями, и ранг системы равен единице;

д) хотя бы одна из трех плоскостей параллельна какой-либо из двух других, т.е. решений нет, и система несовместна.

1.5. Уравнение прямой линии в пространстве

Прямая линия в пространстве может быть задана как пересечение двух плоскостей и, следовательно, определяется системой уравнений вида

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \quad (2.22)$$

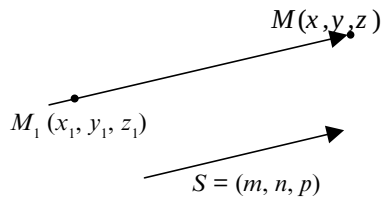


Рис. 2.18

Пересечение плоскостей — прямая линия тогда и только тогда, когда они не параллельны. Это, согласно (2.20), означает, что хотя бы один из определителей отличен от нуля.

Пусть прямая параллельна вектору $S = (m, n, p)$, называемому *направляющим вектором*, и проходит через точку $M_1(x_1, y_1, z_1)$ (рис. 2.18). Тогда ее уравнения могут быть получены из условия коллинеарности векторов $M_1M = (x - x_1, y - y_1, z - z_1)$ (здесь $M(x, y, z)$ — произвольная точка прямой) и $S = (m, n, p)$:

$$\frac{x - x_1}{m} = \frac{y - y_1}{n} = \frac{z - z_1}{p}. \quad (2.23)$$

ОПРЕДЕЛЕНИЕ. Уравнения вида (2.23) называются *каноническими уравнениями прямой линии* в пространстве.

1. Пусть в пространстве задана декартова система координат и две точки $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$.

Чтобы написать уравнение прямой P_1P_2 , примем P_1 за начальную точку, а P_1P_2 за направляющий вектор. Этот вектор не нулевой, если точки не совпадают.

Тогда уравнение прямой будет иметь вид:

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1} = t.$$

2. Условием параллельности двух прямых, заданных своими каноническими уравнениями, будет являться пропорциональность соответствующих координат их направляющих векторов.

Рассмотрим две прямые в пространстве l_1 и l_2 :

$$\frac{x - x_1}{m} = \frac{y - y_1}{n} = \frac{z - z_1}{p} \quad \text{и} \quad \frac{x - x_2}{m_1} = \frac{y - y_2}{n_1} = \frac{z - z_2}{p_1}.$$

Эти прямые параллельны, когда

$$\frac{m}{m_1} = \frac{n}{n_1} = \frac{p}{p_1}.$$

3. Чтобы три точки $P_1(x_1, y_1, z_1)$; $P_2(x_2, y_2, z_2)$ и $P_3(x_3, y_3, z_3)$ лежали на одной прямой, необходимо выполнение условий:

$$\frac{x_3 - x_1}{x_2 - x_1} = \frac{y_3 - y_1}{y_2 - y_1} = \frac{z_3 - z_1}{z_2 - z_1} \quad \text{или}$$

$$\begin{vmatrix} y_1 & z_1 & 1 \\ y_2 & z_2 & 1 \\ y_3 & z_3 & 1 \end{vmatrix} = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = \begin{vmatrix} z_1 & x_1 & 1 \\ z_2 & x_2 & 1 \\ z_3 & x_3 & 1 \end{vmatrix} = 0.$$

1.6. Взаимное расположение прямой и плоскости в пространстве

Пусть задана плоскость $Ax + By + Cz + D = 0$ и прямая (2.22).

1. Прямая параллельна плоскости, когда выполнено условие:

$$\Delta = \begin{vmatrix} A & B & C \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix} = 0.$$

2. Прямая $\frac{x - x_1}{m} = \frac{y - y_1}{n} = \frac{z - z_1}{p}$ и плоскость (2.12) параллель-

ны, если $tA + nB + pC = 0$.

Они взаимно перпендикулярны, если прямая параллельна нормали к плоскости, т.е. если

$$\frac{m}{A} = \frac{n}{B} = \frac{p}{C}.$$

3. Уравнение прямой, перпендикулярной к плоскости (2.12) и проходящей через точку $P_0(x_0, y_0, z_0)$:

$$\frac{x - x_0}{A} = \frac{y - y_0}{B} = \frac{z - z_0}{C}. \quad (2.24)$$

Здесь мы позволили себе привести формулы без доказательств и подробных объяснений. Заметим лишь, что все формулы получены путем исследования решений систем линейных алгебраических уравнений (см. главу 1).

§ 2. Кривые второго порядка

2.1. Эллипс и окружность

Рассмотрим уравнение второй степени с двумя переменными в общем виде

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0, \quad (2.25)$$

в котором A , B и C не равны нулю одновременно.

□ ТЕОРЕМА 2.2. Для любой линии, заданной уравнением второго порядка (2.25), существует декартова прямоугольная система координат, в которой это уравнение принимает один из девяти канонических видов:

$$1) \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1;$$

$$2) \frac{x^2}{a^2} + \frac{y^2}{b^2} = -1;$$

$$3) \frac{x^2}{a^2} + \frac{y^2}{b^2} = 0;$$

$$4) \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1;$$

$$5) \frac{x^2}{a^2} - \frac{y^2}{b^2} = 0;$$

$$6) y^2 = 2px;$$

$$7) y^2 - a^2 = 0;$$

$$8) y^2 + a^2 = 0;$$

$$9) y^2 = 0.$$

В соответствии с этим существуют семь классов линий второго порядка: 1) эллипсы; 3) точки (пары мнимых пересекающихся

прямых); 4) гиперболы; 5) пары пересекающихся прямых; 6) параболы; 7) пары параллельных прямых; 9) прямые (пары совпавших прямых).

Уравнениям вида 2) мнимого эллипса и 8) пары параллельных прямых не удовлетворяет ни одна точка. ■

Эллипс. Итак, мы назвали эллипсом линию, которая в некоторой декартовой системе координат определяется каноническим уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (2.26)$$

при условии, что кривая имеет только действительные точки (рис. 2.19).

Точки пересечения эллипса с осями канонической системы координат $A_1(-a; 0)$, $A_2(a; 0)$, $B_1(0; b)$, $B_2(0; -b)$ называются *вершинами эллипса*. Числа a и b называются соответственно *большой и малой полуосями* эллипса.

Точки $F_1(-c; 0)$ и $F_2(c; 0)$, где

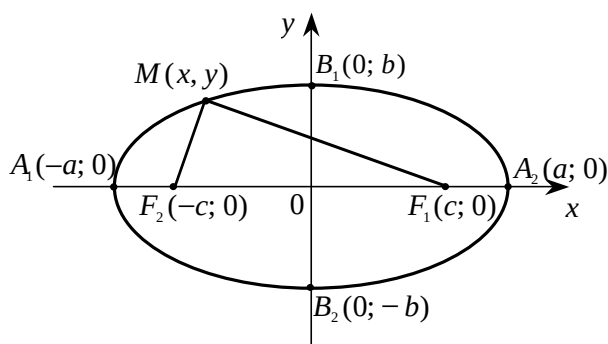


Рис. 2.19

$c = \sqrt{a^2 - b^2}$, (2.27) называются *фокусами* эллипса, а отношение

$$\varepsilon = \frac{c}{a} \quad (2.28)$$

его *эксцентриситетом*.

Эксцентриситет характеризует форму

эллипса.

Очевидно, что $0 \leq \varepsilon < 1$, причем, если $c = 0$, то $\varepsilon = 0$.

□ **ТЕОРЕМА 2.3.** Для того чтобы точка лежала на эллипсе, необходимо и достаточно, чтобы сумма ее расстояний до фокусов равнялась большой оси эллипса $2a$. ■

▷ **ПРИМЕР 2.7.** Определить вид и расположение кривой

$$x^2 + 2y^2 - 4x + 16y = 0. \quad (2.29)$$

Решение. Дополняя члены, содержащие x и y , до полного квадрата, получим

$$(x - 2)^2 + 2(y + 4)^2 = 36,$$

или

$$\frac{(x - 2)^2}{6^2} + \frac{(y + 4)^2}{(3\sqrt{2})^2} = 1.$$

Следовательно, кривая (2.29) представляет эллипс с полуосями $a = 6$ и $b = 3\sqrt{2}$, центр которого находится в точке $O'(2; -4)$ (рис. 2.20). ►

В частном случае при $a = b$ уравнение (2.26) есть уравнение окружности $x^2 + y^2 = a^2$. Для окружности $\varepsilon = 0$.

Продолжим изучение кривых второго порядка, описываемых уравнениями второй степени с двумя переменными.

Окружность. Пусть дана окружность радиуса R с центром $O'(x_0, y_0)$ (рис. 2.21). Найдем ее уравнение.

Для произвольной точки $M(x, y)$ окружности выполняется равенство $O'M = R$.

Используя формулу расстояния между двумя точками (см. пример 2.1), получим уравнение

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = R$$

или (после возведения в квадрат двух положительных частей уравнения) — равносильное уравнение

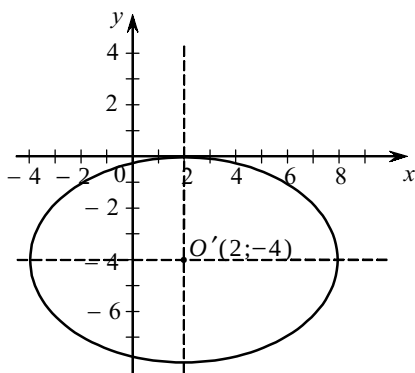


Рис. 2.20

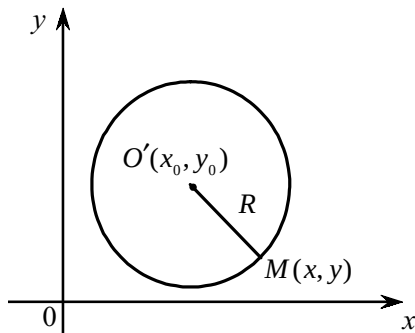


Рис. 2.21

$$(x - x_0)^2 + (y - y_0)^2 = R^2. \quad (2.30)$$

Итак, координаты каждой точки окружности $M(x, y)$ удовлетворяют уравнению (2.30). Нетрудно показать, что координаты любой точки, не лежащей на окружности, этому уравнению не удовлетворяют.

Уравнение (2.30) называется *нормальным уравнением окружности*. В частности, уравнение окружности с центром в начале координат ($x_0 = y_0 = 0$) имеет вид:

$$x^2 + y^2 = R^2. \quad (2.31)$$

▷ ПРИМЕР 2.8. Найти координаты центра и радиус окружности $x^2 + y^2 + 8y - 10 = 0$.

Решение. Дополнив члены, содержащие y , до полного квадрата, получим $x^2 + (y^2 + 8y + 16) - 16 - 10 = 0$, или $x^2 + (y + 4)^2 = 26$, т.е. центр окружности в точке $O(0; -4)$, а ее радиус $R = \sqrt{26}$. ►

2.2. Гипербола и парабола

Гипербола. Гиперболой мы называли линию, которая в некоторой декартовой прямоугольной системе координат определяется каноническим уравнением

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (2.32)$$

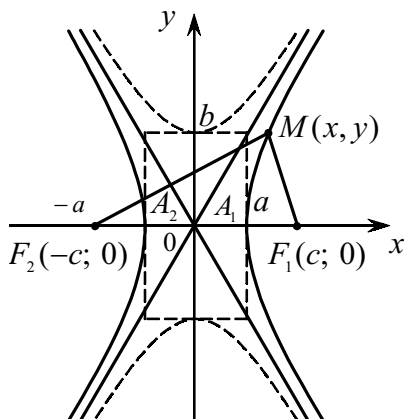


Рис. 2.22

Ось абсцисс канонической системы координат пересекает гиперболу в точках с координатами $A_1(a; 0)$ и $A_2(-a; 0)$, называемых *вершинами* гиперболы. Ось ординат не пересекает гиперболу. Таким образом, гипербола состоит из двух не связанных между собой частей. Они называются ее *ветвями*. Числа a и b называются соответственно вещественной и мнимой полуосями гиперболы (рис. 2.22).

Фокусы гиперболы — точки $F_1(c; 0)$ и $F_2(-c; 0)$, где $c = \sqrt{a^2 + b^2}$, а ее эксцентриситет $\varepsilon = \frac{c}{a}$ принимает любые значения, большие 1.

ОПРЕДЕЛЕНИЕ. Прямые с уравнениями $y = \pm \frac{b}{a}x$ называются **асимптотами** гиперболы.

□ **ТЕОРЕМА 2.4.** Для того чтобы точка M лежала на гиперболе, необходимо и достаточно, чтобы разность ее расстояний до фокусов по абсолютной величине равнялась $d = |F_2M - MF_1| = 2a$. ■

□ **ТЕОРЕМА 2.5.** Для гиперболы оси канонической системы координат являются осями симметрии, а начало канонической системы — центром симметрии. ■

▷ **ПРИМЕР 2.9.** Написать уравнение гиперболы с асимптотами $y = \pm \frac{3}{4}x$, проходящими через точку $A\left(6; \frac{3}{2}\right)$. Найти расстояние между ее вершинами.

Решение. Так как точка $A\left(6; \frac{3}{2}\right)$ лежит на гиперболе, то ее координаты должны удовлетворять уравнению:

$$\frac{36}{a^2} - \frac{9}{4b^2} = 1.$$

Кроме того, $\frac{b}{a} = \frac{3}{4}$, так как асимптоты гиперболы $y = \pm \frac{3}{4}x$,

Решая полученную систему двух уравнений, найдем $a = 4\sqrt{2}$, $b = 3\sqrt{2}$, т.е. уравнение гиперболы $\frac{x^2}{32} - \frac{y^2}{18} = 1$ (рис. 2.23).

Расстояние между вершинами гиперболы равно $2a = 8\sqrt{2}$. ►

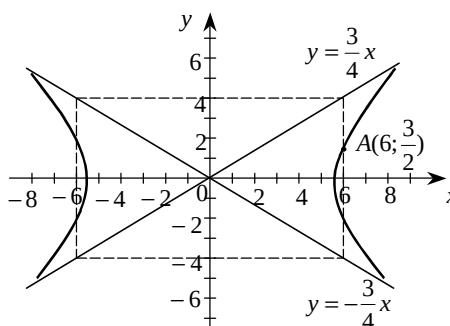


Рис. 2.23

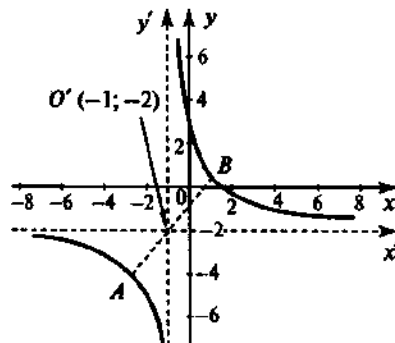


Рис. 2.24

▷ П Р И М Е Р 2.10. Найти координаты центра, вершин и уравнения асимптот гиперболы $y = \frac{3-2x}{x+1}$.

Р е ш е н и е. Преобразуем уравнение, выделив целую часть дробно-линейной функции:

$$y = \frac{-2(x+1)+5}{x+1} = -2 + \frac{m}{x+1},$$

где $m = 5$, или $y + 2 = \frac{5}{x+1}$,

откуда $(x+1)(y+2) = 5$.

Полагая $x+1 = x'$, $y+2 = y'$, получим $x'y' = 5$, т.е. заданное уравнение равносторонней гиперболы с центром $O'(-1; -2)$ и асимптотами $x+1=0$, $y+2=0$ (рис. 2.24). Так как $m = 5 > 0$, то гипербола располагается в I и III квадрантах, а новые координаты ее вершин $(\pm\sqrt{5}; \pm\sqrt{5})$. Переходя к старым координатам по формулам $x = x' - 1$, $y = y' - 2$, найдем старые координаты вершин гиперболы:

$A(-\sqrt{5}-1; -\sqrt{5}-2)$, $B(\sqrt{5}-1; \sqrt{5}-2)$. ►

Здесь нельзя не сказать о каноническом виде гиперболы, известном из школьного курса геометрии: $xy = k$. Можно привести множество примеров из области экономики, сводящихся к построению этих кривых.

Вспомним из курса экономической теории, что инвестиция — это функция от процентной ставки, а процентная ставка — это тот процент денег, который пользователь капитала должен заплатить владельцу.

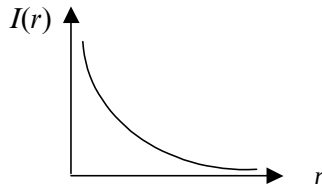


Рис. 2.25

Чем выше r , тем ниже объем инвестиций, так как пользователю капитала при высокой процентной ставке невыгодно брать кредиты. Зависимость инвестиций от r называется инвестиционной функцией $I(r)$ (рис. 2.25), простейшая ее модель — верхняя ветвь гиперболы.

Более подробно этот материал изложен в работе [18].

Парабола. *Параболой* мы назвали линию, которая в некоторой декартовой прямоугольной системе координат определяется каноническим уравнением

$$y^2 = 2px, \quad (2.33)$$

где $p > 0$.

Из уравнения (2.33) вытекает, что для всех точек параболы $x \geq 0$. Парабола проходит через начало канонической системы координат. Эта точка называется *вершиной* параболы.

Форма параболы известна из курса средней школы, где она встречается в качестве графика функции $y = ax^2$. Отличие уравнений объясняется тем, что в канонической системе координат, по сравнению с прежней, оси координат поменялись местами, а коэффициенты связаны равенством $2p = a^{-1}$.

Если точка $O'(x_0, y_0)$ — *вершина* параболы, то парабола имеет вид:

$$(y - y_0)^2 = 2p(x - x_0). \quad (2.34)$$

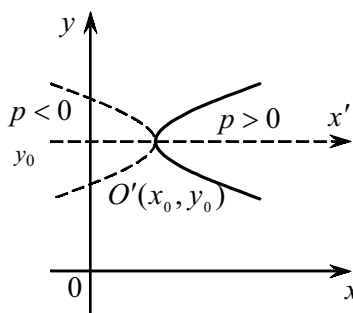


Рис. 2.26

При $p > 0$ ветви параболы направлены вправо, при $p < 0$ — влево (рис. 2.26). Прямая $y = y_0$ является осью симметрии параболы.

Точка $F\left(\frac{p}{2}; 0\right)$ называется фокусом параболы, а прямая $x = -\frac{p}{2}$ —

ее директрисой (рис. 2.27).

Для произвольной точки $M(x, y)$ параболы расстояние до фокуса MF по формуле примера 2.1 равно

$$|MF| = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \sqrt{x^2 - px + \frac{p^2}{4} + 2px} = \sqrt{\left(x + \frac{p}{2}\right)^2} = x + \frac{p}{2}$$

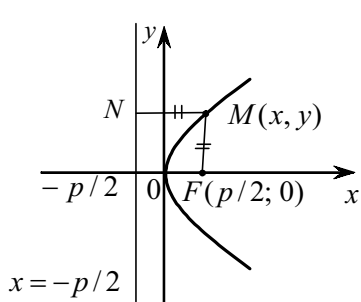


Рис. 2.27

(так как $x + \frac{p}{2} \geq 0$). С другой стороны, расстояние до директрисы

$$MN = x + \frac{p}{2}.$$

□ **ТЕОРЕМА 2.6.** Для того чтобы точка M лежала на параболе, необходимо и достаточно, чтобы она была одинаково удалена от фокуса и от директрисы этой параболы. ■

Если в уравнении (2.33) поменять местами x и y , то получим $x^2 = 2py$ — уравнение параболы с вершиной в начале координат, симметричной относительно оси ординат. Это уравнение обычно

записывают в виде $y = ax^2$, где $a = \frac{1}{2p}$.

Рассмотрим квадратный трехчлен

$$y = Ax^2 + Bx + C \quad (A \neq 0). \quad (2.35)$$

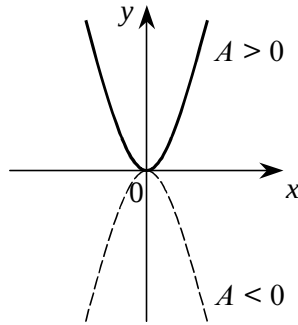


Рис. 2.28

Графиком функции (2.35) при $B = C = 0$ является парабола. При $A > 0$ ветви параболы направлены вверх, при $A < 0$ — вниз (рис. 2.28).

Отсюда $y = A \left(x^2 + \frac{B}{A}x + \frac{C}{A} \right)$. Дополнив выражение, стоящее в скобках, до полного квадрата, получим

$$y = A \left[\left(x + \frac{B}{2A} \right)^2 + \frac{C}{A} - \frac{B^2}{4A^2} \right] = A \left(x + \frac{B}{2A} \right)^2 + \frac{4AC - B^2}{4A}. \quad (2.36)$$

Обозначим $x + \frac{B}{2A} = x'$, $y - \frac{4AC - B^2}{4A} = y'$, тогда в новой системе координат $O'x'y'$ с центром $O' \left(-\frac{B}{2A}, \frac{4AC - B^2}{4A} \right)$ уравнение (2.36) примет вид $y' = A (x')^2$.

Таким образом, график квадратного трехчлена $y = Ax^2 + Bx + C$ есть парабола с вершиной в точке $O' \left(-\frac{B}{2A}, \frac{4AC - B^2}{4A} \right)$ и осью

симметрии $x = -\frac{B}{2A}$, параллельной оси Oy .

▷ П Р И М Е Р 2.11. Построить кривую $y = -3x^2 + 10x - 3$.

Р е ш е н и е. Вынося коэффициент при x^2 и дополняя правую часть уравнения до полного квадрата, получим

$$y = -3 \left(x^2 - \frac{10}{3}x + 1 \right) = -3 \left[\left(x - \frac{5}{3} \right)^2 + 1 - \frac{25}{9} \right] = -3 \left(x - \frac{5}{3} \right)^2 + \frac{16}{3},$$

$$\text{или } y - \frac{16}{3} = -3 \left(x - \frac{5}{3} \right)^2.$$

$$\text{Полагая } x - \frac{5}{3} = x', \quad y - \frac{16}{3} = y',$$

получим $y' = -3(x')^2$.

Таким образом, заданная кривая есть парабола с вершиной в

точке $O' \left(\frac{5}{3}; \frac{16}{3} \right)$ и осью симметрии $O'y'$, параллельной оси Ox

(рис. 2.29). ►

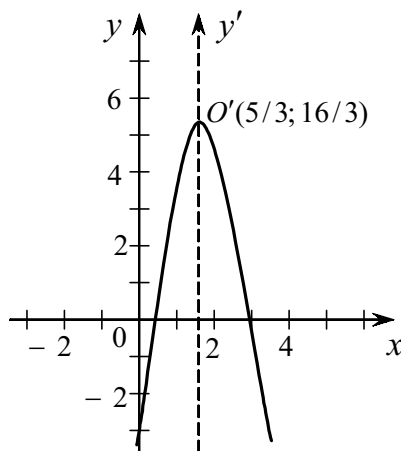


Рис. 2.29

2.3. Примеры решения задач

▷ П Р И М Е Р 2.12. Даны уравнения сторон треугольника $3x - 4y + 24 = 0$ (AB), $4x + 3y + 32 = 0$ (BC), $2x - y - 4 = 0$ (AC). Составить уравнение высоты, медианы и биссектрисы, проведенных из вершины B , и найти их длины.

Р е ш е н и е. 1. Найдем координаты вершин треугольника, решив соответствующие системы уравнений сторон. Так, координаты вершины B определим из системы уравнений прямых AB и BC :

$$\begin{cases} 3x - 4y + 24 = 0, \\ 4x + 3y + 32 = 0, \end{cases}$$

$x = -8, y = 0$, т.е. $B(-8; 0)$. Аналогично найдем координаты вершин A и C , решив системы уравнений прямых AB и AC , AC и BC : $A(8; 12)$, $C(-2; -8)$ (рис. 2.30).

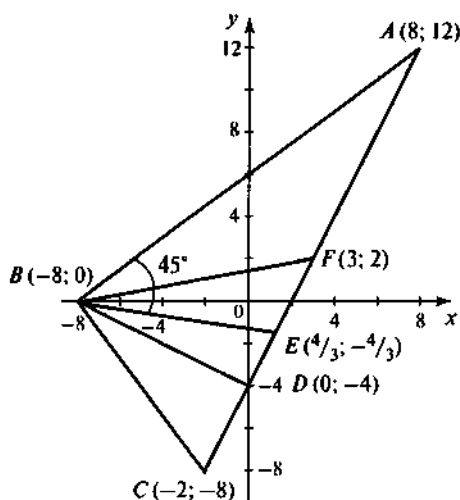


Рис. 2.30

2. Пучок прямых, проходящих через точку $B(-8; 0)$, по формуле (2.5) имеет вид

$$y = k(x + 8). \quad (2.37)$$

Из уравнения прямой AC следует, что ее угловой коэффициент $k_{AC} = 2$. На основании условия перпендикулярности двух пря-

мых $k_{BD} = \frac{-1}{k_{AC}} = -\frac{1}{2}$. Таким образом, уравнение высоты BD при-

мет вид: $y = -\frac{1}{2}(x + 8)$, или $x + 2y + 8 = 0$.

3. Известно, что координаты середины отрезка равны полусуммам соответствующих координат его концов (см. подпараграф 1.3 главы 2), т.е.

$$x_F = \frac{x_A + x_C}{2}, \quad y_F = \frac{y_A + y_C}{2}.$$

Поэтому $x_F = \frac{8 + (-2)}{2} = 3$, $y_F = \frac{12 + (-8)}{2} = 2$, т.е. $F(3; 2)$.

По формуле (2.6) угловой коэффициент

$$k_{BF} = \frac{y_F - y_B}{x_F - x_B} = \frac{2 - 0}{3 - (-8)} = \frac{2}{11}.$$

Подставляя $k = \frac{2}{11}$ в формулу (2.37), получим уравнение медианы BF :

$$y = \frac{2}{11}(x + 8), \text{ или } 2x - 11y + 16 = 0.$$

(Уравнение BF можно было получить и по формуле (2.7) как уравнение прямой, проходящей через две точки: $B(8; 0)$ и $F(3; 2)$.)

4. Из уравнений прямых $3x - 4y + 24 = 0$ (AB) и $4x + 3y + 32 = 0$ (BC) следует, что они перпендикулярны, так как их угловые коэффициенты $k_{AB} = \frac{3}{4}$ и $k_{BC} = -\frac{4}{3}$ обратны по величине и противоположны по знаку. Поэтому биссектриса BE образует с каждой из этих сторон угол 45° . По формуле (2.10)

$$\operatorname{tg} 45^\circ = \frac{k_{AB} - k_{BE}}{1 + k_{AB} \cdot k_{BE}} = \frac{\frac{3}{4} - k_{BE}}{1 + \frac{3}{4}k_{BE}} = 1,$$

откуда $k_{BE} = -\frac{1}{7}$. Теперь по формуле (2.37) получим уравнение биссектрисы BE :

$$y = -\frac{1}{7}(x + 8), \text{ или } x + 7y + 8 = 0.$$

5. Длину медианы BF найдем по формуле примера 2.1 расстояния между двумя точками $B(-8; 0)$ и $F(3; 2)$:

$$d_{BF} = \sqrt{(3+8)^2 + (2-0)^2} = \sqrt{125} = 5\sqrt{5} \approx 11,2.$$

6. Для нахождения длины биссектрисы BE найдем вначале координаты ее точки пересечения E со стороной AC , решив систему уравнений

$$\begin{cases} 2x - y - 4 = 0, \\ x + 7y + 8 = 0, \end{cases}$$

откуда $x = \frac{4}{3}$; $y = -\frac{4}{3}$, т.е. $E\left(\frac{4}{3}; -\frac{4}{3}\right)$

Теперь по формуле примера 2.1

$$d_{BE} = \sqrt{\left(\frac{4}{3} + 8\right)^2 + \left(-\frac{4}{3} - 0\right)^2} = \frac{20\sqrt{2}}{3} \approx 9,4.$$

7. Длину высоты BD можно было найти аналогично тому, как была найдена длина биссектрисы. Но проще это сделать по формуле (2.11) расстояния от точки $B(-8; 0)$ до прямой $2x - y - 4 = 0$ (AC):

$$d_{BD} = \frac{|2(-8) + (-1) \cdot 0 - 4|}{\sqrt{2^2 + (-1)^2}} = 4\sqrt{5} \approx 8,9. \blacktriangleright$$

▷ П Р И М Е Р 2.13. Найти расстояние от начала координат до прямой, проходящей через центр гиперболы $y = \frac{4x-4}{2x+1}$ и вершину параболы $y = -x^2 + 2x - \frac{1}{3}$. Составить уравнение окружности, касающейся гиперболы в ее вершинах.

Р е ш е н и е. 1. В уравнении гиперболы выделим целую часть, получим

$$y = \frac{4(x-1)}{2\left(x + \frac{1}{2}\right)} = \frac{4\left(x + \frac{1}{2}\right) - 6}{2\left(x + \frac{1}{2}\right)} = 2 - \frac{3}{x + \frac{1}{2}},$$

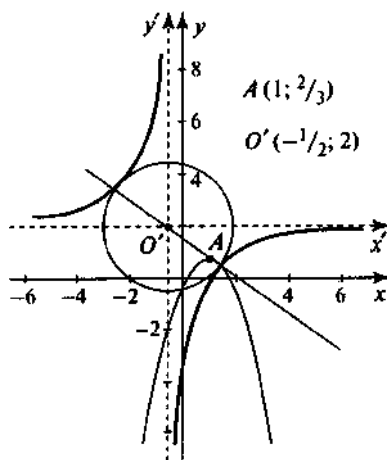


Рис. 2.31

откуда $y - 2 = -\frac{3}{x + \frac{1}{2}},$

или $\left(x + \frac{1}{2}\right)(y - 2) = -3.$

Полагая $x + \frac{1}{2} = x', y - 2 =$
 $= y',$ получим в новой сис-
 теме координат $O'x'y'$ с цен-
 тром $O'\left(-\frac{1}{2}; 2\right)$ гиперболу

$x'y' = -3,$ ветви которой расположены во II и IV квадрантах (рис. 2.31).

2. Выделив полный квадрат, представим уравнение параболы в виде

$$y = -x^2 + 2x - \frac{1}{3} = -(x - 1)^2 - \frac{1}{3} + 1 = -(x - 1)^2 + \frac{2}{3},$$

откуда следует, что вершина параболы находится в точке $A\left(1; \frac{2}{3}\right)$

а ветви ее направлены вниз.

3. Составляем уравнение прямой $O'A$ по формуле (2.7):

$$\frac{y - 2}{\frac{2}{3} - 2} = \frac{x + \frac{1}{2}}{1 + \frac{1}{2}}, \text{ или } 8x + 9y - 14 = 0.$$

4. Найдем расстояние от точки $O(0; 0)$ до прямой $8x + 9y - 14 = 0$ по формуле (2.11):

$$d = \frac{|8 \cdot 0 + 9 \cdot 0 - 14|}{\sqrt{8^2 + 9^2}} = \frac{14}{\sqrt{145}} \approx 1,2.$$

5. Очевидно, что центр искомой окружности должен совпасть с центром гиперболы $O' \left(-\frac{1}{2}; 2 \right)$, а ее радиус R равен расстоянию от точки O' до любой из вершин гиперболы. Для гиперболы $x'y' = m$ координаты любой вершины (по абсолютной величине) $|x'| = |y'| = \sqrt{|m|}$, поэтому расстояние ее от нового начала координат $(0; 0)$ (см. пример 2.1 главы 2) равно $\sqrt{2|m|}$. Следовательно, $R = \sqrt{2|m|} = \sqrt{2 \cdot |-3|} = \sqrt{6}$. Итак, уравнение искомой окружности по формуле (2.30) есть

$$\left(x + \frac{1}{2} \right)^2 + (y - 2)^2 = 6. \blacktriangleright$$

2.4. Кривые спроса и предложения. Точка равновесия

Очень часто нам приходится сталкиваться с построением кривых на плоскости при решении экономических задач. Вспомним некоторые основные положения курса экономической теории.

Допустим, что имеется один продукт, может изменяться только его цена, а все остальные факторы, от которых зависит спрос на данный товар (цены на другие товары, основные производственные фонды, характер применяемой технологии, налоги и дотации, природно-климатические условия и т. д.), остаются неизменными.

Кривые спроса и предложения. Рассмотрим зависимости спроса D (demand) и предложения S (supply) от цены на товар P (price).

Чем меньше цена, тем больше спрос при постоянной покупательной способности населения. Обычно зависимость D от P имеет вид ниспадающей кривой $Q = D(P)$ (рис. 2.32, а), а уравнение линии спроса имеет вид:

$$D = kP^a + c, \tag{2.38}$$

где $a < 0$, k — коэффициент пропорциональности, который определяется опытным путем, $c = const$.

Здесь лишь отметим, что линия спроса может быть как прямой $D(P) = a - kP$, где $0 < P < a/k$, так и частью гиперболы $D(P) = 1/P$, $P > 0$.

В свою очередь предложение растет с увеличением цены на товар, и потому зависимость S от P $Q = S(P)$ имеет следующую характерную форму:

$$S = P^b + d, \quad (2.39)$$

где $b \geq 1$, $d = const$, (рис. 2.32, б).

В формулах (2.38) и (2.39) c и d — так называемые экзогенные величины; они зависят не от цены, а от ряда других причин (благополучие общества, политическая обстановка и т.п.). Вполне понятно, что переменные, входящие в формулы (2.38) и (2.39), положительны (кроме $a < 0$), поэтому графики кривых имеют смысл только в первой координатной четверти.

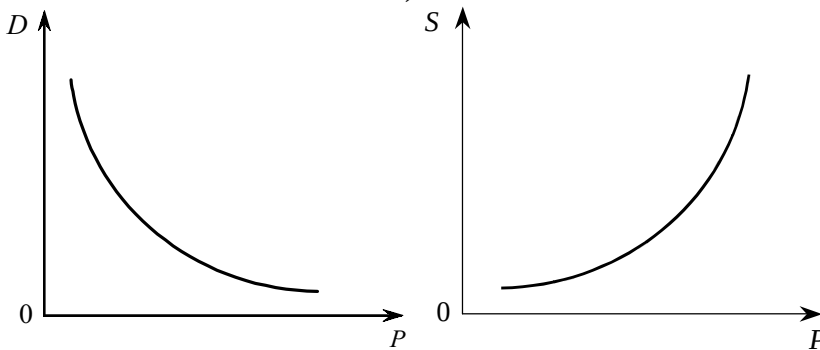


Рис. 2.32

Точка равновесия. Для экономики представляет интерес условие равновесия, т.е. когда спрос равен предложению. Это условие задается уравнением

$$D(P) = S(P) \quad (2.40)$$

и соответствует точке пересечения кривых D и S — это так называемая точка равновесия (точка M на рис. 2.33). Цена P_0 , при которой выполняется условие (2.40), называется равновесной.

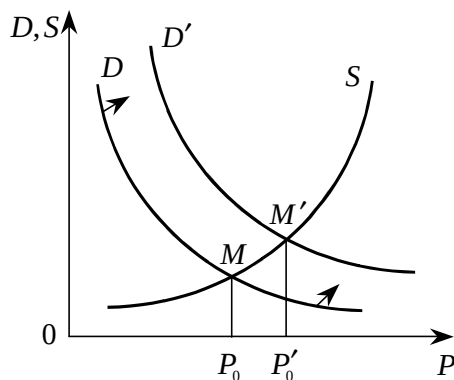


Рис. 2.33

При увеличении благосостояния населения, что соответствует росту величины c в формуле (2.38), точка равновесия M смещается вправо, так как кривая D поднимается вверх; при этом цена на товар растет при неизменной кривой предложения S .

Если цена на рынке выше равновесной, то предложение превышает спрос и возникает затоваривание рынка.

В этой ситуации товаропроизводители (продавцы) многих видов товаров готовы пойти на снижение цены с целью привлечения большего числа покупателей (например, если речь идет о скоропортящихся товарах).

Следовательно, при значениях цены выше равновесной происходит давление на нее в сторону уменьшения.

Если же цена на рынке ниже равновесной, то спрос превышает предложение и товар становится дефицитным. В этой ситуации часть покупателей готова заплатить за товар более высокую цену, но снизить риск и с уверенностью приобрести товар (например, если образуется очередь покупателей, то стоящие в ее конце могут не получить товара). Таким образом, при значениях цены ниже равновесной происходит давление на нее в сторону увеличения. Эти две тенденции приводят к тому, что на рынках многих видов товаров, как правило, устанавливается равновесие, при котором спрос равен предложению.

В силу свойств кривых спроса и предложения равновесное решение является устойчивым в том смысле, что если цена строго

фиксирована и равна равновесной $P = P_0$, то товаропроизводитель, максимизируя прибыль, поставляет на рынок товар в количестве $S(P_0) = Q_0$. Одновременно потребитель, стремясь максимизировать полезность, предъявляет спрос $D(P_0) = Q_0$. При установлении на рынке *совершенной конкуренции* равновесной цены объем товаров, предлагаемый товаропроизводителем и доставляющий ему максимум прибыли по данной цене, в точности равен спросу потребителя.

Динамические неравновесные модели рынка используются для анализа изменения переменных (цена, спрос, предложение) во времени в случае, когда цена в начальный момент отличается от равновесной. При этом процесс установления равновесной цены может быть описан различными моделями [12], [13] при использовании одних и тех же уравнений кривых спроса (2.38) и предложения (2.39).

На исследовании поведения кривых спроса и предложения мы остановимся снова в подпараграфе 15.1 главы 4.

Контрольные вопросы

1. В чем различие при построении уравнения прямой с угловым коэффициентом и уравнения прямой в отрезках?
2. Как вычислить угол между двумя прямыми? Каковы условия параллельности и перпендикулярности двух прямых?
3. Как определить расстояние от точки до прямой?
4. Каким образом все кривые второго порядка объединяются в одно каноническое уравнение? Привести канонические уравнения кривых второго порядка.
5. Где находятся основные характеристики кривых второго порядка: фокусы, эксцентриситеты, полуоси, директрисы, центры? Показать на примерах.

Упражнения

1. Найти уравнение множества точек, равноудаленных от оси Oy и точки $F(4; 0)$.
2. Составить уравнение прямой, проходящей через точку $A(2; 3)$: а) параллельно оси Ox ; б) параллельно оси Oy ; в) составляющей с осью Ox угол 45° .

3. Составить уравнение прямой, проходящей через точки:
а) $A(3; 1)$ и $B(5; 4)$; б) $A(3; 1)$ и $C(3; 5)$; в) $A(3; 1)$ и $D(-4; 1)$.

4. Стороны AB , BC и AC треугольника ABC заданы соответственно уравнениями $4x + 3y - 5 = 0$, $x - 3y + 10 = 0$, $x - 2 = 0$. Определить координаты его вершин.

5. Ординаты всех точек окружности $x^2 + y^2 = 36$ сокращены втрое. Написать уравнение полученной новой кривой.

6. Для гиперболы $3x^2 - 4y^2 = 12$ найти: действительную и мнимую полуоси; координаты фокусов; эксцентриситет; уравнения асимптот.

7. Написать уравнение гиперболы, имеющей вершины в фокусах, а фокусы — в вершинах эллипса $\frac{x^2}{25} + \frac{y^2}{9} = 1$.

8. Составить уравнение параболы, проходящей через точки:
а) $(0; 0)$ и $(-1; -3)$ симметрично относительно оси Ox ; б) $(0; 0)$ и $(2; -4)$ симметрично относительно оси Oy .

9. Привести пример четырех точек пространства, лежащих в одной плоскости.

10. Привести пример трех точек в пространстве, лежащих на одной прямой.

Ответы к упражнениям

1. $y^2 = 8(x - 2)$ 2. а) $y = 3$; б) $x = 2$; в) $y = x + 1$. 3. а) $3x - 2y - 7 = 0$;
б) $x = 3$; в) $y = 1$. 4. $A(2; -1)$; $B(-1; 3)$; $C(2; 4)$. 5. $\frac{x^2}{36} + \frac{y^2}{4} = 1$.

6. $a = 2$, $b = \sqrt{3}$; $F_1(-\sqrt{7}; 0)$, $F_2(\sqrt{7}; 0)$; $\varepsilon = \frac{\sqrt{7}}{2}$; $y = \pm \frac{\sqrt{3}}{2}x$.

7. $\frac{x^2}{16} - \frac{y^2}{9} = 1$. 8. $y^2 = -9x$; $x^2 = -y$.

Г л а в а 3

ОСНОВЫ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Большое число экономических задач сводится к линейным математическим моделям. Традиционно оптимизационные линейные математические модели называются моделями линейного программирования. Этот термин появился в конце 30-х годов прошлого столетия, когда программирование на компьютере еще не было развито, и соответствует не очень удачному переводу английского “programming”. Под линейным программированием понимается линейное планирование, т.е. получение оптимального плана-решения в задачах с линейной структурой.

Вообще говоря, в последнее время задачи линейного программирования принято решать только с использованием компьютерных программ, точнее, пакетов программ. Если для решения использовать готовые пакеты, остается открытым вопрос непосредственно о методах решения, заложенных внутри программ.

Мы постараемся на доступном для понимания уровне изложить некоторые методы решения задач линейного программирования.

§1. Постановка задачи линейного программирования в экономике

В общем виде **задача линейного программирования** ставится следующим образом: максимизировать (минимизировать) функцию

$$f = \sum_{j=1}^n c_j x_j \tag{3.1}$$

при выполнении условий:

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j = b_i & (i = \overline{1, m_1}); \end{cases} \quad (3.2)$$

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j \leq b_i & (i = \overline{m_1 + 1, m_2}); \end{cases} \quad (3.3)$$

$$\begin{cases} \sum_{j=1}^n a_{ij} x_j \leq b_i & (i = \overline{m_2, m}), \end{cases} \quad (3.4)$$

где f — целевая функция, или критерий эффективности (оптимальности) задачи; x_j , $j = \overline{1, n}$ — управляющие переменные, или решения задачи (3.1)–(3.4); b_i , a_{ij} , $i = \overline{1, m}$ — параметры.

Функция (3.1) — линейная, ограничения (3.2)–(3.4) — линейные. Задача содержит n переменных и m ограничений.

Решить задачу линейного программирования — это значит найти значения управляющих переменных x_j , $j = \overline{1, n}$, удовлетворяющих ограничениям (3.2)–(3.4), при которых целевая функция (3.1) принимает минимальное или максимальное значение.

В зависимости от вида целевой функции (3.1) и ограничений (3.2)–(3.4) можно выделить несколько типов задач линейного программирования, или линейных моделей: *общая линейная задача*, *транспортная задача*, и др.

В качестве примера общей линейной задачи рассмотрим **задачу составления плана реализации товара**.

Фирма реализует различные товары, используя при этом определенный набор средств (технических, людских, денежных).

Общий запас средств, запас средств каждого вида, используемых при реализации единицы любого товара, и прибыль от его продажи заданы. Сформировать план реализации товаров, приносящий фирме максимальную прибыль.

Построим математическую модель данной задачи.

1. Цель — максимизация прибыли.

2. Параметры:

n — число различных видов реализуемых товаров;

m — число разных видов средств;

Для изготовления изделий используется четыре типа ресурсов. Количество ресурсов, потребляемых для производства одного изделия, общее количество ресурсов и прибыль от реализации каждого вида изделия заданы в табл. 3.1.

Как организовать производство, чтобы обеспечить заказчиков, не допустить затоваривания, получить при этом максимальную прибыль?

Таблица 3.1

Тип ресурсов	Вид изделий			Всего ресурсов
	1	2	3	
1	500	300	1000	25 000 000
2	1000	200	100	30 000 000
3	150	300	200	20 000 000
4	100	200	400	40 000 000
Прибыль	20	40	50	

Выполним последовательно этапы построения математической модели.

1. Цель — получение максимальной прибыли.

2. Параметрами являются все числовые данные, приведенные в условии задачи.

3. Управляющие переменные:

x_1 — число изделий первого вида;

x_2 — число изделий второго вида;

x_3 — число изделий третьего вида.

4. Ограничения: обеспечить заказчиков, не превысить запас ресурсов, не допустить затоваривания рынка.

В соответствии с этими ограничениями выпишем область допустимых решений задачи:

$$\left\{ \begin{array}{l} x_1 \geq 1000, \\ x_2 \geq 2000, \\ x_3 \geq 2500, \\ x_1 \leq 2000, \\ x_2 \leq 3000, \\ x_3 \leq 5000, \\ 500x_1 + 300x_2 + 1000x_3 \leq 25\,000\,000, \\ 1000x_1 + 200x_2 + 100x_3 \leq 30\,000\,000, \\ 150x_1 + 300x_2 + 200x_3 \leq 20\,000\,000, \\ 100x_1 + 200x_2 + 400x_3 \leq 40\,000\,000. \end{array} \right. \quad (3.7)$$

Первые три неравенства в системе (3.7) соответствуют спросу заказчиков. Неравенства с четвертого по шестое формализуют спрос на рынке. Последние четыре неравенства соответствуют ограничениям по ресурсам.

5. Целевая функция, или критерий эффективности задачи, имеет вид

$$P = 20x_1 + 40x_2 + 50x_3 \rightarrow \max. \quad (3.8)$$

В формуле (3.8) буквой P обозначена прибыль. Ее надо максимизировать. Каждое слагаемое определяет прибыль от производства изделий каждого вида соответственно в количествах x_1 , x_2 , x_3 .

Формулы (3.7), (3.8) — математическая модель поставленной задачи. Ограничения и целевая функция линейны по управляющим переменным, следовательно, данная модель является линейной. (При составлении модели предполагалось, что прибыль линейно зависит от числа реализуемых изделий.)

§2. Графический метод решения задачи линейного программирования

Если число переменных в задаче линейного программирования (ЗЛП) равно двум, а ограничениями является система неравенств, то задачу можно решать графическим методом.

▷ П Р И М Е Р 3.1. При продаже двух видов товара используется четыре типа ресурсов. Норма затрат ресурсов на реализацию единицы товара и общий объем каждого ресурса заданы в табл. 3.2.

Таблица 3.2

Тип ресурса	Норма затрат ресурсов на товары вида		Общее количество ресурсов
	1	2	
1	2	2	12
2	1	2	8
3	4	0	16
4	0	4	12

Прибыль от реализации одной единицы товара первого вида составляет 2 усл. ед., второго вида — 3 усл. ед.

Требуется найти оптимальный план реализации товаров, обеспечивающий торговому предприятию максимальную прибыль.

Р е ш е н и е. Это задача составления плана реализации товара (см. § 1 главы 3) при $n = 2$, $m = 4$.

Математическая модель имеет вид

$$P = 2x_1 + 3x_2 \rightarrow \max, \quad (3.9)$$

$$\left\{ \begin{array}{l} 2x_1 + 2x_2 \leq 12, \\ x_1 + 2x_2 \leq 8, \\ 4x_1 \leq 16, \\ 4x_2 \leq 12, \\ x_1 \geq 0, \\ x_2 \geq 0. \end{array} \right. \quad (3.10)$$

В модели управляющие переменные x_1, x_2 — количество реализуемых изделий первого и второго вида соответственно, P — прибыль.

Система неравенств включает ограничения по ресурсам. Количество ресурсов на реализацию товаров первого и второго вида не превышает общего количества ресурсов каждого типа.

Решим задачу графическим методом.

Построим в плоскости $x_1 O x_2$ область допустимых решений (рис. 3.1). Каждое неравенство системы (3.10) определяет в плоскости $x_1 O x_2$ полуплоскость, лежащую выше или ниже прямой, определяемой соответствующим уравнением. Построим прямые

$$\begin{aligned} 2x_1 + 2x_2 &= 12, \\ x_1 + 2x_2 &= 8, \\ 4x_1 &= 16, \\ 4x_2 &= 12; \\ x_1 &= 0, \\ x_2 &= 0. \end{aligned}$$

Рассмотрим точку с координатами $x_1 = 0; x_2 = 0$. Подставив их в первое неравенство, получаем $0 \leq 12$ — верно, следовательно, иско-

мая полуплоскость лежит ниже прямой $2x_1 + 2x_2 = 12$; остальные полуплоскости находятся аналогичным образом (см. рис. 3.1).

Область $OABCD$ — область решения задачи.

Для нахождения максимального значения P проверим граничные точки из области решений.

Построим две линии уровня:

$$\begin{aligned} 2x_1 + 3x_2 &= 6; \\ 2x_1 + 3x_2 &= 12. \end{aligned}$$

Целевая функция f возрастает в направлении вектора-нормали $\vec{n} = (2; 3)$, следовательно, минимум находится в точке $(0; 0)$.

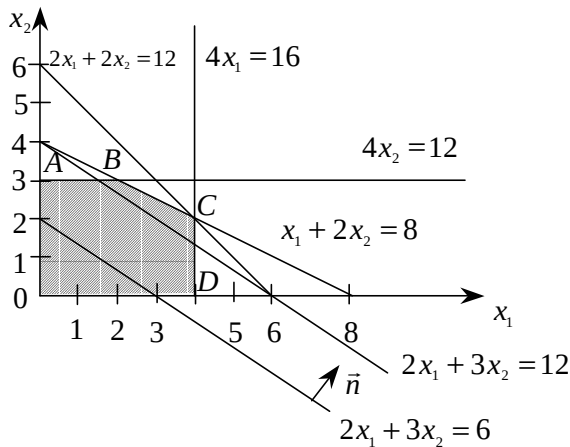


Рис. 3.1

Максимум определяем, передвигая линию уровня в направлении вектора $\vec{n}$ параллельно самой себе до тех пор, пока хотя бы одна ее точка будет принадлежать области допустимых решений.

В данном случае это точка $x_1 = 4, x_2 = 2$; при этом $P = 2 \cdot 4 + 3 \cdot 2 = 14$.

Таким образом, для получения максимальной прибыли в размере 14 усл. ед. надо продать 4 изделия первого вида и 2 изделия второго вида. ►

Графический метод применяется для решения задач линейного программирования следующего вида:

$$f = c_1 x_1 + c_2 x_2 \rightarrow \max (\min), \quad (3.11)$$

$$\begin{cases} a_{i1} x_1 + a_{i2} x_2 \leq b_i, & i = \overline{1, m_1}, \\ a_{i1} x_1 + a_{i2} x_2 \geq b_i, & i = \overline{m_1 + 1, m_2}. \end{cases} \quad (3.12)$$

Алгоритм решения ЗЛП графическим методом:

1. Записывают уравнения прямых, соответствующих ограничениям (3.12), и построить их на плоскости $x_1 O x_2$.

2. Определяют области, в которых выполняются ограничения задачи. Для этого выбирают произвольную точку на плоскости $x_1 O x_2$ и подставляют ее координаты в первую часть одного из неравенств. Если неравенство верно, то искомая полуплоскость находится с той же стороны от прямой, что и точка; в противном случае искомая полуплоскость лежит с противоположной стороны от прямой. Эти действия последовательно выполняются для всех неравенств (3.12).

3. Определить область допустимых решений задачи как область пересечения m полуплоскостей, соответствующих ограничениям задачи.

4. Определить направление возрастания (убывания) целевой функции f . Это можно сделать двумя способами. Можно построить вектор-нормаль $\vec{n} = (c_1, c_2)$: его направление показывает направление возрастания функции f ; в противоположном направлении функция убывает. Можно построить две линии уровня функции $f = K_1, f = K_2$ (K_1, K_2 — произвольные константы, $K_1 \neq K_2$) и по их расположению определить направление возрастания (убывания) функции.

5. Определить граничную точку (точки) области допустимых решений, в которых целевая функция принимает максимальное или минимальное значение.

6. Вычислить значение найденной точки, решив совместно уравнения, задающие прямые, на пересечении которых находится эта точка, или выявив уравнение граничной прямой области допустимых решений, с которой совпадает линия уровня целевой функции.

Возможны следующие варианты области допустимых решений (ОДР) (рис. 3.2).

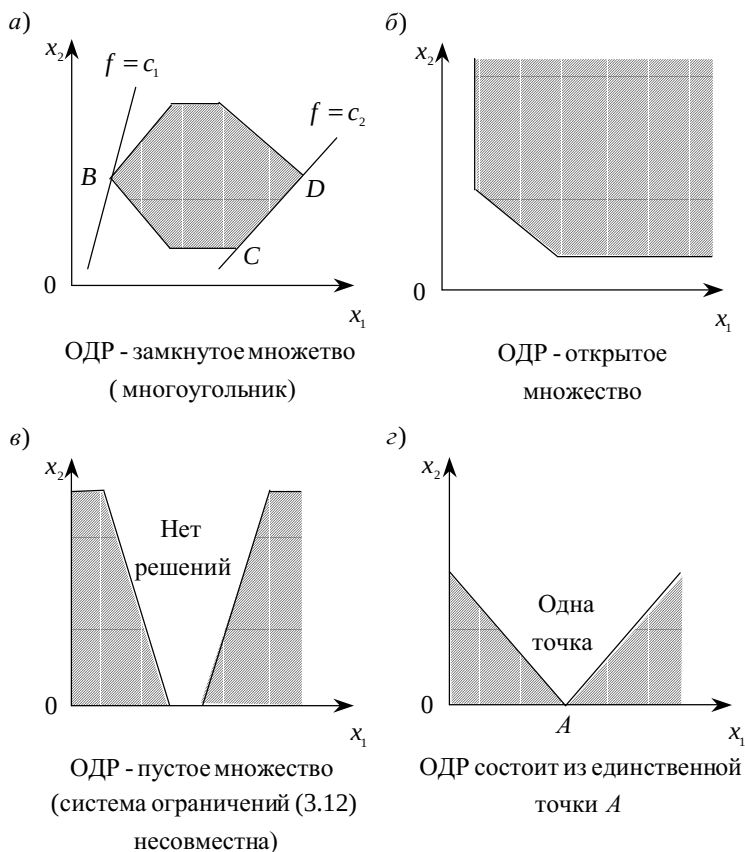


Рис. 3.2

Может быть единственное решение — точка B , бесконечно много решений — отрезок CD (см. рис. 3.2, a), максимальным (минимальным) значением целевой функции может быть ∞ (см. рис. 3.2, b), когда область допустимых решений — открытое множество. Возможно также, что решения не существует (см. рис. 3.2, $в$) или область допустимых решений состоит из единственной точки (см. рис. 3.2, $г$).

§ 3. Примеры построения экономических моделей

В этом параграфе мы покажем, как некоторые экономические задачи сводятся к задачам линейного программирования, т.е. продемонстрируем процесс математического моделирования в некоторых экономических системах. Причем, мы не ставим здесь целью решение этих задач, а лишь сводим их до необходимой математической постановки. Методы их решения на других примерах изложим позже.

1. Рекламная деятельность. Пусть предприятие рекламирует свою продукцию с использованием четырех источников массовой информации: телевидения, радио, газет и расклейки объявлений. Анализ рекламной деятельности в прошлом показал, что эти средства приводят к увеличению прибыли соответственно на 10, 5, 7 и 4 усл. ед. в расчете на 1 усл. ед., потраченную на рекламу. На рекламу выделено 50 000 усл. ед. Администрация предприятия не намерена тратить на телевидение более 40%, а на радио и газеты — более 50% от общей суммы выделенных средств. Как предприятие должно организовать рекламу, чтобы получить максимальную прибыль?

Построим математическую модель задачи.

1. Цель — максимизация прибыли.
2. Параметрами являются все числа, приведенные в условии задачи.
3. Управляющие переменные:
 - x_1 — количество средств, вложенных в рекламу на телевидение;
 - x_2 — количество средств, вложенных в рекламу на радио;
 - x_3 — количество средств, вложенных в рекламу в газетах;
 - x_4 — количество средств, вложенных в рекламу, организованную с помощью расклейки объявлений.

4. Область допустимых решений имеет вид:

$$\begin{cases} x_1 + x_2 + x_3 + x_4 \leq 50000, \\ x_1 \leq 20000, \\ x_2 + x_3 \leq 25000, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{cases} \quad (3.13)$$

Она содержит ограничения по общей сумме выделенных средств, предусмотренных на рекламу по телевидению, на радио и в газетах, и условия неотрицательности переменных.

5. Критерий оптимальности записывается следующим образом:

$$P = 10x_1 + 5x_2 + 7x_3 + 4x_4 \rightarrow \max. \quad (3.14)$$

Формулы (3.13), (3.14) — математическая модель задачи организации рекламной деятельности.

Целевая функция и ограничения линейны по управляющим переменным, следовательно, это задача линейного программирования.

Приведем задачу к каноническому виду, добавив дополнительные слагаемые к левым частям ограничений. Получим

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 = 50000, \\ x_1 + x_6 = 20000, \\ x_2 + x_3 + x_7 = 25000, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{cases} \quad (3.15)$$

Задача (3.14), (3.15) может быть решена симплекс-методом (см. § 5 главы 3). Нам же был интересен сам процесс построения этой модели.

2. Выбор программ капитальных вложений в условиях ограниченности ресурсов. Мы уже ввели некоторые понятия из области финансового анализа (см. подпараграф 1.1 главы 2). Сейчас нам нужно ввести еще несколько понятий. Позже мы покажем, почему может быть полезна *концепция приведенной стоимости* (см. под-

140

параграф 23.4 главы 4, подпараграфы 1.4, 6.1 главы 5). Однако эту концепцию легче воспринять, если понять интуитивно.

Итак, предположим, ваш многоквартирный дом сгорел и вы остались с пустым участком земли стоимостью 50 000 усл. ед. и чеком от страховой компании на сумму 200 000 усл. ед. Вы хотите восстановить здание, но ваш консультант по вопросам недвижимого имущества предлагает вам построить взамен старого дома офисное здание. Затраты на строительство могут составить 300 000 усл. ед., кроме того, нужно учесть еще и стоимость земли, которая в ином случае может быть продана за 50 000 усл. ед. С другой стороны, ваш консультант предполагает, что в будущем ожидается недостаток офисных помещений, и предсказывает, что через год продажа нового здания может принести вам 400 000 усл. ед. Таким образом, вы можете инвестировать сегодня 350 000 усл. ед. в ожидании получить 400 000 усл. ед. в следующем году. Вы должны принять это предложение, если приведенная стоимость ожидаемого дохода, равного 400 000 усл. ед., превышает сумму инвестиций в 350 000 усл. ед. Следовательно, вы должны задаться вопросом: “Сколько будут стоить 400 000 усл. ед. через год и превосходит ли их *приведенная стоимость* сумму 350 000 усл. ед.?”

Вычисление приведенной стоимости. Приведенная стоимость 400 000 усл. ед. через год должна быть меньше нынешних 400 000 усл. ед. *Ведь сегодняшний рубль стоит больше, чем завтрашний*, поскольку сегодня рубль можно инвестировать, и он немедленно начнет приносить доход в виде процента. Это первый основной принцип теории финансов.

Таким образом, приведенная стоимость отсроченного дохода может быть определена умножением суммы дохода на коэффициент дисконтирования, значение которого меньше 1. (Если бы коэффициент был больше 1, рубль сегодня стоил бы меньше, чем завтра). Если через C_1 обозначить ожидаемые доходы в период 1 (год спустя), то:

Приведенная стоимость (PV) = коэффициент дисконтирования $\times C_1$.

Коэффициент дисконтирования равен единице, деленной на сумму единицы и нормы доходности:

$$\text{Коэффициент дисконтирования} = \frac{1}{1+r}.$$

Норма доходности r представляет собой вознаграждение, которое требует инвестор за отсрочку поступления платежей.

Давайте для нашего примера инвестирования в недвижимость сделаем допущение, что доход в сумме 400 000 усл. ед. гарантирован. Офисное здание — не единственный способ получить через год 400 000 усл. ед. Вы можете инвестировать ваши средства в ценные бумаги Правительства России (ГКО) сроком на 1 год. Предположим, что по этим ценным бумагам выплачивается 7% годовых. Сколько вы должны инвестировать в ценные бумаги, чтобы получить 400 000 усл. ед. к концу года? Это легко вычислить: вы должны инвестировать $400\,000/1,07$, что составляет 373 832 усл. ед. Следовательно, при ставке 7% приведенная стоимость 400 000 усл. ед., которые будут получены через год, составляет 373 832 усл. ед.

Давайте допустим также, что, как только вы задействовали земельный участок и начали строительство здания, вы решаете продать ваш проект.

За сколько вы можете его продать? На этот вопрос также легко ответить. Так как данный объект собственности принесет 400 000 усл. ед., инвестор может захотеть приобрести его за 373 832 усл. ед.. Это столько же, сколько вы можете израсходовать, чтобы получить 400 000 усл. ед. дохода, инвестируя в правительственные ценные бумаги. Конечно, вы всегда можете продать свою недвижимость дешевле, но зачем продавать дешевле, чем может предложить рынок? Приведенная стоимость 373 832 усл. ед. представляет собой всего лишь возможную цену, которая удовлетворяет и покупателя, и продавца, поэтому приведенная стоимость недвижимости также является и ее рыночной ценой.

Чтобы вычислить приведенную стоимость, мы дисконтировали ожидаемый будущий доход по норме доходности, которую дают сравнимые альтернативные инвестиции. Эту норму доходности часто называют *ставкой дисконта*, предельной нормой доходности или альтернативными издержками капитала. Ее называют *альтернативными издержками*, поскольку она представляет

собой доход, от которого отказывается инвестор, вкладывая деньги в какой-либо проект, а не в ценные бумаги. В нашем примере альтернативные издержки равнялись 7%. Приведенная стоимость была получена делением 400 000 усл. ед. на 1,07:

$$PV = \text{Коэффициент дисконтирования} \cdot C_1 = \\ = \frac{1}{1+r} \cdot C_1 = \frac{400\,000}{1,07} = 37\,832 \text{ усл. ед.}$$

Чистая приведенная стоимость. Стоимость здания равна 373 832 усл. ед., но это не означает, что вы стали богаче на 373 832 усл. ед. Вы уже вложили 350 000 усл. ед., и, следовательно, чистая приведенная стоимость для вас составляет 23 832 дол. Чистая приведенная стоимость определяется вычитанием из приведенной стоимости требуемых инвестиций:

Чистая приведенная стоимость = приведенная стоимость — требуемые инвестиции = 373 832 – 350 000 = 23 832 усл. ед.

Другими словами, стоимость проекта строительства вашего офисного здания превышает затраты на его осуществление, т.е. он приносит *чистый* прирост стоимости. Формула для расчета чистой приведенной стоимости может быть записана следующим образом:

$$NPV = C_0 + \frac{C_1}{1+r}.$$

При этом напомним, что C_0 — поток денежных средств в период 0 (т. е. сегодня) — обычно является отрицательным числом. Иначе говоря, C_0 — это инвестиции и, следовательно, *отток* денежных средств. В нашем примере $C_0 = -350\,000$ усл. ед.

Замечание о риске и приведенной стоимости. В примере со строительством офисного здания мы сделали одно нереалистичное допущение: на самом деле ваш консультант по вопросам недвижимости не может быть *совершенно уверен* в будущей стоимости офисных зданий. Сумма 400 000 усл. ед. воплощает лучший *прогноз*, но она не гарантирована.

Следовательно, наше заключение о том, сколько в состоянии заплатить за здание инвесторы, является ошибочным. Так как они

определенно могут получить 400 000 усл. ед., купив на 373 832 усл. ед. ценные бумаги ГКО, они не станут покупать ваше здание за такую сумму. Вероятно, вы должны снизить цену, чтобы заинтересовать инвесторов.

Сейчас мы подошли ко второму основному принципу финансов: *надежный рубль стоит больше, чем рисковый*. Большинство инвесторов избегают риска, когда они могут это сделать, не жертвуя доходом. Тем не менее концепции приведенной стоимости и альтернативных издержек имеют значение и для рискованных инвестиций. Все же стоит дисконтировать вырученную сумму по норме доходности сопоставимых инвестиционных проектов. Однако мы должны учитывать ожидаемые поступления и *ожидаемые* нормы доходности других инвестиций.

Не все инвестиции сопряжены с одинаковой степенью риска. Проект строительства офисного здания рискованнее инвестирования в правительственные ценные бумаги, но, вероятно, несет в себе меньший риск, чем бурение наугад нефтяной скважины. Допустим, вы полагаете, что проект связан с таким же риском, как и инвестиции на фондовом рынке, и вы прогнозируете для них норму доходности, равную 12%. Тогда 12% станут соответствующими альтернативными издержками.

Вот от чего вы отказываетесь, не инвестируя в сопоставимые с вашим проектом ценные бумаги. Теперь вы можете пересчитать чистую приведенную стоимость:

$$PV = \frac{400\,000}{1,12} = 357\,143 \text{ усл. ед.}$$

$$NPV = PV - 350\,000 = 7143 \text{ усл. ед.}$$

Если другие инвесторы согласны с вашим прогнозом дохода в 400 000 усл. ед. и с вашей оценкой альтернативных издержек в размере 12%, тогда ваша недвижимость в начале строительства должна стоить 357 143 усл. ед. Если бы вы попытались продать ее дороже, вы не нашли бы покупателя, поскольку тогда ожидаемая норма доходности недвижимости оказалась бы ниже чем 12%, которые можно получить на фондовом рынке. Офисное здание

по-прежнему дает чистый прирост стоимости, но он гораздо меньше, чем показывают наши предыдущие расчеты.

Финансовый менеджер должен учитывать факторы времени и неопределенности и их влияние на стоимость. Это очевидно в нашем примере. Доход в размере 400 000 усл. ед. стоил бы именно столько, если бы мог быть получен немедленно. Если сооружение офисного здания настолько же надежно, как и правительственные ценные бумаги, задержка на 1 год снижает стоимость до 373 832 усл. ед. Если же оно сопряжено с таким же риском, как и инвестиции на фондовом рынке, неопределенность снижает стоимость еще на 16 689 усл. ед., до 357 143 усл. ед.

К сожалению, оценка стоимости активов с учетом факторов времени и неопределенности часто является более сложной задачей, чем предполагает наш пример. Мы вернемся к обсуждению этого вопроса в § 6 главы 5.

Приведенные стоимости и нормы доходности. Мы решили, что строительство офисного здания — стоящее дело, так как его стоимость превышает связанные с ним издержки — т.е. оно имеет положительную чистую приведенную стоимость. Чтобы вычислить стоимость, мы посмотрели, сколько потребуется потратить, чтобы получить доход, равный доходу от инвестирования непосредственно в ценные бумаги. Приведенная стоимость проекта равна его будущей доходности, дисконтированной по норме доходности этих ценных бумаг.

Мы можем уточнить наш критерий, сказав, что наш проект с недвижимостью стоит осуществить, потому что доход превышает издержки. Доходность инвестированного капитала представляет собой просто отношение прибыли к первоначальным расходам:

$$\text{Доходность} = \frac{\text{прибыль}}{\text{инвестиции}} = \frac{400000 - 350000}{350000} = 14\%.$$

Издержки инвестирования опять же представляют собой доход, который был упущен из-за отказа инвестировать в ценные бумаги. В нашем примере, если строительство офисного здания сопряжено с таким риском, как и инвестиции на фондовом рынке,

то упущенный доход равен 12%. Так как 14% доходности от офисного здания превышает 12% альтернативных издержек, нам следует начать возведение фундамента здания.

Здесь мы приведем два равноценных правила, которых необходимо придерживаться при принятии инвестиционных решений:

1. *Правило чистой приведенной стоимости:* осуществлять инвестиции, если они имеют положительную чистую приведенную стоимость.

2. *Правило нормы доходности:* осуществлять инвестиции, норма доходности которых превышает их альтернативные издержки.

Модель выбора программ капитальных вложений в условиях ограниченности ресурсов. При обзоре методов планирования капитальных вложений иногда основываются на допущении, в соответствии с которым богатство акционеров фирмы максимально возрастает, когда она принимает каждый проект, имеющий положительную чистую приведенную стоимость. Однако предположим, что существуют ограничения на осуществление инвестиционных программ, которые не позволяют фирме принять все подобные проекты. В этом случае нам необходим метод отбора группы проектов, который с учетом ограниченного объема ресурсов компании обеспечит максимально возможную чистую приведенную стоимость.

Давайте начнем с очень простого примера. Предположим, что альтернативные издержки инвестирования равны 10%, совокупные ресурсы нашей компании составляют 10 млн. усл. ед., и что перед ней открыты следующие возможности:

Проект	Потоки денежных средств (в млн. дол.)			Чистая приведенная стоимость при $r = 10\%$ (в млн. усл. ед.)	Коэффициент рентабельности
	C_0	C_1	C_2		
<i>A</i>	-10	+30	+5	21	3,1
<i>B</i>	-5	+5	+20	16	4,2
<i>B</i>	-5	+5	+15	12	3,4

У фирмы достаточно ресурсов, чтобы инвестировать либо в проект *A*, либо в проекты *B* и *B*. Хотя чистые приведенные стоимости каждого из проектов *B* и *B* меньше, чем у проекта *A*. Чис-

тая приведенная стоимость этих проектов, взятых вместе, выше. Поэтому ясно, что мы не можем делать выбор исключительно на основе чистых приведенных стоимостей отдельных проектов. Когда средства ограничены, нам нужно получить “наибольшее удовольствие за свои денежки”. Иначе говоря, мы должны выбрать проекты, для которых отношение приведенной стоимости к первоначальным инвестициям является наиболее высоким. Это отношение представляет собой просто коэффициент рентабельности, или коэффициент выгоды-издержки:

$$\text{Коэффициент рентабельности} = \frac{\text{приведенная стоимость}}{\text{инвестиции}}.$$

Среди наших проектов *Б* имеет наибольший коэффициент рентабельности, а проекту *В* принадлежит следующий по величине коэффициент. Следовательно, если наш капитальный бюджет ограничен 10 млн. усл.ед, нам следует принять два этих проекта.

К сожалению, есть некоторые ограничения в применении этих простых методов оценки проектов. Одно из наиболее серьезных ограничений заключается в том, что они не подходят тогда, когда нормируется более чем один вид ресурсов. Например предположим, что бюджетное ограничение в 10 млн. усл. ед. касается потока денежных средств для года 0 и года 1, и что наш выбор расширяется следующим образом:

Проект	Потоки денежных средств (в млн. усл. ед.)			Чистая приведенная стоимость при $r = 10\%$ (в млн. усл. ед.)	Коэффициент рентабельности
	C_0	C_1	C_2		
<i>А</i>	−10	+30	+5	21	3,1
<i>Б</i>	−5	+5	+20	16	4,2
<i>В</i>	−5	+5	+15	12	3,4
<i>Г</i>	0	−40	+60	13	1,4

Одна из стратегий — принять проекты *Б* и *В*; однако в этом случае мы не сможем также принять проект *Г*, затраты по которому превышают наше бюджетное ограничение для 1-го периода. Альтернативный вариант — принять проект *А* в период 0. Хотя

он имеет меньшую чистую приведенную стоимость, чем комбинация проектов *Б* и *В*, он обеспечивает положительный поток денежных средств в размере 30 млн. усл. ед. в 1-й период. Если мы добавим 30 млн. усл. ед. к нашему бюджету в 10 млн. усл. ед., мы можем позволить себе принять и проект *Г*. Проекты *А* и *Г* имеют *меньшие* коэффициенты рентабельности, чем проекты *Б* и *В*, но *большую* совокупную чистую приведенную стоимость.

Причина, по которой метод выбора по коэффициенту рентабельности не сработал в нашем примере, заключается в том, что ресурсы ограничены в *каждом* из двух периодов. В действительности этот метод не подходит всегда, когда имеется *любое дополнительное* ограничение при выборе проектов. Это значит, что он не подходит для случаев, когда два проекта являются взаимоисключающими или когда один проект зависит от другого.

Несколько усовершенствованных моделей выбора в условиях нормирования капитала. Простота метода выбора по коэффициенту рентабельности иногда компенсирует ограниченность его применения. Например, нет необходимости принимать во внимание затраты в последующие годы, если вы не имеете четкого представления о доступности капитала или об инвестиционных возможностях в будущем. Но есть условия, при которых ограниченность метода отбора по коэффициенту рентабельности делает его применение недопустимым. В таких ситуациях нам необходим более общий метод отбора проектов в условиях нормирования капитала.

Начнем знакомиться с проблемой с ее описанием. Предположим, что мы обозначили через x_A долю проекта *А* в нашем примере. Тогда чистая приведенная стоимость инвестиций в этот проект составила бы $21x_A$. Аналогично, чистая приведенная стоимость наших инвестиций в проект *Б* может быть выражена как $16x_B$ и т.д. Наша цель состоит в выборе группы проектов с наибольшей *совокупной* чистой приведенной стоимостью. Другими словами, мы хотим найти значение, при котором максимизируется функция

$$NPV = 21x_A + 16x_B + 12x_C + 13x_D.$$

На наш выбор накладываются некоторые ограничения. Во-первых, совокупный отток денежных средств в период 0 не дол-

жен превышать 10 млн. усл. ед. Иначе говоря: $10x_A + 5x_B + 5x_B + 0 \cdot x_G \leq 10$.

Точно также совокупный отток денежных средств в 1-й период не должен быть больше 10 млн. усл. ед.:

$$-30x_A - 5x_B - 5x_B + 40x_G \leq 10.$$

И, наконец, наши инвестиции в проект не могут иметь отрицательное значение, и мы не можем предпринять более чем один из них. Следовательно, мы имеем:

$$0 \leq x_A \leq 1, \quad 0 \leq x_B \leq 1 \dots$$

Учитывая все эти условия, мы можем представить задачу следующим образом:

$$NPV = 21x_A + 16x_B + 12x_B + 13x_G \rightarrow \max,$$

$$\begin{cases} -30x_A - 5x_B - 5x_B + 40x_G \leq 10, \\ 10x_A + 5x_B + 5x_B + 0x_G \leq 10, \\ 0 \leq x_A \leq 1, \quad 0 \leq x_B \leq 1 \dots \end{cases}$$

Одним из способов решения такой задачи служит метод подстановки различных значений x с выделением тех комбинаций, которые одновременно и удовлетворяют ограничениям, и дают наибольшую чистую приведенную стоимость. Однако разумнее признать, что представленные выше выражения составляют задачу линейного программирования (ЛП). Ее можно решить с помощью компьютера, оснащенного программой для решения задач ЛП, либо одним из методов, предложенных ниже.

(3.18)

Перед нами возникает задача линейного программирования в следующей постановке: найти такие неотрицательные значения $n + m$ переменных $x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_m$, чтобы они удовлетворяли системе уравнений (3.18) и одновременно обращали в минимум линейную функцию этих переменных

$$L = c_1x_1 + c_2x_2 + \dots + c_nx_n.$$

Общее количество переменных равно $n + m$, из них n первоначальных и m добавочных. Функция L выражена только через первоначальные переменные (коэффициенты при добавочных переменных в ней равны нулю).

▷ П Р И М Е Р 3.2. Имеется задача линейного программирования с ограничениями-неравенствами:

(3.19)

Требуется привести эту задачу к виду ОЗЛП.

Решение. Приводим неравенства (3.19) к стандартной форме:

$$\begin{cases} -2x_1 + x_2 - 3x_3 + 6 \geq 0, \\ 3x_2 - x_3 - 1 \geq 0, \\ x_1 - 2x_4 + x_5 + 1 \geq 0, \\ x_1 - x_5 \geq 0. \end{cases} \quad (3.20)$$

Вводим добавочные переменные:

$$\begin{cases} y_1 = -2x_1 + x_2 - 3x_3 + 6, \\ y_2 = 3x_2 - x_3 - 1, \\ y_3 = x_1 - 2x_4 + x_5 + 1, \\ y_4 = x_1 - x_5. \end{cases} \quad (3.21)$$

Задача сводится к тому, чтобы найти неотрицательные значения переменных $x_1, x_2, x_3, x_4, x_5; y_1, y_2, y_3, y_4$, удовлетворяющие уравнениям (3.21). ►

Мы показали, как от задачи линейного программирования с ограничениями-неравенствами можно перейти к задаче с ограничениями-равенствами (т.е. ОЗЛП). Всегда возможен и обратный переход от ОЗЛП к задаче с ограничениями-неравенствами. Если в первом случае мы увеличивали число переменных, то во втором случае будем его уменьшать, устраняя базисные переменные и оставляя только свободные (см. § 4 главы 1).

► П Р И М Е Р 3.3. Имеется задача линейного программирования с ограничениями-равенствами (ОЗЛП)

$$\begin{cases} x_1 + x_2 = 1, \\ x_2 - 2x_3 = -3, \\ x_3 - x_4 + x_5 = 1 \end{cases} \quad (3.22)$$

и минимизируемой функцией

$$L = -x_1 - x_2 + x_5. \quad (3.23)$$

Требуется записать ее как задачу линейного программирования с ограничениями-неравенствами и построить ее решение.

Решение. Так как $m = 3$, $n = 5$, $n - m = 2$, выберем какие-то две из переменных в качестве свободных. Заметим, что переменные x_1, x_2 в качестве свободных выбирать нельзя, так как они связаны первым из уравнений (3.22): значение одной из них полностью определяется значением другой, а свободные переменные должны быть независимыми. По такой же причине нельзя в качестве свободных выбрать переменные x_2, x_3 [их связывает второе уравнение (3.22)]. Выберем в качестве свободных переменные x_1 и x_4 и выразим через них все остальные:

$$\begin{cases} x_2 = -x_1 + 1, \\ x_3 = -\frac{1}{2}x_1 + 2, \\ x_5 = \frac{1}{2}x_1 + x_4 - 1. \end{cases} \quad (3.24)$$

Так как $x_2 \geq 0$, $x_3 \geq 0$, $x_5 \geq 0$, условия (3.24) могут быть заменены неравенствами:

$$\begin{cases} -x_1 + 1 \geq 0, \\ -\frac{1}{2}x_1 + 2 \geq 0, \\ \frac{1}{2}x_1 + x_4 - 1 \geq 0. \end{cases} \quad (3.25)$$

Перейдем в выражении линейной функции L к свободным переменным x_1, x_4 . Подставив в L вместо x_2 и x_5 их выражения (3.24), получим

$$\begin{aligned} L &= -x_1 + x_1 - 1 + \frac{1}{2}x_1 + x_4 - 1 = \frac{1}{2}x_1 + x_4 - 2, \\ L' &= \frac{1}{2}x_1 + x_4. \end{aligned} \quad (3.26)$$

Таким образом, задача сведена к задаче линейного программирования с ограничениями-неравенствами. Ее геометрическая интерпретация показана на рис. 3.3. Основная прямая $L' = 0$ параллельна той стороне ОДР, где L' достигает минимума. Следова-

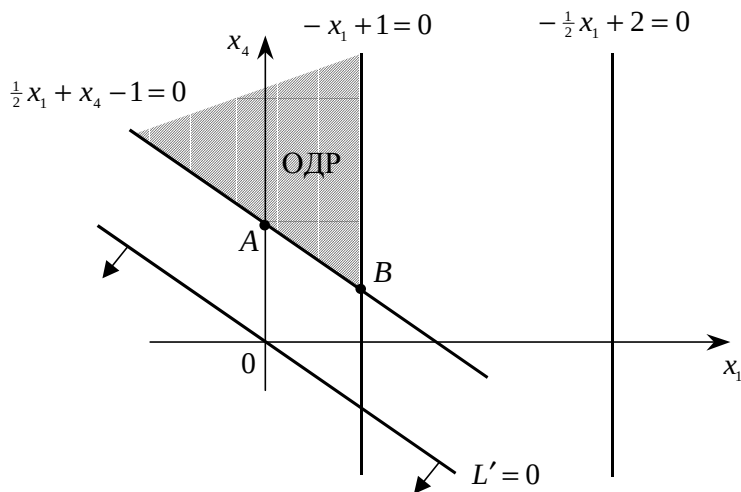


Рис. 3.3

вательно, все точки участка AB дают оптимальное решение. Выбрав в качестве решения, например, координаты точки A , получим $x_1^* = 0$; $x_4^* = 1$; $x_2^* = 1$; $x_3^* = 2$; $x_5^* = 0$.

При таких значениях переменных линейная функция L достигает минимума, равного

$$L_{\min} = \frac{1}{2} \cdot 0 + 1 - 2 = -1. \blacktriangleright$$

Таким образом, мы можем произвольно переходить от ОЗЛП к задаче линейного программирования с ограничениями-неравенствами и наоборот. Если в числе ограничений задачи есть как уравнения, так и неравенства, рекомендуется произвести унификацию, т.е. перейти к какой-либо единообразной форме, например к ОЗЛП.

▷ПРИМЕР 3.4. Имеется задача линейного программирования с переменными x_1, x_2, x_3, x_4 и ограничениями вида

$$\begin{cases} x_1 + x_2 = x_3 + x_4, \\ x_1 - x_2 + x_3 \leq 1, \\ x_2 + x_3 + x_4 \geq 5. \end{cases} \quad (3.27)$$

Минимизируется функция

$$L = x_1 - 2x_2 + x_3 - 3x_4. \quad (3.28)$$

Требуется привести задачу к ОЗЛП.

Р е ш е н и е. Введением добавочных переменных y_1, y_2 приведем условия (3.27) к виду ОЗЛП:

$$\begin{cases} x_1 + x_2 - x_3 - x_4 = 0, \\ x_1 - x_2 + x_3 + y_1 = 1, \\ x_2 + x_3 + x_4 - y_2 = 5. \end{cases}$$

Минимизируемая функция остается в виде (3.28). ►

§5. Симплекс-метод решения задачи линейного программирования

Геометрическая интерпретация, которой мы пользовались при решении задач линейного программирования, перестает быть пригодной для этой цели при числе свободных переменных $n - m > 3$, а затруднительна уже при $n - m = 3$. Для нахождения решения задачи линейного программирования в общем случае (при п р о и з в о л ь н о м числе свободных переменных) применяются не геометрические, а вычислительные методы. Из них наиболее универсальным является так называемый *симплекс-метод*.

Идея симплекс-метода относительно проста. Пусть в задаче линейного программирования имеется n переменных и m независимых линейных ограничений, заданных в форме уравнений. Как известно, оптимальное решение (если оно существует) достигается в одной из опорных точек (вершин ОДР), где по крайней мере $k = n - m$ переменных равны нулю. Выберем какие-то k переменных в качестве свободных и выразим через них остальные m базисных переменных. Пусть в качестве свободных выбраны, например, первые $k = n - m$ переменных $x_1, x_2, \dots, x_k$, а остальные m выражены через них:

[illegible]

Если положить все свободные переменные $x_1, x_2, \dots, x_k$ равными нулю, т.е.

$$x_1 = 0, x_2 = 0, \dots, x_k = 0,$$

получим

$$x_{k+1} = \beta_{k+1}, x_{k+2} = \beta_{k+2}, \dots, x_n = \beta_n.$$

Это решение может быть допустимым или недопустимым. Оно *допустимо*, если все свободные члены $\beta_{k+1}, \beta_{k+2}, \dots, \beta_n$ *неотрицательны*. Предположим, что это условие выполнено. Тогда мы получили *опорное решение*. Но является ли оно оптимальным?

Может быть, да, а может быть, и нет. Чтобы проверить это, выразим минимизируемую линейную функцию L через свободные переменные $x_1, x_2, \dots, x_k$:

$$L = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \dots + \gamma_k x_k. \quad (3.30)$$

Очевидно, что при $x_1 = x_2 = \dots = x_k = 0$ функция $L = \gamma_0$. Посмотрим, нельзя ли улучшить решение, т.е. уменьшить функцию L , увеличивая какие-нибудь из переменных $x_1, x_2, \dots, x_k$ (уменьшать их мы не можем, так как все они равны нулю, а отрицательные значения переменных недопустимы). Если все коэффициенты $\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_k$ в формуле (3.30) *положительны*, то, увеличивая какие-то из переменных $x_1, x_2, \dots, x_k$ сверх нуля, мы не можем уменьшить L ; следовательно, найденное нами опорное решение является *оптимальным*. Если же среди коэффициентов $\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_k$ в формуле (3.30) есть отрицательные, то, увеличивая те переменные $x_1, x_2, \dots, x_k$, коэффициенты при которых отрицательны, мы можем улучшить решение, т.е. уменьшить L .

Пусть, например, коэффициент γ_1 в формуле (3.30) отрицателен. Значит, есть смысл увеличить x_1 , т.е. перейти от данного опорного решения к другому, где переменная x_1 не равна нулю, а вместо нее равна нулю какая-то другая. Увеличение x_1 “полезно” для линейной функции L , оно делает ее меньше. Однако увеличивать x_1 надо осторожно, так, чтобы не стали отрицательными другие переменные $x_{k+1}, x_{k+2}, \dots, x_n$, выраженные через свободные переменные, в частности через x_1 , формулами (3.29).

Посмотрим, “опасно” ли для переменных $x_{k+1}, x_{k+2}, \dots, x_n$ увеличение x_1 , т.е. может ли оно сделать их отрицательными? Да, “опасно”, если коэффициент при x_1 в соответствующем уравнении отрицателен. Если среди уравнений (3.29) нет уравнения с отрицательным коэффициентом при x_1 , то величину x_1 можно увеличивать беспрестанно, а значит, линейная функция L не ограничена снизу и оптимального решения ОЗЛП не существует.

Допустим, что это не так и что среди уравнений (3.29) есть такие, в которых коэффициент при x_1 отрицателен. Для переменных, стоящих в левых частях этих уравнений, увеличение x_1 “опасно” — оно может сделать их отрицательными.

Возьмем одну из таких переменных x_l и посмотрим, до какой степени можно все же увеличить x_1 , пока переменная x_l не станет отрицательной. Рассмотрим l -е уравнение из системы (3.29):

$$x_l = \alpha_{l1}x_1 + \alpha_{l2}x_2 + \dots + \alpha_{lk}x_k + \beta_l.$$

Здесь свободный член $\beta_l \geq 0$, а коэффициент α_{l1} отрицателен. Легко понять, что если оставить $x_2 = \dots = x_k = 0$, то x_1 можно увеличивать только до значения, равного $-\beta_l / \alpha_{l1}$, а при дальнейшем увеличении x_1 переменная x_l станет отрицательной.

Выберем ту из переменных $x_{k+1}, x_{k+2}, \dots, x_n$, которая раньше всех обратится в нуль при увеличении x_1 , т.е. ту, для которой величина $-\beta_l / \alpha_{l1}$ меньше остальных. Пусть такой переменной будет x_r . Тогда имеет смысл “переразрешить” систему уравнений (3.29) относительно других базисных переменных, выведя из числа свободных переменных x_1 и переведя вместо нее в группу свободных переменных x_r . Действительно, мы хотим перейти от опорного решения, задаваемого равенствами $x_1 = x_2 = \dots = x_k = 0$, к опорному решению, в котором уже $x_1 \neq 0$, а $x_2 = \dots = x_k = x_r = 0$. Первое опорное решение мы получили, положив равными нулю все прежние свободные переменные $x_1, x_2, \dots, x_k$; второе мы получим, если обратим в нуль все новые свободные переменные $x_2, \dots, x_k, x_r$. Базисными переменными при этом будут переменные $x_1, x_{k+1}, \dots, x_{r-1}, x_{r+1}, \dots, x_n$.

Предположим, что уравнения типа (3.29) для нового набора базисных и свободных переменных составлены. Тогда можно выразить через новые свободные переменные и линейную функцию L . Если все коэффициенты при переменных в этой формуле положительны, то мы нашли оптимальное решение: оно получится, если все свободные переменные положить равными нулю. Если среди коэффициентов при переменных есть отрицательные, то процедура улучшения решения продолжается: система вновь “переразрешается” относительно других базисных переменных, и так далее, пока не будет найдено оптимальное решение, обращающее функцию L в минимум.

Проследим описанную процедуру постепенного улучшения решения ОЗЛП на конкретном примере.

▷ ПРИМЕР 3.5. Имеется задача линейного программирования с ограничениями-неравенствами:

$$\begin{cases} -5x_1 - x_2 + 2x_3 \leq 2, \\ -x_1 + x_3 + x_4 \leq 5, \\ -3x_1 + 5x_4 \leq 7. \end{cases} \quad (3.31)$$

Требуется минимизировать линейную функцию

$$L = 5x_1 - 2x_3.$$

Решение. Приводя неравенства к стандартному виду (≥ 0) и вводя добавочные переменные y_1, y_2, y_3 , переходим к условиям-равенствам:

$$\begin{cases} y_1 = 5x_1 + x_2 - 2x_3 + 2, \\ y_2 = x_1 - x_3 - x_4 + 5, \\ y_3 = 3x_1 - 5x_4 + 7. \end{cases} \quad (3.32)$$

Число переменных $n = 7$ на 4 превышает число уравнений $m = 3$. Значит, четыре переменные могут быть выбраны в качестве свободных.

Выберем в качестве свободных переменных x_1, x_2, x_3, x_4 . Положив их равными нулю, сразу получим опорное решение: $x_1 = x_2 = x_3 = x_4 = 0, y_1 = 2; y_2 = 5; y_3 = 7$.

При этих значениях переменных $L = 0$.

Является ли это решение оптимальным? Нет! Потому что в выражении линейной функции L коэффициент при x_3 отрицателен. Значит, увеличивая x_3 , можно уменьшить L .

Попробуем увеличивать x_3 . Проследим по уравнениям (3.32), “опасно” ли это для других переменных? Да, “опасно” для y_1 и y_2 — в их уравнения переменная x_3 входит с отрицательным коэффициентом, значит, при увеличении x_3 соответствующие переменные y_1 и y_2 могут стать отрицательными.

Посмотрим, какая из этих переменных — y_1 или y_2 раньше обратится в нуль при увеличении x_3 . Очевидно, y_1 : она станет равной нулю при $x_3 = 1$, а величина y_2 — только при $x_3 = 5$.

Поэтому выбираем переменную y_1 и вводим ее в число свободных вместо x_3 . Чтобы “переразрешить” систему (3.32) относительно x_3, y_2, y_3 , поступим следующим образом. Разрешим первое уравнение (3.32) относительно новой базисной переменной x_3 :

$$x_3 = \frac{5}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}y_1 + 1.$$

Это выражение подставим вместо x_3 во второе уравнение, получим

$$y_2 = -\frac{3}{2}x_1 - \frac{1}{2}x_2 + \frac{1}{2}y_1 - x_4 + 4.$$

Что же касается третьего уравнения, то оно, как не содержащее x_3 , не изменится. Итак, мы привели систему (3.32) к виду

$$\begin{cases} x_3 = \frac{5}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}y_1 + 1, \\ y_2 = -\frac{3}{2}x_1 - \frac{1}{2}x_2 + \frac{1}{2}y_1 - x_4 + 4, \\ y_3 = 3x_1 - 5x_4 + 7 \end{cases} \quad (3.33)$$

со свободными переменными x_1, x_2, y_1, x_4 и базисными x_3, y_2, y_3 .

Выразим линейную функцию L через новые свободные переменные:

$$L = 5x_1 - 5x_1 - x_2 + y_1 - 2,$$

или

$$L = -x_2 + y_1 - 2. \quad (3.34)$$

Положим теперь свободные переменные равными нулю. Линейная функция L станет равной -2 . Это уже лучше, чем прежнее значение $L = 0$. Но является ли это решение оптимальным? Все еще нет, так как коэффициент при x_2 в выражении (3.34) отрицателен. Итак, будем увеличивать x_2 . Посмотрим, для какой из переменных, стоящих в левых частях системы (3.33), это может быть “опасно”. Только для y_2 (в первое уравнение x_2 входит с положительным коэффициентом, а в третье совсем не входит).

Итак, поменяем местами переменные x_2 и y_2 : первую выведем из числа свободных, а вторую — введем. Для этого разрешим второе уравнение (3.33) относительно x_2 и подставим полученное выражение в первое уравнение. Получим еще один вид системы (3.32):

$$(3.35)$$

$$L = 3x_1 + 2y_2 - y_1 + 2x_4 - 8 + y_1 - 2,$$
$$(3.36)$$

в ОЗЛП должны быть записаны в *стандартной* форме:

$$(3.37)$$

4. Подчеркнуть (или выделить иным способом) в разрешающей строке все верхние числа (прежние элементы), за исключением самого разрешающего элемента клетки, а в разрешающем стол-

бце — все нижние числа (новые элементы), за исключением самого разрешающего элемента.

5. Для каждого элемента, не принадлежащего ни к разрешающей строке, ни к разрешающему столбцу, записать в нижнюю часть клетки *произведение выделенных чисел*, стоящих в том же столбце и в той же строке, что и данный элемент.

6. Переписать таблицу, заменив:

а) x_j на y_i и наоборот;

б) элементы разрешающей строки и столбца — числами, стоящими в *нижних* частях тех же клеток;

в) каждый из остальных элементов — суммой чисел, стоящих в *верхней* и *нижней* частях той же клетки.

▷ ПРИМЕР 3.6. В системе уравнений

$$\begin{cases} y_1 = x_1 - x_2 + 2x_3 - 5, \\ y_2 = 2x_1 - x_2 + 1, \\ y_3 = 2x_2 - x_3 - 1, \\ y_4 = -x_1 - x_3 + 2. \end{cases} \quad (3.38)$$

осуществить замену $x_1 \leftrightarrow y_2$, т.е. вывести из числа свободных переменных x_1 и вместо нее ввести y_2 .

Решение. Записываем уравнения (3.38) в форме стандартной таблицы (см. табл. 3.3), оставляя в нижней части каждой клетки свободное место.

Обводим кружком разрешающий элемент -2 , а также разрешающие строку и столбец (в таблице затенены). Вычисляем $\lambda = -\frac{1}{2}$. Вспомогательные записи будем вести в правом нижнем углу клетки (табл. 3.4).

Заполняем, согласно пунктам 1, 2 и 3 алгоритма, нижние части клеток разрешающих строки и столбца.

Выделяем, заключая их в рамку, верхние числа разрешающей строки и нижние числа разрешающего столбца (кроме самой разрешающей клетки).

Таблица 3.3

	Свободный член	y_2 $\updownarrow$ x_1	x_2	x_3
y_1	-5	-1	1	-2
$x_1 \leftrightarrow y_2$	1	(-2)	1	0
y_3	-1	0	-2	1
y_4	2	1	0	1

Таблица 3.4

	Свободный член	y_2 $\updownarrow$ x_1	x_2	x_3
y_1	-5	-1 $-\frac{1}{2}$	1	-2
$x_1 \leftrightarrow y_2$	1 $-\frac{1}{2}$	(-2) $-\frac{1}{2}$	1 $-\frac{1}{2}$	0 0
y_3	-1	0 0	-2	1
y_4	2	1 $\frac{1}{2}$	0	1

Далее заполняем все остальные нижние части клеток, перемножая соответствующие им выделенные числа, стоящие в разрешающей строке и разрешающем столбце на тех же местах, что и данная клетка (табл. 3.5).

Таблица 3.5

	Свободный член	y_2 $\updownarrow$ x_1	x_2	x_3
y_1	-5 $-\frac{1}{2}$	-1 $-\frac{1}{2}$	1 $-\frac{1}{2}$	-2 0
$x_1 \leftrightarrow y_2$	1 $-\frac{1}{2}$	(-2) $-\frac{1}{2}$	1 $-\frac{1}{2}$	0 0
y_3	-1 0	0 0	-2 0	1 0
y_4	2 $\frac{1}{2}$	1 $\frac{1}{2}$	0 $\frac{1}{2}$	1 0

Таблица 3.6

	С в о б о д н ы й ч л е н	y_2	x_2	x_3
y_1	$-\frac{11}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	-2
x_1	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	0
y_3	-1	0	-2	1
y_4	$\frac{5}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1

Заканчиваем преобразование, для чего переписываем табл. 3.5, заменяя x_1 на y_2 : элементы разрешающей строки и столбца — нижними числами тех же клеток, а остальные элементы — суммой верхних и нижних чисел (табл. 3.6). ►

Таким образом, мы научились с помощью табличного алгоритма совершать в уравнениях-ограничениях любую замену $x_j \leftrightarrow y_i$.

Вспомним, что в задаче линейного программирования кроме уравнений-ограничений существует еще и линейная функция

$$L = c_0 + c_1x_1 + c_2x_2 + \dots + c_jx_j + \dots + c_nx_n,$$

которую нужно минимизировать. Если эта функция выражена через прежние свободные переменные $x_1, x_2, \dots, x_n$, то, очевидно, после замены $x_j \leftrightarrow y_i$ ее нужно выразить через новые свободные переменные $x_1, x_2, \dots, x_{j-1}, y_i, x_{j+1}, \dots, x_n$. Нетрудно убедиться, что для этого может быть применен тот же алгоритм, что и для преобразования любой строки стандартной таблицы. Действительно, приводя L к стандартной форме

$$L = c_0 - (\gamma_1x_1 + \gamma_2x_2 + \dots + \gamma_nx_n),$$

где $\gamma_1 = -c_1; \gamma_2 = -c_2; \dots; \gamma_n = -c_n$, мы получаем еще одну строку (добавочную) стандартной таблицы, которая отличается от остальных только тем, что в ней никогда не выбирается разрешающий элемент.

▷ ПРИМЕР 3.7. Сделать замену $x_1 \leftrightarrow y_2$ в системе уравнений

$$\begin{cases} y_1 = x_1 - x_2 + x_3 - 1, \\ y_2 = \frac{1}{2}x_1 - x_3 - 3, \\ y_3 = 3x_2 - 2x_3 \end{cases} \quad (3.39)$$

и в линейной функции

$$L = -x_1 + 2x_2 - x_3 + 1.$$

Решение. Заполняем стандартную таблицу, в верхней строке которой помещаем линейную функцию L (табл. 3.7).

Для выполнения замены $x_1 \leftrightarrow y_2$ в той же таблице делаем дополнительные расчеты (табл. 3.8).

Заменой $x_1 \leftrightarrow y_2$ табл. 3.7 приводится к виду табл. 3.9. ►

С помощью табличного алгоритма обмена переменных в уравнениях ОЗЛП можно решить любую задачу линейного программирования или же убедиться, что она не имеет решения.

Нахождение решения каждой задачи линейного программирования распадается на два этапа:

- 1) отыскание опорного решения;
- 2) отыскание оптимального решения, минимизирующего линейную функцию L .

Таблица 3.7

	Свободный член	y_2 $\updownarrow$ x_1	x_2	x_3
L	1	1	-2	1
y_1	-1	-1	1	-1
$x_1 \leftrightarrow y_2$	-3	$\left(-\frac{1}{2}\right)$	0	1
y_3	0	0	-3	2

Таблица 3.8

	Свободный член	y_2 $\updownarrow$ x_1	x_2	x_3
L	1 -6	1 2	-2 0	1 2
y_1	-1 6	-1 -2	1 0	-1 -2
$x_1 \leftrightarrow y_2$	-3 6	$\left(-\frac{1}{2}\right)$ -2	0 0	1 -2
y_3	0 0	0 0	-3 0	2 0

На *п е р в о м* этапе попутно выясняется, имеет ли вообще данная задача допустимые (неотрицательные) решения; если да, то находится опорное решение, для которого все свободные переменные равны нулю, а все базисные — неотрицательны.

Таблица 3.9

	С в о б о д н ы й ч л е н	y_2	x_2	x_3
L	-5	2	-2	3
y_1	5	-2	1	-3
x_1	6	-2	0	-2
y_3	0	0	-3	2

На *втором* этапе попутно выясняется, ограничена ли минимизируемая функция L ; если нет, то оптимального решения не существует. Если да, то оно находится после того или другого числа замен $x_j \leftrightarrow y_i$.

Оба этапа решения ОЗЛП удобно выполнять с помощью описанного алгоритма преобразования стандартных таблиц.

тельных свободных членов остается прежним, убывали бы, по крайней мере, их абсолютные величины.

Существуют различные способы выбора *разрешающего* элемента для приближения к опорному решению. Остановимся (без строгого доказательства) на одном из них.

Выбор разрешающего элемента. Пусть имеется одно из уравнений (3.40) с отрицательным свободным членом. Отыскиваем в этой строке отрицательный элемент α_{ij} . Если такого элемента нет (все элементы $\alpha_{ij} \geq 0$), это признак того, что система уравнений (3.40) несовместима с неравенствами (3.42). Действительно, при отсутствии отрицательных элементов в строке вся правая часть соответствующего уравнения может быть только отрицательной, а это противоречит условиям неотрицательности переменных.

Предположим, что отрицательный элемент есть. Тогда выберем столбец, в котором он находится, и назовем его *разрешающим*.

Теперь надо выбрать в этом столбце сам разрешающий элемент. Рассмотрим все элементы данного столбца, имеющие одинаковый знак со свободным членом. Из них выберем тот, для которого отношение к нему свободного члена минимально. Строку, в которой стоит выбранный элемент, назовем *разрешающей*.

ОПРЕДЕЛЕНИЕ. Общий элемент разрешающей строки и разрешающего столбца называется *разрешающим элементом*.

Таким образом, выбирается разрешающий столбец, разрешающая строка и разрешающий элемент.

Покажем на примере, как совершается приближение к опорному решению при таком правиле выбора разрешающего элемента. Попутно убедимся в разумности этого правила.

▷ **ПРИМЕР 3.8.** Найти (если оно существует) опорное решение задачи линейного программирования с ограничениями-равенствами

$$\begin{cases} y_1 = 1 - (-x_1 - 2x_2 + x_3), \\ y_2 = -5 - (-2x_1 + x_2 - x_3), \\ y_3 = 2 - (x_1 + x_2), \\ y_4 = 1 - (-x_2 + x_3). \end{cases} \quad (3.43)$$

(Здесь не приводится линейная функция, которую нужно минимизировать, потому что опорное решение находится безотносительно к виду этой функции.)

Р е ш е н и е. Записываем условия (3.43) в виде стандартной таблицы (табл. 3.10).

Таблица 3.10

	Свободный член	x_1	x_2	x_3
y_1	1	-1	-2	1
y_2	-5	-2	1	-1
y_3	2	1	1	0
y_4	1	0	-1	1

В табл. 3.10 имеется отрицательный свободный член -5 в строке y_2 . Согласно правилу выбираем любой отрицательный элемент этой строки, например -2 (в табл. 3.10 он выделен полужирным шрифтом). Тем самым мы выбрали разрешающий столбец x_1 . В качестве “кандидатов” на роль разрешающего элемента рассмотрим все элементы этого столбца, совпадающие по знаку со своим свободным членом; это будут -2 и 1 (ноль не может быть принят в качестве разрешающего элемента).

Вычисляем для каждого “кандидата” отношение к нему свободного члена:

$$(-5) / (-2) = 5 / 2; \quad 2 / 1 = 2.$$

Наименьшее из этих отношений 2 ; значит, элемент 1 выбираем в качестве разрешающего и меняем местами $x_1 \leftrightarrow y_3$ (табл. 3.11). После выполнения действий приходим к табл. 3.12.

В табл. 3.12 по-прежнему один отрицательный свободный член, но по абсолютной величине он уже меньше, чем в табл. 3.10, — значит, мы приближаемся к ОДР.

Таблица 3.11

	Свободный член	y_3 $\updownarrow$ x_1	x_2	x_3
y_1	1 2	-1 1	-2 1	1 0
y_2	-5 4	-2 2	1 2	-1 0
$x_1 \leftrightarrow y_3$	2 2	① 1	1 1	0 0
y_4	1 0	0 0	-1 0	1 0

Таблица 3.12

	Свободный член	y_3	x_2	x_3
y_1	3	1	-1	1
y_2	-1	2	3	-1
x_1	2	1	1	0
y_4	1	0	-1	1

Попробуем избавиться и от этого члена. В строке y_2 имеется только один отрицательный элемент (-1) (выделен полужирным шрифтом). Значит, разрешающим столбцом может быть только столбец x_3 . Вычисляем для всех элементов этого столбца, имеющих одинаковый знак со своим свободным членом, отношение свободного члена к элементу:

$$3 / 1 = 3; (-1) / (-1) = 1; 1 / 1 = 1.$$

Отношение достигает минимума, равного 1, для двух элементов; возьмем в качестве разрешающего первый из них (-1) , стоящий в строке y_2 и столбце x_3 , сделаем замену $x_3 \leftrightarrow y_2$ (табл. 3.13 и 3.14).

В табл. 3.14 все свободные члены неотрицательны, и опорное решение найдено:

$$y_3 = x_2 = y_2 = 0; y_1 = 2; x_3 = 1; x_1 = 2; y_4 = 0. \blacktriangleright$$

Таблица 3.13

	Свободный член	y_3	x_2	y_2 $\updownarrow$ x_3
y_1	3 -1	1 2	-1 3	1 1
$x_3 \leftrightarrow y_2$	-1 1	2 -2	3 -3	(-1) -1
x_1	2 0	1 0	1 0	0 0
y_4	1 -1	0 2	-1 3	1 1

Таблица 3.14

	Свободный член	y_3	x_2	y_2
y_1	2	3	2	1
x_3	1	-2	-3	-1
x_1	2	1	1	0
y_4	0	2	2	1

▷ П Р И М Е Р 3.9. Найти (если оно существует) опорное решение системы

$$\begin{cases} y_1 = -4 - (-x_1 + 2x_2), \\ y_2 = -3 - (x_1 - x_2 + x_3), \\ y_3 = -10 - (2x_1 - x_2 + x_3), \\ y_4 = -2 - (-x_1 + x_2). \end{cases} \quad (3.44)$$

Р е ш е н и е. Записываем систему уравнений (3.44) в виде стандартной таблицы (табл. 3.15).

Таблица 3.15

	Свободный член	x_1	x_2	x_3
y_1	-4	-1	2	0
y_2	-3	1	-1	1
y_3	-10	2	-1	1
y_4	-2	-1	1	0

Выбираем строку с отрицательным свободным членом, например, первую. В ней есть отрицательный элемент (-1). Выбираем столбец x_1 в качестве разрешающего. Вычисляем отношения:

$$(-4) / (-1) = 4; (-2) / (-1) = 2.$$

Последнее отношение минимально; значит, в качестве разрешающего берем элемент (-1) в строке y_4 , и производим замену $x_1 \leftrightarrow y_4$ (табл. 3.16 и 3.17).

Обратим внимание на строку y_3 в табл. 3.17. В ней свободный член отрицателен, но нет ни одного отрицательного элемента (кроме самого свободного члена). Соответствующее уравнение имеет вид

$$y_3 = -14 - (2y_4 + x_2 + x_3).$$

Таблица 3.16

	Свободный член	y_4 $\updownarrow$ x_1	x_2	x_3
y_1	-4 2	-1 -1	2 -1	0 0
y_2	-3 -2	1 1	-1 1	1 0
y_3	-10 -4	2 2	-1 2	1 0
$x_1 \leftrightarrow y_4$	-2 2	(-1) -1	1 -1	0 0

Таблица 3.17

	Свободный член	y_4	x_2	x_3
y_1	-2	-1	1	0
y_2	-5	1	0	1
y_3	-14	2	1	1
x_1	2	-1	-1	0

Может ли при каких бы то ни было неотрицательных значениях y_4, x_2, x_3 величина y_3 быть неотрицательной? Очевидно, нет: при $y_4 = x_2 = x_3 = 0$ получим $y_3 = -14$, а увеличение y_4, x_2, x_3 сверх нуля сделает y_3 еще меньше. Следовательно, система (3.44) несовместима с неравенствами, вытекающими из неотрицательности переменных, и задача линейного программирования с условиями-ограничениями (3.44) *допустимых решений не имеет*. О том же

свидетельствует и строка y_2 табл. 3.17, где тоже нет ни одного отрицательного элемента (кроме самого свободного члена). ►

Таким образом, мы видим, что нет необходимости специально исследовать систему условий ОЗЛП на совместность в области неотрицательных решений; этот вопрос выясняется автоматически, в процессе нахождения опорного решения.

§7. Отыскание оптимального решения основной задачи линейного программирования

В предыдущем параграфе мы научились находить опорное решение системы уравнений ОЗЛП; при этом мы вовсе не занимались минимизируемой функцией L . Теперь приступим к оптимизации решения, т.е. найдем такое опорное решение, которое обращает в минимум линейную функцию

$$L = c_0 - (\gamma_1 x_1 + \gamma_2 x_2 + \dots + \gamma_n x_n).$$

Методика оптимизации решения была изложена в §5 главы 3. Теперь на примерах покажем, как эта оптимизация может быть проведена с помощью табличного алгоритма замены $x_j \leftrightarrow y_i$.

▷ П Р И М Е Р 3.10. Найти решение задачи линейного программирования с уравнениями

$$\begin{cases} y_1 = 2 - (x_1 + x_2 - 2x_3), \\ y_2 = 1 - (x_1 - x_2 + x_3), \\ y_3 = 5 - (x_2 + x_3), \\ y_4 = 2 - (2x_1 - x_2), \end{cases} \quad (3.45)$$

обращающее в минимум линейную функцию

$$L = 0 - (-x_1 + 2x_2 + x_3). \quad (3.46)$$

Р е ш е н и е. Все свободные члены в (3.45) неотрицательны, значит, опорное решение очевидно:

$$x_1 = x_2 = x_3 = 0; y_1 = 2; y_2 = 1; y_3 = 5; y_4 = 2.$$

Является ли оно оптимальным? Нет, так как коэффициенты при x_2 и x_3 в (3.46) положительны, и, увеличивая эти переменные, мы уменьшаем L .

Запишем (3.45) и (3.46) в виде стандартной таблицы (табл. 3.18).

Так как коэффициенты в первой строке при x_2 и x_3 положительны, любую из этих переменных можно вывести из числа свободных. Пусть это будет x_3 . Какой из элементов столбца x_3 выбрать разрешающим? Этот элемент должен быть положительным.

Таблица 3.18

	Свободный член	x_1	x_2	x_3
L	0	-1	2	1
y_1	2	1	1	-2
y_2	1	1	-1	①
y_3	5	0	1	1
y_4	2	2	-1	0

Значит, у нас есть выбор: 1 в строке y_2 или 1 в строке y_3 . Выберем из них тот, для которого отношение к нему свободного члена минимально (обоснование см. в § 5 главы 3).

Отношения равны $1 / 1 = 1$; $5 / 1 = 5$. Минимальное из них 1. Значит, в качестве разрешающего нужно взять элемент 1 в столбце x_3 , строке y_2 . Произведем замену $x_3 \leftrightarrow y_2$ (табл. 3.19 и 3.20).

В верхней строке табл. 3.20 есть положительный коэффициент при x_2 , значит, x_2 надо вывести из свободных переменных. Выбираем в качестве разрешающего тот положительный элемент столбца x_2 , для которого отношение к нему свободного члена минимально. Однако в столбце x_2 единственный положительный элемент 2, его и выбираем в качестве разрешающего (табл. 3.20 и 3.21).

Оказывается, процедура еще не закончена: в первой строке табл. 3.22 имеется положительный элемент в столбце y_2 , значит, переменную y_2 нужно вывести из числа свободных. В качестве разрешающего берем тот из положительных элементов столбца y_2 , для которого отношение к нему свободного члена минимально. Сравнивая отношения

$$6 : \frac{3}{2} = 4, \quad 3 : \frac{1}{2} = 6,$$

Таблица 3.19

	Свободный член	x_1	x_2	y_2 $\updownarrow$ x_3
L	0 -1	-1 -1	2 1	1 -1
y_1	2 2	1 2	1 -2	-2 2
$x_3 \leftrightarrow y_2$	1 1	1 1	-1 -1	(1) 1
y_3	5 -1	0 -1	1 1	1 -1
y_4	2 0	2 0	-1 0	0 1

Таблица 3.20

	Свободный член	x_1	x_2	y_2
L	-1	-2	3	-1
y_1	4	3	-1	2
x_3	1	1	-1	1
y_3	4	-1	(2)	-1
y_4	2	2	-1	0

Таблица 3.21

	Свободный член	x_1	y_3 $\updownarrow$ x_2	y_2
L	-1 -6	-2 $\frac{3}{2}$	3 $-\frac{3}{2}$	-1 $\frac{3}{2}$
y_1	4 2	3 $-\frac{1}{2}$	-1 $\frac{1}{2}$	2 $-\frac{1}{2}$
x_3	1 2	1 $-\frac{1}{2}$	-1 $\frac{1}{2}$	1 $-\frac{1}{2}$
$x_2 \leftrightarrow y_3$	4 2	-1 $-\frac{1}{2}$	(2) $\frac{1}{2}$	-1 $-\frac{1}{2}$
y_4	2 2	2 $-\frac{1}{2}$	-1 $\frac{1}{2}$	0 $-\frac{1}{2}$

Таблица 3.22

	Свободный член	x_1	y_3	y_2
L	-7	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{1}{2}$
y_1	6	$\frac{5}{2}$	$\frac{1}{2}$	($\frac{3}{2}$)
x_3	3	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
x_2	2	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$
y_4	4	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$

выбираем в качестве разрешающего элемент $\frac{3}{2}$ в строке y_1 и столбце y_2 и продолжаем процедуру оптимизации (табл. 3.23 и 3.24).

Таблица 3.23

	Свободный член	x_1	y_3	y_1 $\updownarrow$ y_2
L	-7 -2	$-\frac{1}{2}$ $-\frac{5}{6}$	$-\frac{3}{2}$ $-\frac{1}{6}$	$\frac{1}{2}$ $-\frac{1}{3}$
$y_2 \leftrightarrow y_1$	6 4	$\frac{5}{2}$ $\frac{5}{3}$	$\frac{1}{2}$ $\frac{1}{3}$	$\frac{3}{2}$ $\frac{2}{3}$
x_3	3 -2	$\frac{1}{2}$ $-\frac{5}{6}$	$\frac{1}{2}$ $-\frac{1}{6}$	$\frac{1}{2}$ $-\frac{1}{3}$
x_2	2 2	$-\frac{1}{2}$ $\frac{5}{6}$	$\frac{1}{2}$ $\frac{1}{6}$	$-\frac{1}{2}$ $\frac{1}{3}$
y_4	4 2	$\frac{3}{2}$ $\frac{5}{6}$	$\frac{1}{2}$ $\frac{1}{6}$	$-\frac{1}{2}$ $\frac{1}{3}$

В первой строке табл. 3.24 нет ни одного положительного элемента; значит, оптимальное решение достигнуто; оно таково:

$$x_1 = y_3 = y_1 = 0; y_2 = 4; x_3 = 1; x_2 = 4; y_4 = 6.$$

При этих значениях переменных линейная функция L достигает своего минимального значения, равного

$$L_{\min} = -9. \blacktriangleright$$

Возникает вопрос: как быть, если в столбце, содержащем положительный элемент строки L , не найдется ни одного положительного элемента, чтобы сделать его разрешающим? Легко убе-

Таблица 3.24

	Свободный член	x_1	y_3	y_1
L	-9	$-\frac{4}{3}$	$-\frac{5}{3}$	$-\frac{1}{3}$
y_2	4	$\frac{5}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
x_3	1	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$
x_2	4	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$
y_4	6	$\frac{7}{3}$	$\frac{2}{3}$	$\frac{1}{3}$

даться, что в этом случае функция L не ограничена снизу и ОЗЛП не имеет оптимального решения.

Действительно, в этом случае увеличение переменной, соответствующей данному столбцу, уменьшает линейную функцию L и не может сделать ни одной из базисных переменных отрицательной, т.е. ничто не препятствует неограниченному уменьшению функции L .

Итак, сформулируем **правила нахождения оптимального решения ОЗЛП симплекс-методом**.

1. Если все свободные члены (не считая строки L) в симплекс-методе неотрицательны, а в строке L (не считая свободного члена) нет ни одного положительного элемента, то *оптимальное решение достигнуто*.

2. Если в строке L есть положительный элемент, а в столбце, соответствующем ему, нет ни одного положительного элемента, то линейная функция L не ограничена снизу и *оптимального решения не существует*.

3. Если в этом столбце есть положительные элементы, то следует произвести замену одной из свободных переменных на одну из базисных, причем в качестве разрешающего надо взять такой элемент этого столбца, для которого отношение к нему соответствующего свободного члена минимально.

В заключение остановимся на так называемом *вырожденном* случае, когда один (или более) свободный член в уравнениях-ограничениях получается равным нулю. Это означает, что в данном опорном решении обращаются в нуль не только свободные переменные, но и некоторые базисные. Рассмотрим пример.

▷ П Р И М Е Р 3.11. Найти решение задачи линейного программирования с условиями

$$\begin{cases} y_1 = x_1 - x_2, \\ y_2 = -x_2 + x_3 + 2, \\ y_3 = x_3 + x_4 + 1, \end{cases} \quad (3.47)$$

обращающее в минимум линейную функцию

$$L = 2x_1 - x_2. \quad (3.48)$$

Р е ш е н и е. Записываем (3.47) и (3.48) в виде стандартной таблицы (см. табл. 3.25). Согласно общему правилу, находим в столбце x_2 разрешающий элемент, для которого отношение к нему свободного члена неотрицательно и минимально. Сравнивая отношения 0:1 и 2:1, останавливаемся на разрешающем элементе 1 в строке y_1 , для которого это отношение равно нулю. Осуществляем замену (табл. 3.26 и 3.27).

Таблица 3.25

	Свободный член	x_1	x_2	x_3	x_4
L	0	-2	1	0	0
y_1	0	-1	①	0	0
y_2	2	0	1	-1	0
y_3	1	0	0	-1	-1

Таблица 3.26

	Свободный член	x_1	y_1 $\updownarrow$ x_2	x_3	x_4
L	0 0	-2 1	1 -1	0 0	0 0
$x_2 \leftrightarrow y_1$	0 0	-1 -1	(1) 1	0 0	0 0
y_2	2 0	0 1	1 -1	-1 0	0 0
y_3	1 0	0 0	0 0	-1 0	-1 0

Таблица 3.27

	Свободный член	x_1	y_1	x_3	x_4
L	0	-1	-1	0	0
x_2	0	-1	1	0	0
y_2	2	1	-1	-1	0
y_3	1	0	0	-1	-1

При переходе от табл. 3.25 к табл. 3.27, естественно, не произошло уменьшения линейной функции L (она как была, так и осталась равной нулю), однако все элементы верхней строки стали неположительными, из чего видно, что оптимальное решение достигнуто: минимум функции равен нулю и достигается при $x_1 = y_1 = x_3 = x_4 = 0$; $x_2 = 0$; $y_2 = 2$; $y_3 = 1$. ►

В очень редких случаях может оказаться, что последовательное применение правила выбора разрешающего элемента приводит к тому, что после нескольких замен $x_j \leftrightarrow y_i$ мы вновь возвращаемся к тому же набору базисных и свободных переменных, с которого начали. Это и называется *заикливанием*. Для того чтобы избежать заикливания, достаточно при повторении взять разрешающий элемент не так, как он был взят первый раз (например, в другом столбце). При организации алгоритма линейного программирования на ПЭВМ в программу должно быть введено соответствующее указание.

§8. Транспортная задача линейного программирования

Изложенный в предыдущих параграфах симплекс-метод является универсальным, т.е. применим для решения любых задач линейного программирования. Однако некоторые типы таких задач, в частности так называемую транспортную задачу, можно решить более простыми методами.

Классическая транспортная задача линейного программирования формулируется следующим образом.

Имеется m пунктов отправления: $A_1, A_2, \dots, A_m$, в которых сосредоточены запасы однородного товара (груза) в количестве соответственно $a_1, a_2, \dots, a_m$ единиц. Кроме того, имеется n пунктов назначения: $B_1, B_2, \dots, B_n$, подавших заявки соответственно на $b_1, b_2, \dots, b_n$ единиц товара.

Предполагается, что сумма всех заявок равна сумме всех запасов:

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j. \quad (3.49)$$

Известна стоимость c_{ij} перевозки единицы товара от каждого пункта отправления A_i до каждого пункта назначения B_j . Таблица (матрица) стоимостей перевозки c_{ij} задана следующим образом:

$$C_{m \times n} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix}.$$

Требуется составить такой план перевозок, при котором все заявки были бы выполнены и общая стоимость всех перевозок была минимальна.

При такой постановке задачи показателем эффективности плана перевозок является *с т о и м о с т ь*, поэтому поставленную задачу точнее называют *транспортной задачей по критерию стоимости*.

1. Суммарное количество груза, направляемое из каждого пункта отправления во все пункты назначения, должно быть равно запасу груза в данном пункте. Это даст нам m условий-равенств:

[illegible]

$$\left\{ \begin{array}{l} \sum_{j=1}^n x_{1j} = a_1, \\ \sum_{j=1}^n x_{2j} = a_2, \\ \dots\dots\dots \\ \sum_{j=1}^n x_{mj} = a_m. \end{array} \right. \quad (3.50)$$
[illegible]

или

$$\begin{cases} \sum_{i=1}^m x_{i1} = b_1, \\ \sum_{i=1}^m x_{i2} = b_2, \\ \dots\dots\dots \\ \sum_{i=1}^m x_{in} = b_n. \end{cases} \quad (3.51)$$

3. Суммарная стоимость всех перевозок, т.е. сумма величин x_{ij} , умноженных на соответствующие стоимости c_{ij} , должна быть минимальной:

$$L = c_{11}x_{11} + c_{12}x_{12} + \dots + c_{1n}x_{1n} + c_{21}x_{21} + c_{22}x_{22} + \dots + c_{2n}x_{2n} + \dots + c_{m1}x_{m1} + c_{m2}x_{m2} + \dots + c_{mn}x_{mn} \rightarrow \min,$$

или

$$L = \sum_{i=1}^m \sum_{j=1}^n c_{ij}x_{ij} \rightarrow \min. \quad (3.52)$$

Знак двойной суммы $\sum_{i=1}^m \sum_{j=1}^n$ означает, что суммирование осуществляется по всем комбинациям индексов ($i = 1, \dots, m; j = 1, \dots, n$), т.е. по всем комбинациям пунктов отправления с пунктами назначения.

Функция (3.52) линейна, ограничения-равенства (3.50), (3.51) также линейны. Перед нами — типичная задача линейного программирования с ограничениями-равенствами (ОЗЛП).

Как и всякую другую задачу линейного программирования, ее можно было бы решить симплекс-методом, но некоторые особенности данной задачи позволяют решить ее более просто. Причиной является то, что *все коэффициенты при переменных в уравнениях (3.50), (3.51) равны единице*. Кроме того, имеет значение структура связей между условиями. Нетрудно убедиться, что не все $m + n$ уравнений нашей задачи независимы. Действительно, складывая между собой все уравнения (3.50) и все уравнения (3.51),

мы должны получить одно и то же в силу условия (3.49). Таким образом, условия (3.50), (3.51) связаны одной линейной зависимостью и фактически из этих уравнений только $m + n - 1$, а не $m + n$ являются линейно независимыми. Значит, ранг системы уравнений (3.50), (3.51)

$$r = m + n - 1,$$

а следовательно, можно разрешить эти уравнения относительно $m + n - 1$ базисных переменных, выразив их через остальные, свободные.

Подсчитаем количество свободных переменных:

$$k = mn - (m + n - 1) = mn - m - (n - 1) = m(n - 1) - (n - 1) = (m - 1)(n - 1).$$

Как известно, в задаче линейного программирования оптимальное решение достигается в одной из вершин ОДР, где по крайней мере k переменных обращаются в нуль. Значит, в нашем случае для оптимального плана перевозок по крайней мере $(m - 1)(n - 1)$ значений x_{ij} должны быть равны нулю.

Условимся о терминологии. Значения x_{ij} количества единиц груза, направляемых из пункта A_i в пункт B_j , будем называть *перевозками*.

Любую совокупность значений (x_{ij}) ($i = 1, \dots, m; j = 1, \dots, n$) будем называть *планом перевозок* или просто *планом*.

План (x_{ij}) будем называть *допустимым*, если он удовлетворяет условиям (3.50), (3.51), так называемым *балансовым условиям*: все заявки удовлетворены, все запасы исчерпаны.

Допустимый план будем называть *опорным*, если в нем отличны от нуля не более $r = m + n - 1$ базисных перевозок x_{ij} , а остальные перевозки равны нулю.

План (x_{ij}) будем называть *оптимальным*, если он среди всех допустимых планов приводит к наименьшей стоимости всех перевозок.

Перейдем к изложению **методов решения транспортной задачи (ТЗ)**. Эти методы не требуют манипуляций с симплекс-таблицами, а сводятся к более простым операциям непосредственно с таблицей, где в определенном порядке записаны все условия ТЗ. Такая таблица называется *транспортной* (табл. 3.28). В транспортной таблице записываются:

- пункты отправления (ПО) и назначения (ПН);
- запасы, имеющиеся в пунктах отправления a_i ;
- заявки, поданные пунктами назначения b_j ;
- стоимости перевозки из каждого пункта отправления в каждый пункт назначения c_{ij} .

Стоимости перевозок мы будем помещать в правом верхнем углу каждой клетки, с тем чтобы в самой клетке при составлении плана помещать перевозки x_{ij} .

Таблица 3.28

ПО \ ПН	B_1	B_2	$\dots$	B_n	Запасы a_i
A_1	c_{11}	c_{12}	$\dots$	c_{1n}	a_1
A_2	c_{21}	c_{22}	$\dots$	c_{2n}	a_2
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
A_m	c_{m1}	c_{m2}	$\dots$	c_{mn}	a_m
Заявки b_j	b_1	b_2	$\dots$	b_n	$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$

В правом верхнем углу каждой клетки проставлена стоимость перевозки единицы товара (груза) из ПО A_i в ПН B_j . В правом столбце помещены запасы товара в каждом ПО, в нижней строке — заявки, поданные каждым ПН. Для ТЗ сумма запасов равна сумме заявок; общее значение этой суммы записывается в правой нижней клетке таблицы.

Ранее мы показали, что ранг системы уравнений-ограничений ТЗ равен $r = m + n - 1$, где m — число строк, а n — число столбцов транспортной таблицы. Значит, в каждом опорном плане, включая оптимальный, будут отличны от нуля не более чем $m + n - 1$ перевозок.

Клетки таблицы, в которых мы будем записывать эти отличные от нуля перевозки, условимся называть *базисными*, а остальные (пустые) — *свободными*.

Таким образом, решение ТЗ свелось к следующему. Найти такие значения положительных перевозок, которые, будучи проставлены в базисных клетках транспортной таблицы, удовлетворяли бы следующим условиям:

- сумма перевозок в каждой строке таблицы должна быть равна запасу данного ПО;
- сумма перевозок в каждом столбце должна быть равна заявке данного ПН;
- общая стоимость перевозок — минимальная.

В дальнейшем все действия по нахождению решения ТЗ будут сводиться к преобразованию транспортной таблицы 3.28

При описании этих преобразований удобно пользоваться нумерацией клеток таблицы (подобной нумерации клеток шахматной доски). Клеткой (A_i, B_j) или, короче, клеткой (i, j) мы будем называть клетку, стоящую в i -й строке и j -м столбце транспортной таблицы. Например, самая верхняя левая клетка будет обозначаться $(1, 1)$, стоящая под ней — $(2, 1)$ и т. д.

§9. Нахождение опорного плана

Решение ТЗ, как и всякой задачи линейного программирования, начинается с нахождения опорного решения, или, как мы будем говорить, опорного плана. В отличие от общего случая ОЗЛП с произвольными ограничениями и минимизируемой функцией, решение ТЗ всегда существует. Вполне очевидно, что какой-нибудь допустимый план существовать должен. Среди допустимых планов непременно имеется оптимальный (может быть, не один), потому что линейная функция L — стоимость перевозок заведомо неотрицательна (ограничена снизу нулем). Как построить опорный план? Для этого существуют различные способы. Рассмотрим простейший из них, так называемый способ северо-западного угла. Поясним его на примере.

▷ П Р И М Е Р 3.12. Условия ТЗ заданы табл. 3.29.

Таблица 3.29

ПО \ ПН	B_1	B_2	B_3	B_4	B_5	Запасы a_i
A_1	10	8	5	6	9	48
A_2	6	7	8	6	5	30
A_3	8	7	10	8	7	27
A_4	7	5	4	6	8	20
Заявки b_j	18	27	42	12	26	125

Требуется найти опорное решение ТЗ (построить опорный план).

Р е ш е н и е. Перепишем табл. 3.29 и будем заполнять ее перевозками постепенно, начиная с левой верхней клетки (1, 1) (с северо-западного угла таблицы). Будем рассуждать при этом следующим образом. Пункт B_1 подал заявку на 18 единиц груза. Удовлетворим эту заявку за счет запаса 48, имеющегося в пункте A_1 , и

запишем перевозку 18 в клетке (1, 1) (табл. 3.30). После этого заявка пункта B_1 удовлетворена, а в пункте A_1 осталось еще 30 единиц груза. Удовлетворим за счет них заявку пункта B_2 (27 единиц), запишем 27 в клетке (1, 2); оставшиеся 3 единицы пункта A_1 назначим пункту B_3 . В составе заявки пункта B_3 остались неудовлетворенными 39 единиц, 30 из которых покроем за счет пункта A_2 (чем его запас будет исчерпан) и еще 9 возьмем из пункта A_3 . Из оставшихся 18 единиц пункта A_3 12 выделим пункту B_4 , а 6 единиц назначим пункту B_5 , что вместе со всеми 20 единицами пункта A_4 покроет его заявку (см. табл. 3.30).

На этом распределение запасов закончено: каждый пункт назначения получил груз согласно своей заявке. Это выражается в том, что сумма перевозок в каждой строке равна соответствующему запасу, а в столбце — заявке.

Таблица 3.30

ПН ПО	B_1	B_2	B_3	B_4	B_5	Запасы a_i
A_1	10 18	8 27	5 3	6	9	48
A_2	6	7	8 30	6	5	30
A_3	8	7	10 9	8 12	7 6	27
A_4	7	5	4	6	8 20	20
Заявки b_j	18	27	42	12	26	125

Таким образом, мы составили план перевозок, удовлетворяющий балансовым условиям. Полученное решение ТЗ является допустимым.

Число базисных клеток удовлетворяет условию $r = m + n - 1 = 8$. Число свободных клеток равно $(n - 1)(m - 1) = 12$. Значит, наш план — не только допустимый, но и опорный, и поставленная задача построения опорного плана решена. ►

Возникает вопрос: является ли этот план оптимальным по стоимости? Разумеется, нет! Ведь при его построении мы совсем не учитывали стоимостей перевозок c_{ij} . Стоимость полученного плана (если умножить каждую перевозку на соответствующую стоимость) равна

$$18 \cdot 10 + 27 \cdot 8 + 3 \cdot 5 + 30 \cdot 8 + 9 \cdot 10 + 12 \cdot 8 + 6 \cdot 7 + 20 \cdot 8 = 1039.$$

Попробуем улучшить этот план, перенеся, например, 18 единиц из клетки (1, 1) в клетку (2, 1) и, чтобы не нарушить баланса, перенеся те же 18 единиц из клетки (2, 3) в клетку (1, 3). Получим новый план, приведенный в табл. 3.31.

Нетрудно убедиться, что стоимость нового плана равна

$$27 \cdot 8 + 21 \cdot 5 + 18 \cdot 6 + 12 \cdot 8 + 9 \cdot 10 + 12 \cdot 8 + 6 \cdot 7 + 20 \cdot 8 = 913,$$

т.е. на 126 единиц меньше стоимости плана, приведенного в табл. 3.30.

Таблица 3.31

ПО \ ПН	B_1	B_2	B_3	B_4	B_5	Запасы a_i
A_1	10 27	8	5 21	6	9	48
A_2	6 18	7	8 12	6	5	30
A_3	8	7	10 9	8 12	7 6	27
A_4	7	5	4	6	8 20	20
Заявки b_j	18	27	42	12	26	125

Таким образом, за счет перестановки 18 единиц груза из одних клеток в другие нам удалось понизить стоимость плана. На этом способе уменьшения стоимости в дальнейшем и будет основан алгоритм оптимизации плана перевозок.

Остановимся на одной особенности плана перевозок, которая может встретиться как при построении опорного плана, так и при

его улучшении. Речь идет о так называемом *вырожденном плане*, в котором некоторые из базисных перевозок оказываются равными нулю. Рассмотрим конкретный пример возникновения вырожденного плана.

▷ПРИМЕР 3.13. Дана транспортная таблица (табл. 3.32) (без стоимости перевозок, так как речь идет только о построении опорного плана).

Таблица 3.32

ПО \ ПН	B_1	B_2	B_3	B_4	B_5	Запасы a_i
A_1						20
A_2						30
A_3						25
A_4						20
Заявки b_j	10	10	20	35	20	95

Составить опорный план перевозок.

Р е ш е н и е. Применяя способ северо-западного угла, получим табл. 3.33.

Таблица 3.33

ПО \ ПН	B_1	B_2	B_3	B_4	B_5	Запасы a_i
A_1	10	10				20
A_2			20	10		30
A_3				25		25
A_4					20	20
Заявки b_j	10	10	20	35	20	95

Опорный план составлен. Особенностью его является то, что в нем только шесть, а не восемь ($m + n - 1 = 8$), как должно быть, отличных от нуля перевозок. Значит, некоторые базисные перевозки, оказались равными нулю. Нетрудно заметить, отчего это произошло: при распределении запасов по пунктам назначения в некоторых случаях остатки оказывались равными нулю и в соответствующую клетку не попадали. ►

Такие случаи “вырождения” могут возникать не только при составлении опорного плана, но и при его преобразовании, оптимизации.

В дальнейшем нам удобно будет *всегда* иметь в транспортной таблице $m + n - 1$ базисных клеток, хотя в некоторых из них, может быть, будут нулевые значения перевозок. Для этого можно ничтожно мало изменить запасы или заявки, так чтобы общий баланс не нарушился, а лишние, “промежуточные” балансы уничтожились, т.е. приняли тривиальное значение. Достаточно в нужных местах изменить запасы или заявки, например, на величину ε , а после нахождения оптимального решения положить $\varepsilon = 0$.

Покажем, как перейти от вырожденного плана к невырожденному на примере табл. 3.33. Изменим запасы в первой строке и положим их равными $20 + \varepsilon$ (табл. 3.34). Кроме того, в третьей строке поставим запасы $25 + \varepsilon$. Чтобы “свести баланс”, в четвертой строке ставим запасы $20 - 2\varepsilon$. Для этой таблицы строим опорный план способом северо-западного угла.

Таблица 3.34

ПН ПО	B_1	B_2	B_3	B_4	B_5	Запасы a_i
A_1	10	10	ε			$20 + \varepsilon$
A_2			$20 - \varepsilon$	$10 + \varepsilon$		30
A_3				$25 - \varepsilon$	2ε	$25 + \varepsilon$
A_4					$20 - 2\varepsilon$	$20 - \varepsilon$
Заявки b_j	10	10	20	35	20	95

В табл. 3.34 уже содержится столько базисных переменных, сколько требуется: $m + n - 1 = 8$. В дальнейшем, после оптимизации плана, можно будет положить $\varepsilon = 0$.

§ 10. Улучшение плана перевозок. Цикл пересчета

В предыдущем параграфе мы уже бегло ознакомились со способом улучшения плана, состоящим в переносе (перестановках) некоторых перевозок, без нарушения баланса, из клетки в клетку по некоторому замкнутому циклу. Рассмотрим циклические перестановки подробнее.

Возьмем транспортную таблицу, состоящую, например, из $m = 5$ строк и $n = 6$ столбцов (число строк и столбцов несущественно).

Циклом пересчета в транспортной таблице мы будем называть несколько клеток, соединенных замкнутой ломаной линией, которая в каждой клетке совершает поворот на 90° .

Например, в табл. 3.35 изображены два цикла: первый с четырьмя вершинами (2, 1), (2, 3), (4, 3), (4, 1) и второй — с восемью вершинами (1, 4), (1, 6), (4, 6), (4, 4), (3, 4), (3, 5), (5, 5), (5, 4). Стрелками показано направление обхода цикла.

Нетрудно убедиться, что каждый цикл имеет четное число вершин и, значит, четное число звеньев (стрелок). Условимся отмечать знаком “+” вершины цикла, в которых перевозки увеличиваются, а знаком “−” вершины, в которых они уменьшаются. Цикл с отмеченными вершинами будем называть *означенным*. В табл. 3.36 показано два означенных цикла: первый Π_1 с четырьмя вершинами (1, 1), (1, 2), (3, 2) и (3, 1) и второй Π_2 с восемью вершинами (3, 4), (3, 6), (5, 6), (5, 3), (2, 3), (2, 5), (4, 5) и (4, 4).

Перенести (“перебросить”) какое-то количество единиц груза по означенному циклу — это значит увеличить перевозки, стоящие в положительных вершинах цикла, на это количество единиц и уменьшить на то же количество перевозки, стоящие в отрицательных вершинах. Очевидно, что при переносе любого числа единиц по циклу равновесие между запасами и заявками не меняется: по-прежнему сумма перевозок в каждой строке равна запасам этой строки, а сумма перевозок в каждом столбце — заявке этого столбца. Таким образом, при любом циклическом переносе, оставляющем перевозки неотрицательными, *допустимый план*

Таблица 3.35

ПН ПО	B_1	B_2	B_3	B_4	B_5	B_6	Запасы a_i
A_1	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}	a_1
A_2	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}	c_{26}	a_2
A_3	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}	c_{36}	a_3
A_4	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}	c_{46}	a_4
A_5	c_{51}	c_{52}	c_{53}	c_{54}	c_{55}	c_{56}	a_5
Заявки b_j	b_1	b_2	b_3	b_4	b_5	b_6	$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$

Таблица 3.36

ПН ПО	B_1	B_2	B_3	B_4	B_5	B_6	Запасы a_i
A_1	$+$ c_{11}	$-$ c_{12}	c_{13}	c_{14}	c_{15}	c_{16}	a_1
A_2	c_{21}	c_{22}	$+$ c_{23}	c_{24}	$-$ c_{25}	c_{26}	a_2
A_3	$-$ c_{31}	$+$ c_{32}	c_{33}	$+$ c_{34}	$-$ c_{35}	c_{36}	a_3
A_4	c_{41}	c_{42}	c_{43}	$-$ c_{44}	$+$ c_{45}	c_{46}	a_4
A_5	c_{51}	c_{52}	$-$ c_{53}	c_{54}	c_{55}	$+$ c_{56}	a_5
Заявки b_j	b_1	b_2	b_3	b_4	b_5	b_6	$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$

остается допустимым. Стоимость же плана при этом может меняться — увеличиваться или уменьшаться.

Назовем *ценой цикла* увеличение стоимости перевозок при перемещении одной единицы груза по означенному циклу. Очевидно, что цена цикла равна алгебраической сумме стоимостей, стоящих в вершинах цикла, причем в положительных вершинах сто-

имость берется со знаком “+”, а в отрицательных — со знаком “-”. Например, для цикла C_1 в табл. 3.36 цена равна

$$c_{11} - c_{12} + c_{32} - c_{31},$$

для цикла C_2

$$c_{34} - c_{36} + c_{56} - c_{53} + c_{23} - c_{25} + c_{45} - c_{44}.$$

Обозначим цену цикла через γ . При перемещении одной единицы груза по циклу стоимость перевозок увеличивается на величину γ , а при перемещении k единиц груза — соответственно, на $k\gamma$.

Очевидно, для улучшения плана имеет смысл перемещать перевозки только по тем циклам, цена которых отрицательна. Каждый раз, когда нам удастся совершить такое перемещение, стоимость плана уменьшается на соответствующую величину $k\gamma$.

Так как перевозки не могут быть отрицательными, мы будем пользоваться только такими циклами, отрицательные вершины которых лежат в базисных клетках таблицы, где стоят положительные перевозки¹. Если циклов с отрицательной ценой в таблице больше не осталось, это означает, что дальнейшее улучшение плана невозможно, т.е. оптимальный план достигнут.

Метод последовательного улучшения плана перевозок и состоит в том, что в таблице отыскиваются циклы с отрицательной ценой, по ним перемещаются перевозки, и план улучшается до тех пор, пока циклов с отрицательной ценой уже не останется.

При улучшении плана циклическими переносами, как правило, пользуются приемом, заимствованным из симплекс-метода: при каждом шаге (цикле) заменяют одну свободную переменную на базисную, т.е. заполняют одну свободную клетку и взамен того освобождают одну базисную клетку. При этом общее число базисных клеток остается неизменным: $m + n - 1$. Этот прием удобен тем, что для него легче находить подходящие циклы.

Можно доказать, что *для любой свободной клетки транспортной таблицы всегда существует цикл (и притом единственный), одна из вершин которого лежит в этой свободной клетке, а все ос-*

¹ В случае вырождения, как мы увидим далее, может оказаться полезным *ф и к т и в н ы й* перенос по циклу, отрицательная вершина которого лежит в клетке с нулевой перевозкой.

тальные — в базисных клетках. Если цена такого цикла (с плюсом в свободной клетке) отрицательна, то план можно улучшить перемещением перевозок по данному циклу. Количество единиц груза k , которое можно переместить, определяется минимальным значением перевозок, стоящих в отрицательных вершинах цикла (если переместить большее число единиц груза, возникнут отрицательные перевозки).

▷ П Р И М Е Р 3.14. Найти оптимальный план для транспортной задачи, приведенной в табл. 3.37.

Таблица 3.37

ПО \ ПН	B_1	B_2	B_3	B_4	Запасы a_i
A_1	10	7	6	8	31
A_2	5	6	5	4	48
A_3	8	7	6	7	38
Заявки b_j	22	34	41	20	117

Р е ш е н и е. Составляем опорный план способом северо-западного угла (табл. 3.38).

Таблица 3.38

ПО \ ПН	B_1	B_2	B_3	B_4	Запасы a_i
A_1	10 22	7 9	6	8	31
A_2	5	6 25	5 23	4	48
A_3	8	7	6 18	7 20	38
Заявки b_j	22	34	41	20	117

Стоимость этого плана

$$L_1 = 22 \cdot 10 + 9 \cdot 7 + 25 \cdot 6 + 23 \cdot 5 + 18 \cdot 6 + 20 \cdot 7 = 796.$$

Число базисных переменных, как и полагается в невырожденном случае, равно $r = m + n - 1 = 3 + 4 - 1 = 6$.

Попробуем улучшить план, заняв свободную клетку (2, 4) с минимальной стоимостью 4. Цикл, соответствующий этой клетке, показан в табл. 3.38. Цена этого цикла равна $\gamma = 4 - 7 + 6 - 5 = -2$.

По этому циклу мы можем переместить максимум 20 единиц груза (чтобы не получить в клетке (3, 4) отрицательной перевозки). Новый, улучшенный план показан в табл. 3.39.

Таблица 3.39

ПО \ ПН	B_1	B_2	B_3	B_4	Запасы a_i
A_1	10 22	7 9	6	8	31
A_2	5 +	6 - 25	5 3	4 20	48
A_3	8	7	6 38	7	38
Заявки b_j	22	34	41	20	117

Стоимость этого плана $L_2 = 796 + 20 \cdot (-2) = 756$. В нем по-прежнему шесть базисных клеток.

Для дальнейшего улучшения плана обратим внимание на свободную клетку (2, 1) со стоимостью 5. Цикл, соответствующий этой клетке, показан в табл. 3.39. Цена его $\gamma = 7 - 6 + 5 - 10 = -4$. По этому циклу переместим 22 единицы груза, чем уменьшим стоимость перевозок до $L_3 = 756 + 22 \cdot (-4) = 668$ (табл. 3.40).

Попробуем дальше улучшить этот план, подсчитывая цены циклов, начинающихся положительной вершиной в свободной клетке. Просматриваем имеющиеся свободные клетки табл. 3.40 и определяем цену цикла для каждой из них. Все эти цены или положительные, или нулевые (предлагаем читателю проверить это

Таблица 3.40

ПО \ ПН	B_1	B_2	B_3	B_4	Запасы a_i
A_1	10	7 31	6	8	31
A_2	5 22	6 3	5 3	4 20	48
A_3	8	7	6 38	7	38
Заявки b_j	22	34	41	20	117

самостоятельно). Следовательно, никакой циклический перенос перевозок не может улучшить план перевозок. Таким образом, план, полученный в табл. 3.40, является оптимальным. ►

Указанный метод отыскания оптимального решения ТЗ называется *распределительным*; он состоит в отыскании свободных клеток с отрицательной ценой цикла и в перенесении перевозок по этому циклу.

► ПРИБЕР 3.15. Найти оптимальный план перевозок для ТЗ, условия которой приведены в табл. 3.41.

Таблица 3.41

ПО \ ПН	B_1	B_2	B_3	Запасы a_i
A_1	10	5	4	40
A_2	6	4	5	23
A_3	7	3	6	20
Заявки b_j	20	20	43	83

Р е ш е н и е. Строим опорный план способом северо-западного угла; он получается вырожденным. Чтобы избежать этого, нарушаем баланс запасов и заявок на ε в первой и третьей строках, не нарушая общего баланса (сумма запасов равна сумме заявок). После этого строим опорный план тем же способом северо-западного угла (табл. 3.42), в нем ровно столько базисных переменных, сколько нужно: пять. Улучшаем план перевозок переносом $20 - \varepsilon$ единиц груза по циклу, показанному в табл. 3.42; получим новый, улучшенный план (см. табл. 3.43).

План, приведенный в табл. 3.41, еще не оптимален, так как цикл с началом в свободной клетке (2, 1) имеет отрицательную цену:

$$\gamma = 6 - 10 + 4 - 5 = -5.$$

Таблица 3.42

ПО \ ПН	B_1	B_2	B_3	Запасы a_i
A_1	¹⁰ 20	⁵ 20	⁴ + ε	$40 + \varepsilon$
A_2	⁶	⁴	⁵ 23	23
A_3	⁷	³ +	⁶ - 20 - ε	$20 - \varepsilon$
Заявки b_j	20	20	43	83

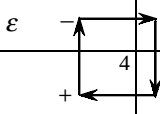
Таблица 3.43

ПО \ ПН	B_1	B_2	B_3	Запасы a_i
A_1	¹⁰ 20	⁵ ε	⁴ + 20	$40 + \varepsilon$
A_2	⁶ +	⁴	⁵ - 23	23
A_3	⁷	³ 20 - ε	⁶	$20 - \varepsilon$
Заявки b_j	20	20	43	83

Перемещаем по этому циклу 20 единиц груза; получаем табл. 3.44.

Таблица 3.44

ПО \ ПН	B_1	B_2	B_3	Запасы a_i
A_1	10	5	4	$40 + \varepsilon$
A_2	20	6	5	23
A_3	7	3	6	$20 - \varepsilon$
Заявки b_j	20	20	43	83



Цена цикла, начинающегося в клетке (2, 2) табл. 3.44, также отрицательна: $4 - 5 + 4 - 5 = -2$. Однако по этому циклу можно перенести только перевозку, равную ε . Тем не менее сделаем это и получим новый план (табл. 3.45).

Таблица 3.45

ПО \ ПН	B_1	B_2	B_3	Запасы a_i
A_1	10	5	$40 + \varepsilon$	$40 + \varepsilon$
A_2	20	ε	$3 - \varepsilon$	23
A_3	7	$20 - \varepsilon$	6	$20 - \varepsilon$
Заявки b_j	20	20	43	83

В табл. 3.45 все циклы, соответствующие свободным клеткам, имеют неотрицательную цену, поэтому план, приведенный в табл. 3.45, является оптимальным. Полагая в нем $\varepsilon = 0$, получим окончательный оптимальный план (табл. 3.46) с минимальной стоимостью перевозок

$$L_{\min} = 40 \cdot 4 + 20 \cdot 6 + 3 \cdot 5 + 20 \cdot 3 = 355. \blacktriangleright$$

Таблица 3.46

П О \ П Н	B_1	B_2	B_3	Запасы a_i
A_1	¹⁰	⁵	⁴ 40	40
A_2	⁶ 20	⁴	⁵ 3	23
A_3	⁷	³ 20	⁶	20
Заявки b_j	20	20	43	83

Заметим, что примененный здесь метод “ликвидации вырождения” путем ε -изменения запасов не совсем удобен, так как требует дополнительных действий с ε -измененными данными. Проще было бы при заполнении табл. 3.42 не изменять запасы, а вообразить их себе измененными и вместо ε поставить в базисной клетке (3, 3) нуль. Базисная клетка с нулевой перевозкой будет отличаться от свободной тем, что в ней нуль будет проставлен, а в свободной — нет. Дальнейшие манипуляции с транспортной таблицей будут точно такими же, как если бы в базисных клетках стояли только положительные перевозки. Разница лишь в том, что, когда одна из отрицательных вершин цикла окажется в базисной клетке с нулевой перевозкой, нужно переносить по этому циклу нулевую перевозку (фиктивный перенос). Если в транспортной таблице немного (одна-две) базисных переменных обращаются в нуль, можно рекомендовать этот простой метод вместо ε -изменений запасов (заявок). (Предлагаем читателю самостоятельно решить пример 3.15 упрощенным способом.) Следует иметь в виду, что при большом количестве базисных переменных, обращающихся в нуль, упрощенный метод становится менее удобным, так как легко запутаться с расстановкой по таблице нулевых базисных перевозок (т.е. ошибочно проставить базисные клетки там, где они находиться не могут).

§ 11. Двойственность в линейном программировании

11.1. Понятие двойственности

С каждой задачей линейного программирования тесно связана другая линейная задача, называемая *двойственной*. Первоначальная задача называется *исходной*.

Связь исходной и двойственной задач состоит в том, что коэффициенты C_j функции цели исходной задачи являются свободными членами системы ограничений двойственной задачи, свободные члены b_i системы ограничений исходной задачи служат коэффициентами функции цели двойственной задачи, а матрица коэффициентов системы ограничений двойственной задачи является транспонированной матрицей коэффициентов системы ограничений исходной задачи. Решение двойственной задачи может быть получено из решения исходной и наоборот.

В качестве примера рассмотрим задачу использования ресурсов. Предположим, предприятие имеет m видов ресурсов в количестве b_i ($i = 1, 2, \dots, m$) единиц, из которых производится n видов продукции. Для производства 1 ед. j -й продукции расходуется a_{ij} ед. i -го ресурса, а ее стоимость составляет C_j ед. Составить план выпуска продукции, обеспечивающий ее максимальный выпуск в стоимостном выражении. Обозначим через x_j ($j = 1, 2, \dots, n$) количество единиц j -й продукции. Тогда исходную задачу сформулируем так.

Найти вектор $X = (x_1, x_2, \dots, x_n)$, который удовлетворяет ограничениям

$$\left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2, x_j \geq 0 (j=1,2,\dots,n) \\ \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m, \end{array} \right.$$

и составляет максимальное значение линейной функции

$$Z = C_1x_1 + C_2x_2 + \dots + C_nx_n.$$

идущих на изготовление единицы j -й продукции, равна $\sum_{i=1}^m a_{ij} y_i$.

выполняться неравенство $\sum_{i=1}^m a_{ij} y_i \geq C_j, j = 1, 2, \dots, n$. Стоимость всех

имеющихся ресурсов выразится величиной $\sum_{i=1}^m b_i y_i$. Итак, *двойст-*

Найти вектор $Y = (y_1, y_2, \dots, y_m)$, который удовлетворяет ограничениям

$$\begin{cases} a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \geq C_1, \\ a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \geq C_2, y_i \geq 0 \quad (i=1,2,\dots,n) \\ \\ a_{m1}y_1 + a_{m2}y_2 + \dots + a_{mn}y_n \geq C_m, \end{cases}$$

$$f = b_1 y_1 + b_2 y_2 + \dots + b_m y_m.$$

Исходная задача. Сколько и какой продукции x_j ($j = 1, 2, \dots, n$) необходимо произвести, чтобы при заданных стоимостях C_j , ($j = 1, 2, \dots, n$) единицы продукции и размерах имеющихся ресурсов b_i ($i = 1, 2, \dots, m$) максимизировать выпуск продукции в стоимостном выражении.

207

b_i и величинах стоимости единицы продукции C_j минимизировать общую стоимость затрат?

Переменные y_i называются *оценками* или *учетными, неявными* ценами.

Многие задачи линейного программирования первоначально ставятся в виде исходных или двойственных задач, поэтому имеет смысл говорить о паре двойственных задач линейного программирования.

11.2. Несимметричные двойственные задачи

Двойственные задачи бывают симметричными и несимметричными в зависимости от системы ограничений.

В несимметричных двойственных задачах система ограничений исходной задачи задается в виде равенств, а двойственной – в виде неравенств, причем в последней переменные могут быть и отрицательными. Для простоты изложения постановку задачи условимся записывать в матричной форме.

Исходная задача. Найти матрицу-столбец $X = (x_1, x_2, \dots, x_n)$, которая удовлетворяет ограничениям

$$AX = A_0, X \geq 0$$

и минимизирует линейную функцию

$$Z = CX.$$

Двойственная задача. Найти матрицу-строку $Y = (y_1, y_2, \dots, y_m)$, которая удовлетворяет ограничениям

$$YA \leq C$$

и максимизирует линейную функцию $f = Y A_0$.

В обеих задачах $C = (c_1, c_2, \dots, c_n)$ – матрица-строка, $A_0 = (b_1, b_2, \dots, b_m)$ – матрица-столбец, $A = (a_{ij})$ – матрица коэффициентов системы ограничений. Связь между оптимальными планами пары двойственных задач устанавливает следующая теорема.

ТЕОРЕМА (Теорема двойственности).

Если из пары двойственных задач одна обладает оптимальным планом, то и другая имеет решение, причем для экстремальных значений линейных функций выполняется соотношение

$$\text{Min } Z = \text{max } f.$$

Если линейная функция одной из задач не ограничена, то другая не имеет решения.

Если рассматривать задачу, двойственную к двойственной задаче, то снова придем к исходной задаче.

Рассмотрим исходную задачу.

Исходная задача. Найти минимальное значение линейной функции $Z = x_2 - x_4 - 3x_5$ при ограничениях

$$\begin{cases} x_1 + 2x_2 - x_4 + x_5 = 1, \\ -4x_2 + x_3 + 2x_4 - x_5 = 2, \quad x_j \geq 0, (j = 1, 2, \dots, 6), \\ 3x_2 + x_5 + x_6 = 5. \end{cases}$$

Здесь матрица-строка $C = (0; 1; 0; -1; -3; 0)$, матрица-столбец

$$A_0 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 2 & 0 & -1 & 1 & 0 \\ 0 & -4 & 1 & 2 & -1 & 0 \\ 0 & 3 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Выпишем транспонированную матрицу к A .

$$A^T = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -4 & 3 \\ 0 & 1 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Двойственная задача. Найти максимальное значение линейной функции $f = y_1 + 2y_2 + 5y_3$ при ограничениях

$$\begin{cases} y_1 \leq 0, \\ 2y_1 - 4y_2 + 3y_3 \leq 1, \\ y_2 \leq 0, \\ -y_1 + 2y_2 \leq -1, \\ y_1 - y_2 + y_3 \leq -3, \\ y_3 \leq 0. \end{cases}$$

11.3. Симметричные двойственные задачи

Разновидностью двойственных задач линейного программирования являются двойственные симметричные задачи, в которых система ограничений как исходной, так и двойственной задач задается неравенствами, причем на двойственные переменные налагается условие неотрицательности.

Исходная задача. Найти матрицу-столбец $X = (x_1, x_2, \dots, x_n)$, которая удовлетворяет системе ограничений $AX \geq A_0$, $X \geq 0$ и минимизирует линейную функцию $Z = CX$.

Двойственная задача. Найти матрицу-строку $Y = (y_1, y_2, \dots, y_m)$, которая удовлетворяет системе ограничений $YA \leq C$, $Y \geq 0$ и максимизирует линейную функцию $f = YA_0$.

Систему неравенств с помощью дополнительных переменных можно преобразовать в систему уравнений, поэтому всякую пару симметричных двойственных задач можно преобразовать в пару несимметричных, для которых теорема двойственности уже доказана.

Используя симметричность, можно выбрать задачу, более удобную для решения. Объем задачи, решаемой с помощью ЭВМ, ограничен числом включаемых строк, поэтому задача, довольно громоздкая в исходной постановке, может быть упрощена в двойственной формулировке. При вычислениях без помощи машин использование двойственности упрощает вычисления.

Исходная задача. Найти минимальное значение линейной функции $Z = x_1 + 2x_2 + 3x_3$ при ограничениях

$$\begin{cases} 2x_1 + 2x_2 - x_3 \geq 2, \\ x_1 - x_2 - 4x_3 \leq -3, \\ x_1 + x_2 - 2x_3 \geq 6, \\ 2x_1 + x_2 - 2x_3 \geq 3, x_j \geq 0, (j = 1, 2, 3). \end{cases}$$

Очевидно, для того чтобы записать двойственную задачу, сначала необходимо систему ограничений исходной задачи привести к виду (3.53). Для этого второе неравенство следует умножить на -1 .

Двойственная задача. Найти максимум линейной функции $f = 2y_1 + 3y_2 + 6y_3 + 3y_4$ при ограничениях

$$\begin{cases} 2y_1 - y_2 + y_3 + 2y_4 \leq 1, \\ 2y_1 + y_2 + y_3 + y_4 \geq 2, \\ -y_1 + 4y_2 - 2y_3 - 2y_4 \geq 3, y_i \geq 0, (i = 1, 2, 3, 4). \end{cases}$$

11.4. Виды математических моделей двойственных задач

На основании рассмотренных несимметричных и симметричных двойственных задач можно заключить, что математические модели пары двойственных задач могут следующие варианты.

Несимметричные задачи

(1) Исходная задача

$$Z_{\min} = CX$$

$$AX = A_0$$

$$X \geq 0.$$

Двойственная задача

$$f_{\max} = Y A_0$$

$$YA \leq C$$

(2) Исходная задача

$$Z_{\max} = CX$$

$$AX = A_0$$

$$X \geq 0$$

Двойственная задача

$$f_{\max} = Y A_0$$

$$YA \geq C$$

Симметричные задачи

(3) Исходная задача

$$Z_{\min} = CX$$

$$AX \geq A_0$$

$$X \geq 0$$

Двойственная задача

$$f_{\max} = Y A_0$$

$$YA \leq C$$

$$Y \geq 0$$

(4) Исходная задача

$$Z_{\max} = CX$$

$$AX \leq A_0$$

$$X \geq 0$$

Двойственная задача

$$f_{\min} = Y A_0$$

$$YA \geq C$$

$$Y \geq 0$$

Таким образом, прежде чем записать двойственную задачу для данной исходной, систему ограничений исходной задачи необходимо привести к соответствующему виду. Запишем, например, математическую модель двойственной задачи для заданной исходной.

Найти минимальное значение линейной функции $Z = 2x_1 + x_2 + 5x_3$ при ограничениях

$$\begin{cases} x_1 - x_2 - x_3 \leq 4, \\ x_1 - 5x_2 + x_3 \geq 5, \\ 2x_1 - x_2 + 3x_3 \geq 6, \quad x_j \geq 0 \quad (j = 1, 2, 3). \end{cases}$$

Рассматриваемая задача относится к симметричным двойственным задачам на отыскание минимального значения линейной функции. Для того чтобы можно было записать двойственную задачу, ее модель должна иметь вид (3). Переход осуществляется умножением первого неравенства на -1 .

Исходная задача:

$Z_{\min} = 2x_1 + x_2 + 5x_3$ при ограничениях

$$\begin{cases} -x_1 + x_2 + x_3 \geq -4, \\ x_1 - 5x_2 + x_3 \geq 5, \\ 2x_1 - x_2 + 3x_3 \geq 6, \quad x_j \geq 0 \quad (j = 1, 2, 3). \end{cases}$$

Двойственная задача:

$f_{\max} = -4y_1 + 5y_2 + 6y_3$ при ограничениях

$$\begin{cases} -y_1 + y_2 + 2y_3 \leq 2, \\ y_1 - 5y_2 - y_3 \leq 1, \\ y_1 + y_2 + 3y_3 \leq 5, \quad y_j \geq 0 \quad (i = 1, 2, 3). \end{cases}$$

После постановки задачи рекомендуется использовать симплекс-метод для ее решения.

Контрольные вопросы

1. Как формулируется общая задача линейного программирования? Что такое план, опорный план?
2. Как геометрически интерпретируется задача линейного программирования?
3. В чем заключается симплекс-метод решения задачи линейного программирования?
4. Как строятся опорные планы?
5. Как строится оптимальный план? Каковы условия оптимальности?
6. Каков алгоритм симплекс-метода?
7. В чем суть методов решения транспортной задачи?
8. В чем заключается понятие двойственности в линейном программировании?

Упражнения

1. Найти минимальное значение линейной функции:

а) $L = 4x_1 + 5x_2 + 6x_3$ при ограничениях

$$\begin{cases} x_1 + x_2 + x_3 \geq 5, \\ x_1 - x_2 + x_3 \geq 1, \\ x_1 - x_2 - 4x_3 \leq -3, \\ x_1 - x_2 + 8x_3 \geq 4, \end{cases} \quad x_j \geq 0 \ (j = 1, 2, 3);$$

б) $L = 2y_1 + 4y_2 + 12y_4$ при ограничениях

$$\begin{cases} y_1 + 2y_2 + y_3 + 4y_4 \geq 10, \\ 2y_1 + y_2 - 2y_3 + 3y_4 \geq 4, \end{cases} \quad y_i \geq 0 \ (i = 1, 2, 3, 4).$$

2. Применяя симплекс-метод, решить задачи:

а) $L_{\min} = 2x_1 - 2x_2 + 3x_3 + 4x_4$ при ограничениях

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 \leq 2, \\ 2x_1 - x_2 + 2x_3 - 3x_4 \geq 3, \\ 3x_1 + 4x_2 - 5x_3 + 2x_4 \leq 4, \end{cases} \quad x_j \geq 0 \ (j = 1, 2, 3, 4);$$

б) $L_{\min} = 5x_1 - x_2 - 4x_3$ при ограничениях

$$\begin{cases} -x_2 + 2x_3 \geq 9, \\ -x_1 + x_2 \geq 1, \\ x_1 + x_2 - 3x_3 \geq 8, \\ x_1 - x_3 \leq 4, \end{cases} \quad x_j \geq 0 \ (j = 1, 2, 3);$$

в) $L_{\min} = 2x_1 + 3x_2 + \frac{5}{2}x_3$ при ограничениях

$$\begin{cases} 2x_1 + x_2 + 3x_3 \geq 6, \\ 2x_1 + 4x_2 + 3x_3 \geq 16, \\ 3x_1 + 4x_2 + 2x_3 \geq 12, \end{cases} \quad x_j \geq 0 \ (j = 1, 2, 3);$$

г) $L_{\min} = 5x_1 + 6x_2 + 3x_3$ при ограничениях

$$\begin{cases} 3x_1 + x_2 + 2x_3 \geq 9, \\ x_1 + 2x_2 + 2x_3 \geq 8, \\ x_1 + 6x_2 \geq 12. \end{cases} \quad x_j \geq 0 \ (j = 1, 2, 3).$$

Записать для них оптимальные планы.

Г л а в а 4

ОСНОВЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА

§1. Множества и операции над ними

1.1. Основные определения

Мы уже немного говорили о множествах (см. §1 главы 1). Система, семейство, совокупность — эти термины можно считать синонимами слова “множество”. *Множество* можно определить как совокупность объектов, объединенных по определенному признаку. Например, множество всех наборов товаров рыночной корзины, множество зрителей в данном театре, множество студентов определенного высшего учебного заведения, совокупность отличников и “хорошистов” в некоторой школе, совокупность коммерческих банков, имеющих уставный фонд не менее 100 млн. рублей. Множество может содержать конечное или бесконечное число объектов.

Пусть X и Y — два множества. Тогда между ними можно определить следующие соотношения. Если оба множества состоят из одних и тех же элементов, то они совпадают, что соответствует записи $X = Y$. Если все элементы множества X содержатся в множестве Y , то X целиком содержится в Y , или $X \subset Y$ (X является *подмножеством* множества Y). Если ни один элемент множества X не содержится в Y , то и само множество X не содержится в Y , или $X \not\subset Y$.

В математике используется понятие *пустого множества*, обозначаемого символом $\emptyset$. Это множество, в котором не содержится ни одного элемента, поэтому оно является подмножеством любого множества.

Введем также понятие суммы множеств и их пересечения.

ОПРЕДЕЛЕНИЕ. *Суммой* или *объединением множеств* X и Y называется совокупность элементов, входящих как в множество X , так и в множество Y (обладающих либо свойством множества X , либо свойством множества Y); сумма этих множеств обозначается $X \cup Y$.

Например, пусть X — множество государственных предприятий с годовым оборотом не менее S денежных единиц, а Y — множество негосударственных предприятий с тем же нижним порогом годового оборота. Тогда $X \cup Y$ будет множеством всех предприятий с указанным нижним ограничением S . Отметим, что добавление пустого множества $\emptyset$ к любому множеству X не меняет этого множества, т.е. $X \cup \emptyset = X$.

ОПРЕДЕЛЕНИЕ. *Пересечением множеств X и Y (или их общей частью) является совокупность элементов, входящих как в множество X , так и в множество Y ; это множество обозначается $X \cap Y$ (заштрихованная область на рис. 4.1).*

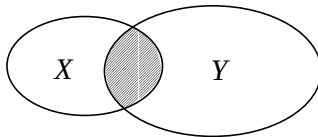


Рис. 4.1

Например, если X — это множество предприятий с годовым оборотом T не менее s , а Y — совокупность предприятий с годовым оборотом не более S , причем $s < S$, то в пересечение $X \cap Y$ войдут объекты с годовым оборотом T , удовлетворяющим неравенству $s \leq T \leq S$.

Отсутствие элементов со свойствами множеств X и Y одновременно означает, что пересечение этих множеств представляет собой пустое множество $\emptyset$.

ОПРЕДЕЛЕНИЕ. *Разностью множеств X и Y называется множество Z , содержащее все элементы множества X , не содержащиеся в Y ; эта разность обозначается $Z = X \setminus Y$.*

В общем случае сложение и пересечение определяются для любого конечного числа множеств путем последовательного попарного проведения соответствующих операций.

При записи математических выражений целесообразно применять логическую символику. Вместо выражения “любое x из множества X ” употребляют запись $\forall x \in X$, где перевернутая латинская буква $\forall$ (квантор общности) взята от начала английского слова Any — любой. Аналогично вместо выражения “существует элемент x из множества X ” кратко пишут: $\exists x \in X$, где перевернутая латинская буква $\exists$ (квантор существования) является начальной в английском слове Existence — существование.

Будем говорить, что множество вещественных чисел X ограничено сверху (снизу), если существует такое число d , при котором для $\forall x \in X$ выполняется неравенство $x \leq d$ ($x \geq d$). Число d тогда называется *верхней (нижней) гранью* множества X . Множества, ограниченные снизу и сверху, называются *ограниченными*. Любой конечный промежуток ограничен. Интервалы $(a, +\infty)$ и $(-\infty, b)$ представляют собой множества, ограниченные соответственно снизу (сверху), но не ограниченные сверху (снизу). Вся числовая прямая не ограничена ни снизу, ни сверху.

Любое ограниченное сверху (снизу) множество имеет бесконечное число верхних (нижних) граней. Действительно, если число d является верхней гранью множества X , то и любое число $d_1 > d$, согласно определению верхней грани, также будет верхней гранью этого множества. Наименьшая верхняя грань множества X , ограниченного сверху, называется *точной верхней гранью* этого множества; она обозначается символом $\sup X$. Наибольшая нижняя грань ограниченного снизу множества X называется *точной нижней гранью* этого множества и обозначается символом $\inf X$. Эти символы образованы от латинских слов: *supremum* — наивысший и *infimum* — наинизший.

Приведем примеры:

1. Пусть $X = (a, b)$. В таком случае числа a и b являются соответственно точными нижней и верхней гранями множества X , т.е. $a = \inf X$, $b = \sup X$.

2. Пусть $X = (-\infty, b)$. Тогда нижних граней (в том числе и точной нижней грани) множество X не имеет, а число b является его точной верхней гранью: $b = \sup X$.

ОПРЕДЕЛЕНИЕ. Множество X называется *конечным*, если существует такое натуральное число n (называемое числом элементов множества X), что между элементами множества X и элементами множества $\{1, 2, \dots, n-1, n\}$ можно установить взаимно однозначное соответствие.

ОПРЕДЕЛЕНИЕ. *Счетным* называется множество, элементам которого можно поставить в соответствие ряд натуральных чисел.

Например, последовательность $\left\{\frac{1}{n}\right\}$, где $n \in N$ — счетное множество, так как можно установить взаимно однозначное соответствие между элементами последовательности и множеством натуральных чисел:

$$\begin{array}{ccccccc} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \dots & \frac{1}{n} & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ 1 & 2 & 3 & 4 & \dots & n & \dots \end{array}$$

1.2. Арифметическое m -мерное пространство

Назовем m -мерной точкой систему из m вещественных чисел: $M(x_1, x_2, \dots, x_m)$. Сами числа $x_1, x_2, \dots, x_m$ являются координатами этой точки M . Множество всех мыслимых m -мерных точек составляют m -мерное пространство, которое иногда называют *арифметическим пространством*.

ОПРЕДЕЛЕНИЕ. *Расстоянием* MM' между двумя m -мерными точками $M(x_1, x_2, \dots, x_m)$ и $M'(x'_1, x'_2, \dots, x'_m)$ называется

$$MM' = M'M = \sqrt{\sum_{i=1}^m (x'_i - x_i)^2} = \sqrt{(x'_1 - x_1)^2 + (x'_2 - x_2)^2 + \dots + (x'_m - x_m)^2}.$$

Если взять еще одну точку $M''(x''_1, x''_2, \dots, x''_m)$, то для расстояний $MM', M'M'', MM''$ выполняется неравенство:

$$MM'' \leq MM' + M'M'',$$

напоминающее теорему: сторона треугольника не превосходит суммы двух других сторон.

ОПРЕДЕЛЕНИЕ. Точка $M'(x'_1, x'_2, \dots, x'_m)$ называется *внутренней точкой множества* μ в m -мерном пространстве, если она принадлежит множеству M вместе с некоторой достаточно малой ее окрестностью.

ОПРЕДЕЛЕНИЕ. Множество, целиком состоящее из внутренних точек, будем называть *открытым множеством*.

ОПРЕДЕЛЕНИЕ. Точка M_0 называется *точкой сгущения* множества μ , если в каждой ее окрестности содержится хотя одна точка множества μ , отличная от M_0 .

ОПРЕДЕЛЕНИЕ. Точки сгущения для открытого множества, не принадлежащего ему, называются *пограничными точками* этого множества. Пограничные точки в их совокупности образуют *границу множества*.

ОПРЕДЕЛЕНИЕ. Открытое множество вместе с его границей называется *замкнутым множеством*.

Замкнутому множеству принадлежат уже все его точки сгущения.

§2. Бюджетные и производственные множества

2.1. Определение бюджетного множества

В § 1 главы 1 мы уже говорили о *пространстве товаров* $\mathbf{G}$ и ввели понятие цены набора $X - C(X)$.

Вспоминая, что такое линейное пространство, мы видим, что пространство товаров $\mathbf{G}$ представляет собой часть арифметического линейного пространства $\mathbf{R}^n$ — так называемый *неотрицательный октант*, $\mathbf{G} = \{X \in \mathbf{R}^n : X \geq 0\}$. Поэтому при работе с пространством товаров можно использовать структуру линейного пространства (соблюдая некоторые естественные ограничения). Так, для любого $X \in \mathbf{G}$ подмножество $L_X = \{\lambda X : 0 \leq \lambda\}$ называется *лучом*, проходящим через X ; для любых двух точек X, Y любая точка $\alpha X + \beta Y \in \mathbf{G}$ называется их *линейной комбинацией*, а множество $[X, Y] = \{\alpha X + \beta Y : \alpha, \beta \geq 0, \alpha + \beta = 1\}$ называется *отрезком*, соединяющим X и Y .

ОПРЕДЕЛЕНИЕ. Подмножество $W \subseteq \mathbf{G}$ назовем *выпуклым*, если вместе с любыми $X, Y \in W$ весь соединяющий их отрезок лежит в W .

В зависимости от цены набора все пространство товаров можно разбить на непересекающиеся классы.

С обыденной точки зрения каждый товар должен быть желателен для покупателя и должен обладать определенной потребительской полезностью. Это свойство товаров выражается в некоторой мере через цены на них.

Пусть вектор цен есть P . Зафиксируем какую-нибудь денежную сумму Q и назовем ее *доходом*.

ОПРЕДЕЛЕНИЕ. Множество наборов товаров стоимости не более Q при данных ценах P называется *бюджетным множеством* B ; множество наборов товаров стоимости, равной Q , называется *границей* $\bar{G}$ этого бюджетного множества.

Бюджетное множество и его граница зависят от цен и дохода, так что точнее их было бы обозначить $B(P, Q)$ и $\bar{G}(P, Q)$.

Бюджетное множество и его границу можно определить как

$$\underline{B}(P, Q) = \{(x_1, \dots, x_n) : x_1, \dots, x_n \geq 0, p_1 x_1 + \dots + p_n x_n \geq Q\},$$

$$\bar{G}(P, Q) = \{(x_1, \dots, x_n) : x_1, \dots, x_n \geq 0, p_1 x_1 + \dots + p_n x_n \geq Q\}$$

или с помощью векторных неравенств и равенств:

$$B(P, Q) = \{X : X \geq 0, PX \leq Q\}, \quad \bar{G}(P, Q) = \{X : X \geq 0, PX = Q\}.$$

Проблема рационального поведения потребителя заключается в решении вопроса о том, какое количество товаров или услуг он хочет и может приобрести при заданных ценах и его доходе.

Выбор потребителем некоторого набора товаров во многом зависит от его вкусов, желаний. Потребитель различает наборы товаров, предпочитая один набор товаров другому. Запись $X \preceq Y$ означает, что потребитель предпочитает набор X набору Y либо не делает между ними различий. Из-за последнего обстоятельства отношение " $\preceq$ " называется *слабым предпочтением*. Оно формирует еще два отношения: *отношение равноценности* (или *безразличия*) $X \sim Y$, если и только если $X \preceq Y$ и $Y \preceq X$, и *отношение предпочтения* (или *строгого предпочтения*) — $X \prec Y$, если и только если $X \preceq Y$ и неверно, что $X \sim Y$. Какими же свойствами обладают эти три отношения?

Математики называют отношение *рефлексивным*, если $X \preceq Y$ для всякого X ; *симметричным*, если $X \preceq Y$ влечет, что и $Y \preceq X$; *транзитивным*, если $X \preceq Y$ и $Y \preceq Z$ влечет $X \preceq Z$; *совершенным* (или *полным*), если для любых двух наборов X, Y либо $X \preceq Y$, либо $Y \preceq X$.

АКСИОМА. *Отношение слабого отношения рефлексивно, транзитивно и совершенно. 2. Отношение равноценности рефлексивно, симметрично и транзитивно. 3. Отношение предпочтения транзитивно. 4. Для любого $X \in \mathbf{G}$ множество предпочтительности $P_X = \{Y : X \preceq Y\}$ выпукло. 5. Каждый товар желателен для индивида: если $X \preceq Y$, то и $Y \preceq X$, а если к тому же $X \neq Y$ (т.е. $x_i < y_i$ для некоторого i), то $Y \prec X$.*

Подчеркнем, что это именно аксиома, выражающая фундаментальные свойства системы предпочтений индивида, вообще говоря, живого человека. Рефлексивность означает, что любой набор товара равноценен сам по себе, а совершенность означает, что индивид способен сравнить по привлекательности любые два набора товаров. Выпуклость означает, что лучше иметь комбина-

цию товаров, пусть в меньших количествах, чем какой-то один из этих товаров (лучше иметь немножко соли, сахара, кофе, хлеба, чем одну только соль, один сахар, кофе, хлеб, хотя бы и в большем количестве).

Свойство транзитивности не совсем очевидно, однако интуитивно осознается потребителем при выборе товаров.

Отношение равноценности рефлексивно, симметрично и транзитивно. Любое отношение, обладающее этими тремя свойствами, называется *эквивалентностью*. Любая эквивалентность на множестве разбивает это множество на непересекающиеся подмножества, называемые *классами эквивалентности*. Итак, отношение равноценности является эквивалентностью и разбивает пространство товаров на непересекающиеся подмножества, называемые классами равноценности (или безразличия), а в случае двух или трех товаров эти классы называются *линиями* (или *поверхностями*) *равноценности* (или *безразличия*). Каждый отдельный класс равноценности состоит из наборов товаров, одинаково привлекательных для потребителя, — он не отдает предпочтения ни одному из этих наборов. При этом каждый набор из пространства товаров попадает в какой-нибудь класс равноценности, именно в тот, где собраны наборы, одинаково ценные с ним (для данного индивида).

2.2. Производственные множества и их свойства

Напомним, что наша экономика работает в пространстве товаров $G = \{X = (x_1, \dots, x_n): x_1, \dots, x_n \geq 0\}$. Ранее рассматривалось пространство товаров, состоящее из неотрицательных n -мерных векторов. Рассмотрим теперь вектор T размерности n , первые m компонентов которого неположительны: $x_1, \dots, x_m \leq 0$, а последние $(n - m)$ компонентов — неотрицательны: $x_{m+1}, \dots, x_n \geq 0$. Вектор $X = (x_1, \dots, x_m)$ назовем вектором затрат, а вектор $Y = (x_{m+1}, \dots, x_n)$ — вектором выпуска. Сам же вектор $T = (X, Y)$ назовем *вектором затрат-выпуска*, или *технологией*.

По своему смыслу технология (X, Y) есть способ переработки ресурсов в готовую продукцию: “смешав” ресурсы в количестве X , получим продукцию в размере Y .

Каждый конкретный производитель характеризуется некоторым множеством τ технологий, которое называется *производственным множеством*.

Производственное множество отражает широту возможностей производителя: чем оно больше, тем шире эти возможности.

Считается, что производственное множество должно удовлетворять следующим условиям:

а) оно замкнуто — это означает, что если вектор T затрат-выпуска сколь угодно точно приближается к векторам из τ , то и T принадлежит τ ;

б) $\tau \cap \Omega = \{0\}$, где $\Omega = \{T : T \geq 0\}$; это формальное условие понятно с содержательной и экономической точек зрения: если бы в τ был вектор $T \geq 0$, $T \neq 0$, то это означало бы, что что-то можно производить, ничего не затрачивая;

в) $\tau \cap (-\tau) = \{0\}$, т.е. если $T \in \tau$, $T \neq 0$, то $-T \notin \tau$ — нельзя поменять местами затраты и выпуск, т.е. производство — необратимый процесс;

г) множество выпукло; это предположение ведет, кроме всего прочего, к уменьшению отдачи от перерабатываемых ресурсов с ростом объемов производства (к увеличению норм расхода затрат на готовую продукцию). В частности, предположение о выпуклости ведет к уменьшению производительности труда с ростом объема производства.

§3. Числовые последовательности

3.1. Числовые последовательности и операции над ними

Числовые последовательности представляют собой бесконечные упорядоченные множества чисел. Примерами последовательностей могут служить: последовательность всех членов арифметической (геометрической) прогрессии, последовательность приближенных значений $\sqrt{2}$, т.е. $x_1 = 1$, $x_2 = 1,4$, $x_3 = 1,41$, ..., последовательность периметров правильных n -угольников, вписанных в данную окружность ($n \geq 3$). Уточним понятие числовой последовательности.

ОПРЕДЕЛЕНИЕ. Если каждому числу n из натурального ряда чисел $1, 2, 3, \dots, n, \dots$ поставлено в соответствие вещественное число x_n , то множество вещественных чисел

$$x_1, x_2, x_3, \dots, x_n, \dots \quad (4.1)$$

называется **числовой последовательностью**, или просто **последовательностью**.

Числа $x_1, x_2, x_3, \dots, x_n, \dots$ будем называть **элементами** (или **членами**) последовательности (4.1), символ x_n — **общим элементом** (или **общим членом**) последовательности, а число n — его **номером**. Сокращенно последовательность (4.1) будем обозначать символом $\{x_n\}$. Например, символ $\{1/n\}$ обозначает последовательность чисел

$$1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$$

Иными словами, под последовательностью можно понимать бесконечное множество занумерованных элементов или множество пар чисел (n, x_n) , в которых первое число последовательно принимает значения $1, 2, 3, \dots$. Последовательность считается заданной, если указан способ получения любого ее элемента. Например, формула $x_n = 1 + (-1)^n$ определяет последовательность $0, 2, 0, 2, \dots$.

Геометрически последовательность вещественных чисел изображается на числовой оси в виде последовательности точек, координаты которых равны соответствующим членам последова-

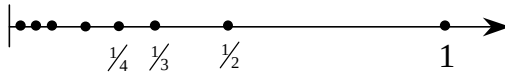


Рис. 4.2

тельности. На рис. 4.2 изображена последовательность $\{x_n\} = \{1/n\}$ на числовой прямой.

Арифметические действия над числовыми последовательностями.

Пусть даны последовательности $\{x_n\}$ и $\{y_n\}$.

Суммой последовательностей назовем последовательность $\{x_n + y_n\} = x_1 + y_1, x_2 + y_2, \dots, x_n + y_n, \dots$, *разностью* — последовательность $\{x_n - y_n\} = x_1 - y_1, x_2 - y_2, \dots, x_n - y_n, \dots$.

Произведением последовательности $\{x_n\}$ на число t назовем последовательность $\{tx_n\} = tx_1, tx_2, tx_3, \dots, tx_n, \dots$.

Произведением последовательностей назовем последовательность $\{x_n y_n\} = x_1 y_1, x_2 y_2, x_3 y_3, \dots, x_n y_n, \dots$.

Частным назовем последовательность

$$\left\{ \frac{x_n}{y_n} \right\} = \frac{x_1}{y_1}, \frac{x_2}{y_2}, \dots, \frac{x_3}{y_3}, \dots, \frac{x_n}{y_n}, \dots,$$

если все члены последовательности $\{y_n\}$ отличны от нуля.

3.2. Ограниченные и неограниченные последовательности

Поскольку числовые последовательности являются числовыми множествами, то естественно ввести понятие их ограниченности, как это было сделано для множеств в §2 главы 4.

ОПРЕДЕЛЕНИЕ. Последовательность вещественных чисел $\{x_n\}$ называется *ограниченной сверху (снизу)*, если существует такое число M (число m), что любой элемент x_n этой последовательности удовлетворяет неравенству $x_n \leq M$ ($x_n \geq m$).

ОПРЕДЕЛЕНИЕ. Последовательность x_n называется *ограниченной*, если она ограничена и сверху и снизу, т.е. существуют такие числа m и M , что любой элемент $\{x_n\}$ этой последовательности удовлетворяет неравенствам $m \leq x_n \leq M$.

Пусть $A = \max\{|m|, |M|\}$, тогда условие ограниченности последовательности можно записать в форме $|x_n| \leq A$.

ОПРЕДЕЛЕНИЕ. Последовательность $\{x_n\}$ называется **неограниченной**, если для любого положительного числа A существует элемент x_n этой последовательности, удовлетворяющий неравенству $|x_n| > A$, (т.е. либо $x_n > A$, либо $x_n < -A$).

Из приведенных определений следует, что все элементы ограниченной сверху последовательности принадлежат промежутку $(-\infty, M]$, а элементы последовательности, ограниченной снизу, лежат на промежутке $[m, +\infty)$. Неограниченная последовательность может быть ограничена лишь с одной из сторон, например сверху (снизу). Все элементы ограниченной последовательности принадлежат отрезку $[m, M]$.

Приведем примеры ограниченных и неограниченных последовательностей:

1) последовательность $-1, -2, -3, \dots, -n, \dots$ не ограничена снизу, но ограничена сверху;

2) последовательность $1, 1/2, 1/3, \dots, 1/n, \dots$ ограничена, так как любой элемент $x_n = 1/n$ этой последовательности удовлетворяет неравенствам $0 \leq x_n \leq 1$;

3) последовательность $-1, 2, -3, \dots, (-1)^n n \dots$ неограниченная. Действительно, каково бы ни было число $A > 0$, среди элементов x_n этой последовательности найдутся элементы, для которых будет выполняться неравенство $|x_n| > A$.

3.3. Предел последовательности. Бесконечно большие и бесконечно малые последовательности

ОПРЕДЕЛЕНИЕ. Число a называется **пределом последовательности** $\{x_n\}$, если для любого положительного числа ε существует такой номер $N(\varepsilon)$, что при всех $n > N(\varepsilon)$ выполняется неравенство

$$|x_n - a| < \varepsilon. \quad (4.2)$$

Если предел последовательности равен числу a , это записывается так:

$$\lim_{n \rightarrow \infty} x_n = a, \text{ или } x_n \rightarrow a \text{ при } n \rightarrow \infty. \quad (4.3)$$

ОПРЕДЕЛЕНИЕ. Последовательность $\{x_n\}$ называется *бесконечно большой*, если для любого положительного числа A существует такой номер $N(A)$, что при $n > N(A)$ (для всех элементов последовательности с номерами $n > N(A)$) выполняется неравенство $|x_n| > A$.

Очевидно, что любая бесконечно большая последовательность является неограниченной. Обратное утверждение неверно: неограниченная последовательность может и не быть бесконечно большой. Например, неограниченная последовательность $1, 2, 1, 3, 1, 4, \dots, 1, n, \dots$ не является бесконечно большой, так как при $A > 1$ неравенство $|x_n| > A$ не выполняется для всех элементов x_n с нечетными номерами.

ОПРЕДЕЛЕНИЕ. Последовательность $\{\alpha_n\}$ называется *бесконечно малой*, если для любого положительного числа ε существует такой номер $N(\varepsilon)$, что при $n > N(\varepsilon)$ выполняется неравенство $|\alpha_n| < \varepsilon$.

3.4. Основные свойства бесконечно малых последовательностей

□ **ТЕОРЕМА 4.1.** Сумма и разность двух бесконечно малых последовательностей являются бесконечно малыми последовательностями. ■

СЛЕДСТВИЕ. Сумма любого конечного числа бесконечно малых последовательностей есть бесконечно малая последовательность.

□ **ТЕОРЕМА 4.2.** Произведение двух бесконечно малых последовательностей есть бесконечно малая последовательность. ■

СЛЕДСТВИЕ. Произведение конечного числа бесконечно малых последовательностей есть бесконечно малая последовательность.

□ **ТЕОРЕМА 4.3.** Произведение бесконечно малой последовательности на ограниченную последовательность является бесконечно малой последовательностью. ■

СЛЕДСТВИЕ. Произведение бесконечно малой последовательности на постоянное число является бесконечно малой последовательностью.

§4. Сходящиеся последовательности

4.1. Понятие сходящейся последовательности

ОПРЕДЕЛЕНИЕ. Последовательность $\{x_n\}$ называется *сходящейся*, если существует такое число a , что в любой его ε -окрестности находятся все элементы этой последовательности, начиная с некоторого номера, зависящего от ε .

Таким образом, последовательность, имеющая конечный предел, является *сходящейся*, а последовательность, не имеющая конечного предела, является *расходящейся*.

Интересно дать геометрическую интерпретацию сказанного определения. Поскольку последовательность представляет собой бесконечное множество чисел, то, если она сходится, в любой ε -окрестности точки a на числовой прямой находится бесконечное число точек — элементов этой последовательности, тогда как вне ε -окрестности остается конечное число элементов (рис. 4.3). Поэтому предел последовательности часто называют *точкой сгущения*.

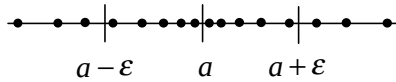


Рис. 4.3

З а м е ч а н и я:

1. Бесконечно большая последовательность не имеет конечного предела. Часто говорят, что она имеет *бесконечный предел*; записывается это в следующем виде:

$$\lim_{n \rightarrow \infty} x_n = \infty.$$

Если при этом, начиная с некоторого номера, все члены последовательности положительны (отрицательны), то пишут

$$\lim_{n \rightarrow \infty} x_n = +\infty \quad \left(\lim_{n \rightarrow \infty} x_n = -\infty \right).$$

2. Принято считать, что всякая бесконечно малая последовательность сходится и имеет своим пределом число $a = 0$.

3. Предел последовательности, у которой $x_1 = x_2 = \dots = c$, где c — некоторая константа, равен c , т.е. $\lim_{n \rightarrow \infty} c = c$.

4.2. Основные свойства сходящихся последовательностей

□ ТЕОРЕМА 4.4. Если все элементы бесконечно малой последовательности $\{x_n\}$ равны одному и тому же числу C , то $C = 0$. ■

□ ТЕОРЕМА 4.5. Сходящаяся последовательность имеет только один предел. ■

□ ТЕОРЕМА 4.6. Предел суммы двух (или нескольких) сходящихся последовательностей равен сумме пределов этих последовательностей. ■

□ ТЕОРЕМА 4.7. Предел произведения двух (или нескольких) сходящихся последовательностей равен произведению пределов этих последовательностей. ■

□ ТЕОРЕМА 4.8. Предел частного двух сходящихся последовательностей равен частному пределов этих последовательностей (при условии, что предел делителя не равен нулю). ■

▷ ПРИМЕР 4.1. Найти предел $\lim_{n \rightarrow \infty} \frac{3n^2 + 2n + 4}{4n^2 + n - 3}$.

Р е ш е н и е. При $n \rightarrow \infty$ числитель и знаменатель дроби стремятся к бесконечности, т.е. применить сразу теорему о пределе частного нельзя, так как она предполагает существование конечных пределов последовательностей. Преобразуем данную последовательность, разделив числитель и знаменатель на n^2 . Применяя теоремы о пределе частного и пределе суммы, последовательно найдем:

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{3n^2 + 2n + 4}{4n^2 + n - 3} &= \lim_{n \rightarrow \infty} \frac{3 + 2/n + 4/n^2}{4 + 1/n - 3/n^2} = \frac{\lim_{n \rightarrow \infty} (3 + 2/n + 4/n^2)}{\lim_{n \rightarrow \infty} (4 + 1/n - 3/n^2)} = \\ &= \frac{\lim_{n \rightarrow \infty} 3 + \lim_{n \rightarrow \infty} (2/n) + \lim_{n \rightarrow \infty} (4/n^2)}{\lim_{n \rightarrow \infty} 4 + \lim_{n \rightarrow \infty} (1/n) - \lim_{n \rightarrow \infty} (3/n^2)} = \frac{3 + 0 + 0}{4 + 0 - 0} = \frac{3}{4}. \blacktriangleright\end{aligned}$$

▷ ПРИМЕР 4.2. Найти предел последовательности

$$\{x_n\} = \left\{ \frac{\sqrt{n} \cos n}{n+1} \right\}.$$

Решение. Здесь, как и в предыдущем примере, числитель и знаменатель не имеют конечных пределов, поэтому сначала необходимо выполнить соответствующие преобразования. Поделив числитель и знаменатель на n , получим

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} \frac{(1/\sqrt{n}) \cos n}{1 + 1/n} = \frac{\lim_{n \rightarrow \infty} (1/\sqrt{n}) \lim_{n \rightarrow \infty} \cos n}{\lim_{n \rightarrow \infty} 1 + \lim_{n \rightarrow \infty} (1/n)}.$$

Поскольку в числителе стоит произведение бесконечно малой последовательности на ограниченную последовательность, то в силу теоремы 4.3 окончательно получим

$$\lim_{n \rightarrow \infty} x_n = \frac{0}{1+0} = 0. \blacktriangleright$$

▷ ПРИМЕР 4.3. Найти предел последовательности

$$\{x_n\} = \{\sqrt{n+1} - \sqrt{n}\}.$$

Решение. Здесь применить непосредственно теорему о пределе суммы (разности) последовательностей нельзя, так как конечных пределов слагаемых в формуле для $\{x_n\}$ не существует. Сначала умножим и разделим формулу для $\{x_n\}$ на сопряженное выражение $\sqrt{n+1} + \sqrt{n}$, затем каждое слагаемое числителя и знаменателя разделим на $\sqrt{n}$:

$$\begin{aligned}
\lim_{n \rightarrow \infty} x_n &= \lim_{n \rightarrow \infty} \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \lim_{n \rightarrow \infty} \frac{n+1-n}{\sqrt{n+1} + \sqrt{n}} = \\
&= \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n+1} + \sqrt{n}} = \frac{\lim_{n \rightarrow \infty} (1/\sqrt{n})}{\lim_{n \rightarrow \infty} (\sqrt{1+1/n}) + 1} = \frac{0}{1+1} = 0. \blacktriangleright
\end{aligned}$$

§5. Монотонные последовательности

5.1. Определение монотонных последовательностей

Монотонные последовательности классифицируются по соотношению между соседними элементами. Такие соотношения могут иметь один из следующих видов:

$$x_n < x_{n+1}, \quad (4.4)$$

$$x_n \leq x_{n+1}, \quad (4.5)$$

$$x_n > x_{n+1}, \quad (4.6)$$

$$x_n \geq x_{n+1}. \quad (4.7)$$

ОПРЕДЕЛЕНИЕ. Последовательность $\{x_n\}$ называется: *возрастающей*, если при $\forall n$ для ее элементов выполняется условие (4.4); *неубывающей*, если выполняется условие (4.5); *убывающей*, если выполняется условие (4.6); *невозрастающей*, если выполняется условие (4.7).

Все эти последовательности называются *монотонными*. Последовательности типов (4.4) и (4.6) называются *строго монотонными*. Монотонные последовательности всегда ограничены либо сверху (невозрастающие и убывающие последовательности), либо снизу (возрастающие и неубывающие последовательности).

Приведем примеры монотонных последовательностей:

1) последовательность $1, \frac{1}{3}, \frac{1}{5}, \dots, \frac{1}{(2n-1)}, \dots$ — убывающая и ограниченная сверху числом 1 и снизу числом 0 ($0 < \frac{1}{2n-1} \leq 1$ при $\forall n$);

2) последовательность $1, 1, \frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{n}, \frac{1}{n}, \dots$ — невозрастающая и ограниченная сверху и снизу;

3) последовательность $1, 2, 3, \dots, n, \dots$ — возрастающая ограниченная снизу и неограниченная сверху;

4) последовательность $1, 1, 2, 2, \dots, n, n, \dots$ — неубывающая ограниченная снизу и неограниченная сверху;

5) последовательность $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots, \frac{n}{(n+1)}, \dots$ — ограниченная (снизу числом $\frac{1}{2}$, сверху — числом 1);

б) последовательность $-1, \frac{1}{2}, -\frac{1}{3}, \dots, (-1)^n/n, \dots$ — ограниченная, но не монотонной, поскольку для всех n не выполняется ни одно из условий (4.4)–(4.7).

5.2. Признак сходимости монотонной последовательности. Формула сложных процентов

□ ТЕОРЕМА 4.9. *Монотонная ограниченная последовательность сходится.* ■

Применяя теорему 4.9, можно доказать сходимость последовательности $\{x_n\}$, общий член которой выражается формулой

$$x_n = \left(1 + \frac{1}{n}\right)^n.$$

Предел этой последовательности называют числом e , т.е.

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n. \quad (4.8)$$

Число e — иррациональное число (бесконечная непериодическая дробь), которое часто встречается в математике, экономике и др., для которого найдено большое количество десятичных знаков: $e = 2,71828182845\dots$. Если внимательно его рассмотреть, то можно заметить повторение после запятой четырех цифр: 1828 — это год рождения Льва Николаевича Толстого¹.

Известно, что **формула сложных процентов** имеет вид

$$Q = Q_0 \left(1 + \frac{p}{100}\right)^n, \quad (4.9)$$

где Q — сумма вкладов по истечении n периодов; Q_0 — первоначальная сумма вклада в банк; p — процент начисления за определенный период (месяц, год); n — количество периодов хранения вклада.

¹ Прим. д. ф.-м. н., проф. В. Я. Якубова.

Иначе $i = \frac{P}{100}$ называется *процентной ставкой*, а коэффициент $(1 + i)^n$, входящий в правую сторону соотношения (4.9), называется *коэффициентом наращивания*.

В § 1 главы 2 мы говорили о простой процентной ставке (2.1). Если ввести наши обозначения, то формула (2.1) будет выглядеть так:

$$Q = Q_0(1 + i_{np}n), \quad (2.1^*)$$

где i_{np} — простая процентная ставка.

Посмотрим, как изменяется количество денег на банковском счете со временем при начислении дохода по одинаковым, простой и сложной, процентным ставкам. Сравнение формул (2.1*) и (4.9) показывает, что в первый год деньги растут быстрее, если доход начисляется по простой процентной ставке. К концу первого года доходы, полученные по обеим ставкам, одинаковы. В дальнейшем деньги растут быстрее (причем, существенно), если начисление дохода происходит по сложной процентной ставке.

Формулы типа (4.9) используются также в демографических расчетах (прирост народонаселения) и в экономических прогнозах (увеличение валового национального продукта).

Пусть первоначальный депозит Q_0 помещен в банк под $p = 100\%$ годовых, тогда через год сумма депозита составит $2Q_0$. Предположим, что через полгода счет закрыт с результатом

$$Q_1 = Q_0 \left(1 + \frac{1}{2}\right) = \frac{3}{2} Q_0, \text{ и эта сумма вновь помещена в качестве депозита в том же банке. В конце года депозит будет составлять}$$

$$Q_2 = Q_0 \left(1 + \frac{1}{2}\right)^2 = 2,25 Q_0. \text{ Будем уменьшать срок размещения депозита в банке при условии его последующего размещения после изъятия. При ежеквартальном повторении этих операций депозит в конце года составит}$$

$$Q_3 = Q_0 \left(1 + \frac{1}{3}\right)^3 \approx 2,37 Q_0.$$

Если банк настолько щедр, что позволяет повторять операцию “изъятие-размещение” в течение года сколько угодно раз, то при ежемесячном манипулировании сумма за год составит

$$Q_{12} = Q_0 \left(1 + \frac{1}{12}\right)^{12} \approx 2,61 Q_0; \text{ при ежедневном посещении банка}$$

$$Q_{365} = Q_0 \left(1 + \frac{1}{365}\right)^{365} \approx 2,714 Q_0; \text{ при ежечасной активности}$$

$$Q_{8720} = Q_0 \left(1 + \frac{1}{8720}\right)^{8720} \approx 2,718 Q_0 \text{ и т.д. Нетрудно заметить,}$$

что последовательность значений увеличения первоначального вклада $\{q_n\} = \{Q_n/Q_0\}$ совпадает с последовательностью, пределом которой является число e при $n \rightarrow \infty$. Таким образом, доход, который можно получить при непрерывном использовании процентов на проценты, может составить за год не более чем

$$\lim_{n \rightarrow \infty} \frac{(Q_n - Q_0)}{Q_0} \cdot 100\% = (e - 1) \cdot 100\% \approx 172\%.$$

В общем случае, если p — процент начисления и год разбит на n частей, то через t лет сумма депозита достигнет величины

$$Q_n = Q_0 \left(1 + \frac{r}{n}\right)^{nt},$$

где $r = p/100$. Это выражение можно преобразовать:

$$Q_n = Q_0 \left[\left(1 + \frac{r}{n}\right)^{\frac{n}{r}} \right]^{rt}.$$

Если ввести новую переменную $m = \frac{n}{r}$, то при $n \rightarrow \infty$ получим

$m \rightarrow \infty$, или

$$\lim_{n \rightarrow \infty} Q_n = \lim_{m \rightarrow \infty} Q_0 \left[\left(1 + \frac{1}{m} \right)^m \right]^{rt} = Q_0 e^{rt}. \quad (4.10)$$

Расчеты, выполненные по этой формуле, называют вычислениями *по непрерывным процентам*¹.

▷ ПРИМЕР 4.4. Пусть темп инфляции составляет 1% в день. Насколько уменьшится первоначальная сумма через полгода?

Р е ш е н и е. Применение формулы сложных процентов дает

$$Q = Q_0 \left(1 - \frac{1}{100} \right)^{182},$$

где Q_0 — первоначальная сумма; 182 — число дней в полугодии.

Преобразовав это выражение, получим

$$Q = Q_0 \left[\left(1 - \frac{1}{100} \right)^{-100} \right]^{-182/100} \approx Q_0 / e^{1,82} \approx Q_0 / 6,$$

т.е. инфляция уменьшит первоначальную сумму примерно в 6 раз. ►

Сравним начисление дохода по простой, сложной, непрерывной процентным ставкам.

1. Предположим, что вы инвестируете² 1000 рублей по годовой ставке с непрерывным начислением, равной 10% ($r = 0,10$), на один год ($t = 1$). Стоимость в конце года (по формуле (4.10)) равна $1000 \cdot e^{0,10} = 1105$ рублей. Говоря другими словами, инвестирование под 10% годовых с непрерывным начислением процентов в течение года равнозначно инвестированию с годовым начисле-

¹ Когда мы говорим о непрерывных выплатах, мы предполагаем, что деньги образуют постоянный поток, подобно воде, льющейся из водопроводного крана. На самом деле такого быть не может. Финансовые менеджеры *делают вид*, что начисление процентных выплат происходит непрерывно, а не раз в час, день или неделю, потому что, во-первых, это упрощает расчеты и, во-вторых, это дает очень точное приближение к чистой приведенной стоимости (см. § 3 главы 3) частых выплат.

² Инвестировать — делать долгосрочные вложения своего капитала.

нием по простой процентной ставке 10,5%. (Предлагаем читателю убедиться в этом самостоятельно).

2. Теперь предположим, что вы вкладываете 1000 руб. по ставке 11% с непрерывным начислением процента ($r = 0,11$) в течение одного года ($t = 1$). Следовательно, стоимость инвестиций на конец года составляет $1000 \cdot e^{0,11} = 1116$ руб. Если применить к первоначальной сумме формулу (4.9) с начислением сложного процента по ставке 11,6% *один раз* в год, то можно заметить, что результат равнозначен результату инвестирования по годовой ставке 11% с непрерывным начислением процента.

3. И наконец, допустим, что вы инвестируете 1000 руб. по годовой ставке 11% с непрерывным начислением процента ($r = 0,11$) на два года ($t = 2$). Конечная стоимость инвестиций будет $1000 \cdot e^{0,22} = 1246$ руб.

Принцип непрерывного начисления процента имеет особое значение при планировании долгосрочных вложений, когда более обоснованно считать, что поток денежных средств поступает равномерно в течение года, а не в конце года.

§6. Понятие функции

6.1. Определение функциональной зависимости. Способы задания функции

Пусть X — некоторое числовое множество, и каждому числу $x \in X$ по какому-либо закону f поставлено в соответствие одно и только одно число $y \in Y$. Тогда будем говорить, что определена функциональная зависимость y от x по закону $y = f(x)$. При этом x называют *независимой переменной* (или *аргументом*), y — *зависимой переменной* или *функцией*, множество X — областью определения (существования) функции. Множество Y — множество всех значений функции, которые она принимает, когда x пробегает все множество X , — называется *множеством значений* (областью изменения) функции. Совокупность точек координатной плоскости xOy , удовлетворяющих уравнению $y = f(x)$, называется *графиком* этой функции.

Для обозначения функции и независимой переменной могут быть использованы и другие буквы. Примеры записи функций: $y = y(x)$, $y = F(x)$, $y = g(x)$.

Задать функцию — значит указать закон, по которому, согласно определению, каждому значению аргумента из области определения ставится в соответствие (вычисляется) значение зависимой переменной y . Существует три основных способа задания функций: табличный, аналитический и графический.

Пример *табличного* способа задания функции:

x	1	3	4
y	0	-1	-2

Аналитический способ состоит в задании связи между аргументом и функцией в виде формул. Следует подчеркнуть, что функция может определяться и набором формул: разным участкам области определения функции соответствуют разные формулы.

Примеры *аналитического* и *графического* способов задания функций:

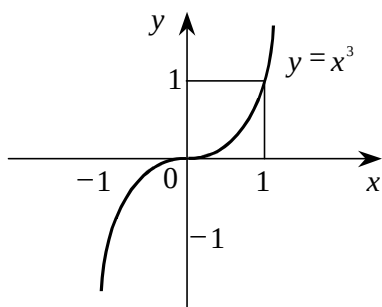


Рис. 4.4

1. $y = x^3$. Эта функция задана на всей прямой $-\infty < x < \infty$. Множество значений этой функции — тоже бесконечная числовая прямая $-\infty < y < \infty$. График функции называется *кубической параболой* (рис. 4.4).

2. $y = \sqrt{1 - x^2}$. Функция задана на отрезке $[-1, 1]$, множество ее значений — отрезок $[0, 1]$. Это половина окружности, лежащая в верхней полуплоскости (рис. 4.5).

$$3. y = \text{sign } x = \begin{cases} 1, & \text{если } x > 0, \\ 0, & \text{если } x = 0, \\ -1, & \text{если } x < 0. \end{cases}$$

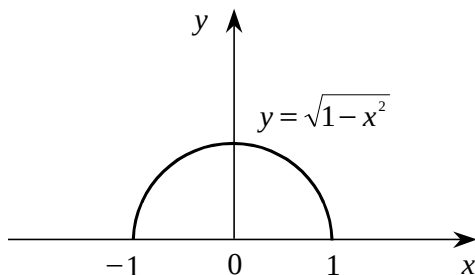


Рис. 4.5

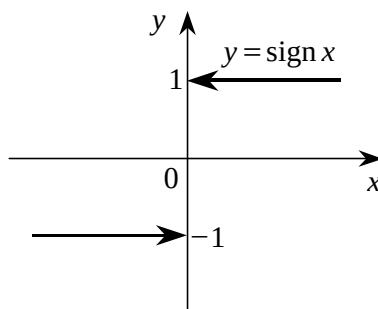


Рис. 4.6

Термин *sign* — происходит от латинского *signum* — знак. Функция задана на всем бесконечном промежутке $(-\infty, \infty)$, а множество ее значений состоит из трех чисел: $-1, 0, 1$ (рис. 4.6). Стрелки означают, что полупрямые не достигают точек на оси ординат, так как при $x = 0$ значение функции определено по другому соответствию.

4. $y = [x]$ — целая часть значений аргумента. Функция задана для всех вещественных значений x , а множество ее значений состоит из целых чисел (рис. 4.7).

$$5. y = \begin{cases} x & \text{при } x \leq 0, \\ \sqrt{x} & \text{при } x > 0. \end{cases}$$

Функция определена на бесконечном промежутке $(-\infty, \infty)$, область ее изменения — также бесконечный промежуток $(-\infty, \infty)$ (рис. 4.8).

6. Функция Дирихле

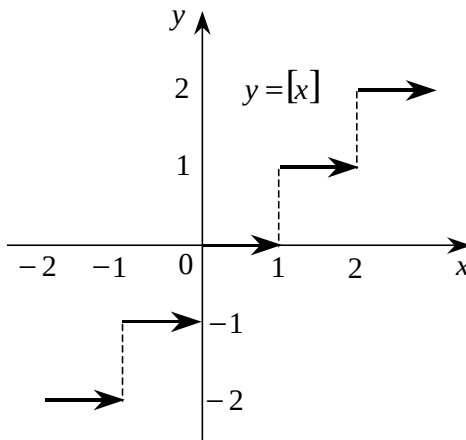


Рис. 4.7

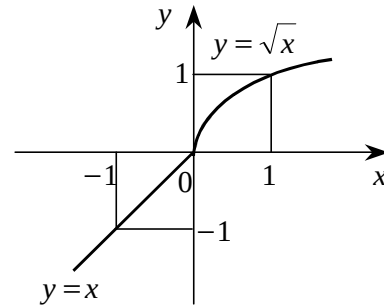


Рис. 4.8

$$y = \begin{cases} 0, & \text{если } x \text{ — иррациональное число,} \\ 1, & \text{если } x \text{ — рациональное число.} \end{cases}$$

Функция Дирихле определена на всей числовой оси, а множество ее значений состоит из двух чисел: 0 и 1. Изобразить ее графически невозможно.

6.2. Некоторые свойства функций

По аналогии с монотонными последовательностями (§ 5 главы 4) вводятся понятия монотонных функций.

Говорят, что функция $f(x)$ *не убывает* (*не возрастает*) на множестве X , если для любых $x_1, x_2 \in X$, таких, что $x_1 < x_2$, справедливо неравенство $f(x_1) \leq f(x_2)$ ($f(x_1) \geq f(x_2)$).

ОПРЕДЕЛЕНИЕ. Неубывающие и невозрастающие функции называются **монотонными** функциями.

ОПРЕДЕЛЕНИЕ. Если для $\forall x_1, x_2 \in X : x_1 < x_2$ справедливо неравенство $f(x_1) < f(x_2)$ ($f(x_1) > f(x_2)$), то функция $f(x)$ называется **возрастающей** (**убывающей**) на множестве X . Такие функции называются также **строго монотонными**.

Приведем несколько примеров.

1. Функция $f(x) = \operatorname{sign} x$ является неубывающей на всей числовой прямой (см. рис. 4.6).

2. Функция $f(x) = x^2$ является убывающей на промежутке $(-\infty; 0)$ и возрастающей на промежутке $[0; +\infty)$.

3. Функция $f(x) = x^3$ является возрастающей на всей числовой прямой (см. рис. 4.4).

ОПРЕДЕЛЕНИЕ. Если на некотором промежутке X определена функция $z = \varphi(x)$ с множеством значений Z , и на этом множестве определена функция $y = f(z)$, то функция $y = f[\varphi(x)]$ называется **сложной функцией** (или **суперпозицией функций**), а переменная z -промежуточной переменной сложной функции.

Например, $y = \sin \sqrt{1-x}$ — сложная функция, определенная на полубесконечном интервале $(-\infty; 1]$, так как $y = f(z) = \sin z$, $z = \varphi(x) = \sqrt{1-x}$.

Позже мы снова встретимся с понятиями монотонности (подпараграф 14.2) и сложной функции.

6.3. Производственная функция и функция полезности

Приведем примеры использования функций в экономике.

Определение функции полезности. Мы в подпараграфе 2.1. главы 4 рассмотрели бюджетное множество, на котором строится система предпочтений потребителя.

Система предпочтений индивида указывает, какой из двух наборов предпочтительнее для него. Во многих случаях, однако, удобно оценивать привлекательность набора товаров количеством, т.е. приписать каждому набору X из пространства товаров $\mathbf{G}$ какое-то число $u(X)$. Получается функция $u: \mathbf{G} \rightarrow \mathbf{R}$. Главное требование к такой функции, чтобы она отражала отношение предпочтения (слабого) на $\mathbf{G}$, т.е. удовлетворяла условиям:

$u(X) \leq u(Y)$, если и только если $X \preceq Y$;

$u(X) = u(Y)$, если и только если $X \sim Y$;

$u(X) < u(Y)$, если и только если $X \prec Y$.

Такая функция называется *функцией полезности*; и она постоянна на каждом классе равноценности, т.е. ее можно представить как функцию, “перебирающую” классы равноценности в сторону все большего предпочтения наборов товаров.

Работать с функцией полезности гораздо удобнее, чем с системой предпочтений. Однако, если на систему предпочтений не накладывать никаких ограничений, кроме рассмотренных ранее (а именно: транзитивность, совершенность и рефлексивность), функции полезности может и не существовать. Тем не менее при некоторых естественных условиях, наложенных на систему предпочтений, функция полезности существует.

Скажем, что система предпочтений *непрерывна*, если для всякого $X \in \mathbf{G}$ множество предпочтительности $P_X = \{Y \in \mathbf{G} : X \preceq Y\}$ и множество предпочтительности $N_X = \{Z \in \mathbf{G} : X \preceq Z\}$ замкнуты (рис. 4.9). Кстати, как легко видеть, пересечение этих двух множеств есть класс равноценности.

Чуть позже (в § 10 главы 4) мы сформулируем условия, при которых функция полезности существует.

Сейчас отметим, что функция полезности, если она существует, не определяется единственным образом. Главное требование к функции полезности — она должна отражать систему предпочтений. Поэтому, если $u(X)$ — функция полезности, то

$$v(X) = k u(X) + b \quad (4.11)$$

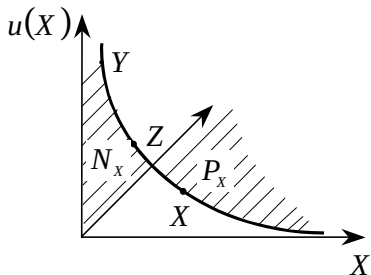


Рис. 4.9

(где $k > 0$, а b — константа) также есть функция полезности; и вообще, если $y = f(x)$ — произвольная строго возрастающая числовая функция на $\mathbf{R}$, то сложная функция $f(u(X))$ также есть функция полезности.

Свойства функции полезности. Основные свойства функции полезности вытекают из ее связи с системой предпочтений.

Сформулируем свойства желательности каждого товара:

$$X \leq Y \Rightarrow X \preceq Y; \quad (4.12)$$

$$X < Y, X \neq Y \Rightarrow X \prec Y. \quad (4.13)$$

Для функции полезности отсюда следует, что она *неубывающая*, т.е. $X \leq Y$ влечет $u(X) \leq u(Y)$, а если к тому же $X \neq Y$, то $u(X) < u(Y)$.

Определение производственной функции. Мы в подпараграфе 2.2 главы 4 рассмотрели производственное множество (заштриховано на рис. 4.10), на котором существует производственная функция. Итак, пространство затрат m -мерно. Каждой точке пространства затрат $X = (x_1, \dots, x_m)$ соответствует единственный максимальный выпуск, произведенный при использовании этих затрат. Эта связь и называется *производственной функцией*. Однако обычно производственную функцию понимают не столь ограничительно, и всякую функциональную связь между затратами и выпуском считают производственной функцией.

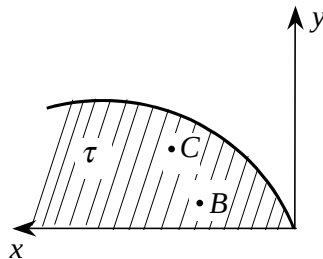


Рис. 4.10

Свойства производственной функции. Предполагается, что производственная функция f удовлетворяет двум аксиомам.

АКСИОМА 1. Существует подмножество пространства затрат, называемое экономической областью E , в которой увеличение

любого вида затрат не приводит к уменьшению выпуска. Таким образом, если X_1, X_2 — две точки этой области, то $X_1 \geq X_2$ влечет $f(X_1) \geq f(X_2)$.

АКСИОМА 2. Существует выпуклое подмножество S экономической области, для которой подмножества $\{X \in S: f(X) \geq a\}$ выпуклы для всех $a \geq 0$.

Чуть позже (в § 17) мы вернемся к экономическому обоснованию этих аксиом.

§7. Предел функции

7.1. Предел функции в бесконечности

С понятием предела числовой последовательности $\{y_n = f(n)\}$ тесно связано понятие предела функции $y = f(x)$ в бесконечности. Если в первом случае переменная n , возрастая, принимает лишь целые значения, то во втором случае переменная x , изменяясь, принимает любые значения.

ОПРЕДЕЛЕНИЕ. Число A называется *пределом функции* $y = f(x)$ *при* x , *стремящемся к бесконечности*, если для любого сколь угодно малого положительного числа $\varepsilon > 0$ найдется такое положительное число $\delta > 0$ (зависящее от ε ; $\delta = \delta(\varepsilon)$), что для всех x , таких, что $|x| > \delta$, верно неравенство

$$|f(x) - A| < \varepsilon. \quad (4.14)$$

Этот предел функции обозначается $\lim_{x \rightarrow \infty} f(x) = A$, или $f(x) \rightarrow A$, $x \rightarrow \infty$.

▷ ПРИМЕР 4.5. Доказать, что $\lim_{x \rightarrow \infty} \frac{5x+1}{x} = 5$.

Решение. Для любого $\varepsilon > 0$ неравенство (4.14), т.е. в данном случае $\left| \frac{5x+1}{x} - 5 \right| < \varepsilon$, выполняется при $\frac{1}{|x|} < \varepsilon$, или $|x| > \frac{1}{\varepsilon}$.

Итак, для любого $\varepsilon > 0$ существует такое число $\delta = \frac{1}{\varepsilon} > 0$, что для всех x , таких, что $|x| > \delta$, будет верно неравенство $|f(x) - 5| < \varepsilon$, где $f(x) = \frac{5x+1}{x}$, а это и означает, что $\lim_{x \rightarrow \infty} f(x) = 5$. ►

7.2. Предел функции в точке

Пусть функция $y = f(x)$ задана в некоторой окрестности точки x_0 , кроме, быть может, самой точки x_0 .

ОПРЕДЕЛЕНИЕ. Число A называется *пределом функции $f(x)$ при x , стремящемся к x_0 (или в точке x_0)*, если для любого, даже сколь угодно малого положительного числа $\varepsilon > 0$ найдется такое положительное число $\delta > 0$ (зависящее от ε , $\delta = \delta(\varepsilon)$), что для всех x , не равных x_0 и удовлетворяющих условию

$$|x - x_0| < \delta, \quad (4.15)$$

выполняется неравенство

$$|f(x) - A| < \varepsilon. \quad (4.16)$$

Этот предел функции обозначается $\lim_{x \rightarrow x_0} f(x) = A$ или $f(x) \rightarrow A$ при $x \rightarrow x_0$.

▷ ПРИМЕР 4.6. Доказать, что $\lim_{x \rightarrow 1} (2x + 3) = 5$,

Р е ш е н и е. Пусть $\varepsilon = 0,1$. Тогда неравенство (4.16), т.е. в данном случае $|(2x + 3) - 5| < 0,1$, будет выполняться при $|x - 1| < 0,05$. Аналогично при $\varepsilon = 0,01$ то же неравенство (4.16) будет верно при $|x - 1| < 0,005$. Таким образом, для $\forall \varepsilon > 0$ неравенство $|(2x + 3) - 5| < \varepsilon$ будет выполняться при $|x - 1| < \frac{\varepsilon}{2}$.

Итак, при $\forall \varepsilon > 0$ существует такое число $\delta = \frac{\varepsilon}{2}$ (для $\varepsilon = 0,1$ $\delta = 0,05$; для $\varepsilon = 0,01$ $\delta = 0,005$ и т. д.), что для всех $x \neq 1$ и удовлетворяющих условию $|x - 1| < \delta$ верно неравенство $|f(x) - 5| < \varepsilon$, где $f(x) = 2x + 3$; а это и означает, что $\lim_{x \rightarrow 1} f(x) = 5$. ►

7.3. Основные свойства предела функции

Пусть $f(x)$, $\varphi(x)$ — функции, для которых существуют пределы при $x \rightarrow x_0$ (или при $x \rightarrow \infty$): $\lim_{x \rightarrow x_0 (\infty)} f(x) = A$, $\lim_{x \rightarrow x_0 (\infty)} \varphi(x) = B$.

Сформулируем **основные теоремы о пределах**.

1. *Предел алгебраической суммы конечного числа функций равен такой же сумме пределов этих функций:*

$$\lim_{x \rightarrow x_0 (\infty)} [f(x) + \varphi(x)] = A + B.$$

2. *Предел произведения конечного числа функций равен произведению пределов этих функций.*

$$\lim_{x \rightarrow x_0 (\infty)} [f(x) \varphi(x)] = AB.$$

В частности, постоянный множитель можно выносить за знак предела, т.е.

$$\lim_{x \rightarrow x_0 (\infty)} (C f(x)) = CA.$$

3. *Предел частного двух функций равен частному пределов этих функций (при условии, что предел делителя не равен нулю):*

$$\lim_{x \rightarrow x_0 (\infty)} \frac{f(x)}{\varphi(x)} = \frac{A}{B} \quad (B \neq 0).$$

4. *Если $\lim_{u \rightarrow u_0} f(u) = A$, $\lim_{x \rightarrow x_0} \varphi(x) = u_0$, то предел сложной функции*

$$\lim_{x \rightarrow x_0} f(\varphi(x)) = A.$$

(Напомним, если заданы две функции $f(u)$ и $u = \varphi(x)$, то функция $f(\varphi(x))$ называется сложной функцией (см. подпараграф 6.2 главы 4).)

5. *Если $\exists$ некоторая окрестность точки x_0 (или при достаточно больших x): $f(x) \leq \varphi(x)$, то $A \leq B$, т.е.*

$$\lim_{x \rightarrow x_0 (\infty)} f(x) \leq \lim_{x \rightarrow x_0 (\infty)} \varphi(x).$$

§8. Два замечательных предела

8.1. Первый замечательный предел

□ ТЕОРЕМА 4.10. Предел функции $\frac{\sin x}{x}$ в точке $x = 0$ существует и равен единице, т.е.

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1. \blacksquare \quad (4.17)$$

▷ ПРИМЕР 4.7. Найти предел функции $\sin(ax)/(bx)$, $b \neq 0$, при $x \rightarrow 0$.

Решение. Преобразуем данную дробь так, чтобы в знаменателе был аргумент синуса; только тогда можно будет применить первый замечательный предел, поскольку при $x \rightarrow 0$ пределом ax также является нуль:

$$\lim_{x \rightarrow 0} \frac{\sin(ax)}{bx} = \lim_{x \rightarrow 0} \frac{\sin(ax)}{ax} \cdot \frac{a}{b} = \frac{a}{b} \lim_{x \rightarrow 0} \frac{\sin(ax)}{ax} = 1 \cdot \frac{a}{b} = \frac{a}{b}. \blacktriangleright$$

▷ ПРИМЕР 4.8. Найти $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2}$.

Решение. Преобразуем данную дробь, чтобы можно было применить теорему 4.10:

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} &= \lim_{x \rightarrow 0} \frac{2 \sin^2(x/2)}{x^2} = \lim_{x \rightarrow 0} \frac{1}{2} \left(\frac{\sin(x/2)}{x/2} \right)^2 = \\ &= \frac{1}{2} \lim_{x \rightarrow 0} \frac{\sin(x/2)}{x/2} \lim_{x \rightarrow 0} \frac{\sin(x/2)}{x/2} = \frac{1}{2} \cdot 1 \cdot 1 = \frac{1}{2}. \blacktriangleright \end{aligned}$$

▷ П Р И М Е Р 4.9. Найти $\lim_{x \rightarrow 0} \frac{\cos x - \cos 3x}{x^2}$.

Р е ш е н и е. Как и в первых двух примерах, преобразуем данную дробь, чтобы “подогнать” ее под первый замечательный предел:

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\cos x - \cos 3x}{x^2} &= \lim_{x \rightarrow 0} \frac{2 \sin 2x \sin x}{x^2} = \\ &= 2 \cdot 2 \cdot \lim_{x \rightarrow 0} \frac{\sin 2x}{2x} \lim_{x \rightarrow 0} \frac{\sin x}{x} = 4 \cdot 1 \cdot 1 = 4. \blacktriangleright \end{aligned}$$

8.2. Второй замечательный предел

□ Т Е О Р Е М А 4.11. Предел функции $f(x) = \left(1 + \frac{1}{x}\right)^x$ при $x \rightarrow \infty$ существует и равен числу e (ср. с (4.8)). ■

▷ П Р И М Е Р 4.10. Найти $\lim_{x \rightarrow 0} (1+x)^{1/x}$.

Р е ш е н и е. Применим здесь замену переменной, полагая $1/x = y$. Тогда $y \rightarrow \infty$ при $x \rightarrow 0$, т.е.

$$\lim_{x \rightarrow 0} (1+x)^{1/x} = \lim_{y \rightarrow \infty} (1+1/y)^y = e. \blacktriangleright$$

С Л Е Д С Т В И Е 1 (из теоремы 4.11). $\lim_{x \rightarrow 0} (1+x)^{1/x} = e$.

▷ П Р И М Е Р 4.11. Найти $\lim_{x \rightarrow \infty} (1+2/x)^x$.

Р е ш е н и е. Заменяем переменную, положив $y = \frac{x}{2}$. При $x \rightarrow \infty$ и $y \rightarrow \infty$, откуда

$$\begin{aligned} \lim_{x \rightarrow \infty} (1+2/x)^x &= \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x/2}\right)^x = \left[\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x/2}\right)^{x/2} \right]^{2/x \cdot x} = \\ &= \lim_{y \rightarrow \infty} (1+1/y)^{2y} = \lim_{y \rightarrow \infty} \left[(1+1/y)^y \right]^2 = \end{aligned}$$

$$= \lim_{y \rightarrow \infty} (1+1/y)^y \lim_{y \rightarrow \infty} (1+1/y)^y = e \cdot e = e^2. \blacktriangleright$$

СЛЕДСТВИЕ 2 (из теоремы 4.11). $\lim_{x \rightarrow \infty} (1+a/x)^x = e^a$.

▷ ПРИМЕР 4.12. Найти $\lim_{x \rightarrow 0} \left[\frac{1}{x} \log_a (1+x) \right], a \neq 1$.

Р е ш е н и е. Преобразуем дробь под знаком предела:

$$\begin{aligned} \lim_{x \rightarrow 0} \left[\frac{1}{x} \log_a (1+x) \right] &= \lim_{x \rightarrow 0} \log_a (1+x)^{1/x} = \\ &= \log_a \left[\lim_{x \rightarrow 0} (1+x)^{1/x} \right] = \log_a e = \frac{1}{\ln a} \end{aligned}$$

(см. теорему 4.15 и пример 4.10). $\blacktriangleright$

§9. Бесконечно малые и бесконечно большие функции

ОПРЕДЕЛЕНИЕ. Функция $f(x)$ называется *бесконечно малой функцией* (или просто *бесконечно малой*) в точке $x = a$, если предел ее в этой точке равен нулю: $\lim_{x \rightarrow a} f(x) = 0$.

Аналогично определяются бесконечно малые при $x \rightarrow \infty$, $x \rightarrow \pm\infty$, $x \rightarrow a+$ и $x \rightarrow a-$.

ОПРЕДЕЛЕНИЕ. Функция $f(x)$ называется *бесконечно большой функцией* (или просто *бесконечно большой*) в точке a , если для любой сходящейся к a последовательности $\{x_n\}$ значений аргумента соответствующая последовательность $\{f(x_n)\}$ значений функции является бесконечно большой последовательностью.

В этом случае пишут $\lim_{x \rightarrow a} f(x) = \infty$ ($\lim_{x \rightarrow a} f(x) = +\infty$ или $\lim_{x \rightarrow a} f(x) = -\infty$) и говорят, что функция имеет в точке a бесконечный предел ($+\infty$ или $-\infty$). По аналогии с конечными односторонними пределами определены и односторонние бесконечные пределы:

$$\lim_{x \rightarrow a+} f(x) = +\infty, \quad \lim_{x \rightarrow a-} f(x) = -\infty, \quad \lim_{x \rightarrow a-} f(x) = +\infty, \\ \lim_{x \rightarrow a+} f(x) = -\infty.$$

Аналогично определяются бесконечно большие функции при $x \rightarrow \infty$, $x \rightarrow +\infty$ и $x \rightarrow -\infty$.

▷ П Р И М Е Р 4.13. Найти:

$$\text{а) } \lim_{x \rightarrow 7} \frac{3x+5}{x-5}; \quad \text{б) } \lim_{x \rightarrow 5} \frac{3x+5}{x-5}; \quad \text{в) } \lim_{x \rightarrow \infty} \frac{\sin x}{x}; \quad \text{г) } \lim_{x \rightarrow 0} x \cos \frac{1}{x}.$$

Р е ш е н и е. а) На основании непрерывности функции в точке $x = 7$ искомый предел равен значению функции в этой точке, т.е.

$$\lim_{x \rightarrow \infty} \frac{3x+5}{x-5} = \frac{3 \cdot 7 + 5}{7 - 5} = 13.$$

б) При $x \rightarrow 5$ числитель $(3x + 5)$ стремится к $3 \cdot 5 + 5 = 20$ (т.е. является ограниченной функцией), а знаменатель $(x - 5)$ — к нулю (т.е. является бесконечно малой величиной); очевидно, их отно-

шение есть величина бесконечно большая, т.е. $\lim_{x \rightarrow 5} \frac{3x+5}{x-5} = \infty$.

в) $\lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$, ибо отношение ограниченной функции $\sin x$ ($|\sin x| \leq 1$) к бесконечно большой величине x (при $x \rightarrow \infty$) есть величина бесконечно малая.

г) $\lim_{x \rightarrow 0} x \cos \frac{1}{x} = 0$, так как произведение бесконечно малой ве-

личины x (при $x \rightarrow 0$) на ограниченную функцию $\cos \frac{1}{x}$ ($\left| \cos \frac{1}{x} \right| \leq 1$) есть величина бесконечно малая.

Заметим, что этот предел нельзя вычислять с помощью теоремы о пределе произведения (см. подпараграф 7.3 главы 4), поскольку

$\lim_{x \rightarrow 0} \cos \frac{1}{x}$ не существует (при $x \rightarrow 0$ аргумент косинуса $\frac{1}{x}$ изменяется вдоль числовой оси до бесконечности, при этом значения $\cos \frac{1}{x}$ колеблются от -1 до 1 и от 1 до -1 , не стремясь ни к какому числу (пределу)). ►

В рассмотренных примерах предел находился сразу: в виде числа или символа ∞ . Но чаще при вычислении пределов мы сталкиваемся с неопределенностями, когда результат вычисления предела не ясен: например, в случае отношения двух бесконечно малых функций (условное обозначение $\frac{0}{0}$) или бесконечно больших ($\frac{\infty}{\infty}$).

Кроме отмеченных неопределенностей вида $\frac{0}{0}$ и $\frac{\infty}{\infty}$ в математическом анализе рассматриваются также неопределенности вида $\infty - \infty$, $0 \cdot \infty$, 1^∞ , ∞^0 , 0^0 .

▷ ПРИМЕР 4.14. Найти:

$$\text{а) } \lim_{x \rightarrow 1} \frac{2x^2 - x - 1}{(x-1)^2}; \quad \text{б) } \lim_{x \rightarrow 2} \frac{\sqrt{x+2} - \sqrt{6-x}}{x^2 - 4}; \quad \text{в) } \lim_{x \rightarrow 1} \frac{\sqrt{x} - 1}{\sqrt[3]{x} - 1}.$$

Р е ш е н и е: а) Для раскрытия неопределенности вида $\frac{0}{0}$ разложим числитель на множители и сократим дробь на множитель $(x-1)$ (сокращение возможно, так как при $x \rightarrow 1$ выражение $(x-1)$ стремится к нулю, но не равно ему).

$$\lim_{x \rightarrow 1} \frac{2x^2 - x - 1}{(x-1)^2} = \left[\frac{0}{0} \right] = \lim_{x \rightarrow 1} \frac{(2x+1)(x-1)}{(x-1)^2} = \lim_{x \rightarrow 1} \frac{2x+1}{x-1} = \infty.$$

б) Для раскрытия неопределенности вида $\frac{0}{0}$ умножим числитель и знаменатель на выражение, сопряженное числителю:

$$\begin{aligned} \lim_{x \rightarrow 2} \frac{\sqrt{x+2} - \sqrt{6-x}}{x^2 - 4} &= \left[\frac{0}{0} \right] = \lim_{x \rightarrow 2} \frac{(\sqrt{x+2} - \sqrt{6-x})(\sqrt{x+2} + \sqrt{6-x})}{(x^2 - 4)(\sqrt{x+2} + \sqrt{6-x})} = \\ &= \lim_{x \rightarrow 2} \frac{2(x-2)}{(x-2)(x+2)(\sqrt{x+2} + \sqrt{6-x})} = \lim_{x \rightarrow 2} \frac{2}{(x+2)(\sqrt{x+2} + \sqrt{6-x})} = \frac{1}{8}. \end{aligned}$$

в) Для раскрытия неопределенности вида $\frac{0}{0}$ удобно предварительно сделать замену $t = \sqrt[6]{x}$ (тогда $\sqrt{x} = t^3$, $\sqrt[3]{x} = t^2$, $x \rightarrow 1$, $t \rightarrow 1$), а затем полученные многочлены разложить на множители:

$$\begin{aligned} \lim_{x \rightarrow 1} \frac{\sqrt{x} - 1}{\sqrt[3]{x} - 1} &= \left[\frac{0}{0} \right] = \lim_{t \rightarrow 1} \frac{t^3 - 1}{t^2 - 1} = \lim_{t \rightarrow 1} \frac{(t-1)(t^2 + t + 1)}{(t-1)(t+1)} = \\ &= \lim_{t \rightarrow 1} \frac{t^2 + t + 1}{(t+1)} = \frac{3}{2}. \end{aligned} \quad \blacktriangleright$$

▷ ПРИМЕР 4.15. Найти:

$$\text{а) } \lim_{x \rightarrow \infty} \frac{3x^2 + 2}{4x^5 + 3x + 1}; \quad \text{б) } \lim_{x \rightarrow \infty} \frac{a_1 x^n + b_1 x^{n-1} + \dots + k_1 x + l_1}{a_2 x^m + b_2 x^{m-1} + \dots + k_2 x + l_2},$$

$$a_1 \cdot a_2 \neq 0;$$

$$\text{в) } \lim_{x \rightarrow \infty} \frac{\sqrt[4]{x^9 + 1}}{x^2 + \sqrt{x}}; \quad \text{г) } \lim_{x \rightarrow \infty} \frac{4x + \sin x}{x - \cos x};$$

$$\text{д) } \lim_{x \rightarrow +\infty} f(x) \text{ и } \lim_{x \rightarrow -\infty} f(x), \text{ где } f(x) = \frac{2^{x+1} + 3^{x+1}}{2^x + 3^x}.$$

Р е ш е н и е: а) Имеем неопределенность вида $\frac{\infty}{\infty}$. Учитывая, что поведение числителя и знаменателя при $x \rightarrow \infty$ определяется членами с наибольшими показателями степеней (соответственно $3x^2$ и $4x^5$), разделим числитель и знаменатель на x с наибольшим показателем степени числителя и знаменателя. Воспользовавшись теоремы о пределах, получим

$$\lim_{x \rightarrow \infty} \frac{3x^2 + 2}{4x^5 + 3x + 1} = \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow \infty} \frac{\frac{3}{x^3} + \frac{2}{x^5}}{4 + \frac{3}{x^4} + \frac{1}{x^5}} = \frac{0 + 0}{4 + 0 + 0} = 0.$$

б) Используя тот же прием, что и в пункте “а”, можно показать, что

$$\lim_{x \rightarrow \infty} \frac{a_1 x^n + b_1 x^{n-1} + \dots + k_1 x + l_1}{a_2 x^m + b_2 x^{m-1} + \dots + k_2 x + l_2} = \left[\frac{\infty}{\infty} \right] = \begin{cases} 0, & n < m, \\ a_1 / a_2, & n = m, \\ \infty, & n > m, \end{cases}$$

т.е. предел отношения двух многочленов $\lim_{x \rightarrow \infty} \frac{P_n(x)}{Q_m(x)}$ равен 0, отношению коэффициентов при старших степенях x или ∞ , если показатель степени числителя n соответственно меньше, равен или больше степени знаменателя m .

Рекомендуем запомнить это правило.

в) Имеем неопределенность вида $\frac{\infty}{\infty}$. Здесь выражению в числителе условно можно приписать степень $n = \frac{9}{4}$, а в знаменателе степень $m = 2$; так как $n > m$, то на основании правила, сформулированного в пункте “б”, искомый предел равен ∞ .

Действительно, разделив и числитель и знаменатель на x^2 , получим

$$\lim_{x \rightarrow \infty} \frac{\sqrt[4]{x^9 + 1}}{x^2 + \sqrt{x}} = \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow \infty} \frac{\sqrt[4]{x + \frac{1}{x^8}}}{1 + \frac{1}{x\sqrt{x}}} = \infty.$$

г) Для раскрытия неопределенности вида $\frac{\infty}{\infty}$ разделим числитель и знаменатель на x , получим

$$\lim_{x \rightarrow \infty} \frac{4x + \sin x}{x - \cos x} = \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow \infty} \frac{4 + \frac{\sin x}{x}}{1 - \frac{\cos x}{x}} = \frac{4 + 0}{1 - 0} = 4,$$

так как $\lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$, $\lim_{x \rightarrow \infty} \frac{\cos x}{x} = 0$ (см. пример 4.13, пункт “в”).

д) При $x \rightarrow \infty$ имеем неопределенность вида $\frac{\infty}{\infty}$. Поведение числителя и знаменателя определяется в этом случае вторыми слагаемыми, которые возрастают быстрее первых. Разделив числитель и знаменатель на 3^x и воспользовавшись теоремами о пределах, получим

$$\lim_{x \rightarrow +\infty} \frac{2^{x+1} + 3^{x+1}}{2^x + 3^x} = \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow +\infty} \frac{2 \cdot \left(\frac{2}{3}\right)^x + 3}{\left(\frac{2}{3}\right)^x + 1} = \frac{0 + 3}{0 + 1} = 3,$$

поскольку $\lim_{x \rightarrow +\infty} \left(\frac{2}{3}\right)^x = 0$, так как функция $y = a^x$ при $0 < a < 1$

убывает, что известно из школьного курса математики.

При $x \rightarrow -\infty$ имеем неопределенность вида $\frac{0}{0}$. Поведение числителя и знаменателя в этом случае определяется первыми слагаемыми, которые убывают медленнее других. Разделив числитель и знаменатель на 2^x и воспользовавшись теоремами о пределах, получим

$$\lim_{x \rightarrow -\infty} \frac{2^{x+1} + 3^{x+1}}{2^x + 3^x} = \left[\frac{0}{0} \right] = \lim_{x \rightarrow -\infty} \frac{2 + 3\left(\frac{3}{2}\right)^x}{1 + \left(\frac{3}{2}\right)^x} = \frac{2+0}{1+0} = 2, \text{ так как}$$

$$\lim_{x \rightarrow -\infty} \left(\frac{3}{2}\right)^x = \lim_{t \rightarrow +\infty} \left(\frac{2}{3}\right)^t = 0. \blacktriangleright$$

▷ ПРИМЕР 4.16. Найти:

а) $\lim_{x \rightarrow \infty} (\sqrt{x^2 + 1} - \sqrt{x^2 - 1})$; б) $\lim_{x \rightarrow -\infty} (\sqrt{x^2 + 2} + x)$;

в) $\lim_{x \rightarrow 1} \left(\frac{1}{1-x} - \frac{2}{1-x^2} \right)$.

Р е ш е н и е: а) Для раскрытия неопределенности вида $\infty - \infty$ умножим и разделим выражение в скобках на сопряженное выражение, получим

$$\begin{aligned} \lim_{x \rightarrow \infty} (\sqrt{x^2 + 1} - \sqrt{x^2 - 1}) &= [\infty - \infty] = \\ \lim_{x \rightarrow \infty} \frac{(\sqrt{x^2 + 1} - \sqrt{x^2 - 1})(\sqrt{x^2 + 1} + \sqrt{x^2 - 1})}{\sqrt{x^2 + 1} + \sqrt{x^2 - 1}} &= \\ = \lim_{x \rightarrow \infty} \frac{2}{\sqrt{x^2 + 1} + \sqrt{x^2 - 1}} &= 0. \end{aligned}$$

б) При $x \rightarrow -\infty$ имеем неопределенность вида $\infty - \infty$, так как квадратный корень из неотрицательного числа всегда неотрицателен. Решение аналогично решению пункта “а”:

$$\begin{aligned} \lim_{x \rightarrow -\infty} \left(\sqrt{x^2 + 2} + x \right) &= [\infty - \infty] = \lim_{x \rightarrow -\infty} \frac{(\sqrt{x^2 + 2} + x)(\sqrt{x^2 + 2} - x)}{\sqrt{x^2 + 2} - x} = \\ &= \lim_{x \rightarrow -\infty} \frac{2}{\sqrt{x^2 + 2} - x} = 0. \end{aligned}$$

Обращаем внимание на то, что при $x \rightarrow -\infty$ в знаменателе нет неопределенности, так как он представляет сумму бесконечно больших положительных величин — величину, бесконечно большую.

$$\begin{aligned} \text{в) } \lim_{x \rightarrow 1} \left(\frac{1}{1-x} - \frac{2}{1-x^2} \right) &= [\infty - \infty] = \lim_{x \rightarrow 1} \frac{(1+x)-2}{1-x^2} = \\ &= \lim_{x \rightarrow 1} \frac{x-1}{(1-x)(1+x)} = \left[\frac{0}{0} \right] = \lim_{x \rightarrow 1} \left(-\frac{1}{(1+x)} \right) = -\frac{1}{2}. \quad \blacktriangleright \end{aligned}$$

Простейшие примеры с использованием первого замечательного предела были рассмотрены ранее. Рассмотрим несколько более сложные задачи.

▷ ПРИМЕР 4.17. Найти:

$$\text{а) } \lim_{x \rightarrow \infty} x \sin \frac{1}{x}; \quad \text{б) } \lim_{x \rightarrow 0} \frac{\sin 4x}{\sin 8x}.$$

$$\begin{aligned} \text{Решение. а) } \lim_{x \rightarrow \infty} x \sin \frac{1}{x} &= [\infty \cdot 0] = \lim_{x \rightarrow \infty} \frac{\sin \frac{1}{x}}{\frac{1}{x}} = \\ &= \left[\frac{0}{0} \right] = \lim_{y \rightarrow 0} \frac{\sin y}{y} = 1 \end{aligned}$$

(сделали замену $y = \frac{1}{x}$, $x \rightarrow \infty \Rightarrow y \rightarrow 0$).

$$\text{б) } \lim_{x \rightarrow 0} \frac{\sin 4x}{\sin 8x} = \left[\frac{0}{0} \right] = \frac{1}{2} \lim_{x \rightarrow 0} \left(\frac{\sin 4x}{4x} : \frac{\sin 8x}{8x} \right) = \frac{1}{2} (1:1) = \frac{1}{2}.$$

▷ ПРИМЕР 4.18. Найти:

$$\text{а) } \lim_{x \rightarrow 0} \frac{\sin x^6}{\sin^5 x}; \quad \text{б) } \lim_{x \rightarrow \infty} x(\ln(1+x) - \ln x).$$

Р е ш е н и е:

$$\text{а) } \lim_{x \rightarrow 0} \frac{\sin x^6}{\sin^5 x} = \left[\frac{0}{0} \right] = \lim_{x \rightarrow 0} \left[\left(x \cdot \frac{\sin x^6}{x^6} \right) : \left(\frac{\sin x}{x} \right)^5 \right] = (0 \cdot 1) : 1^5 = 0;$$

б) Имеем неопределенность вида $\infty(\infty - \infty)$. Получим неопределенность вида $\infty \cdot 0$, если с помощью свойств логарифма представим предел в виде

$$\begin{aligned} \lim_{x \rightarrow \infty} x(\ln(1+x) - \ln x) &= \lim_{x \rightarrow \infty} \left(x \ln \frac{1+x}{x} \right) = \lim_{x \rightarrow \infty} x \cdot \ln \left(1 + \frac{1}{x} \right) = \\ &= [\infty \cdot 0] = \lim_{x \rightarrow \infty} \ln \left(1 + \frac{1}{x} \right)^x. \end{aligned}$$

На основании непрерывности логарифмической функции перейдем к пределу под символом логарифма, т.е.

$$\lim_{x \rightarrow \infty} \ln \left(1 + \frac{1}{x} \right)^x = \ln \left[\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x} \right)^x \right] = \ln e = 1. \quad (4.18)$$

В следующем параграфе мы рассмотрим понятие непрерывной функции.

§ 10. Непрерывность функции

10.1. Определение непрерывной функции.

Точки разрыва. Асимптоты

Понятие непрерывности функции является фундаментальным в математическом анализе. Рассмотрим функцию $f(x)$, определенную в некотором промежутке X , и пусть x_1 — точка этого промежутка.

Напомним, что понятие $\lim_{x \rightarrow x_1} f(x)$ означает, что значения x_1 переменная x не принимает, оно может даже не принадлежать области определения функции. Однако особый интерес представляет именно случай, когда $\lim_{x \rightarrow x_1} f(x) = f(x_1)$.

ОПРЕДЕЛЕНИЕ. *Приращением* некоторой переменной величины называется разность между новым значением этой величины и ее прежним значением, т.е. $x_1 - x$.

Для обозначения приращения используется греческая буква Δ ; $\Delta x = x_1 - x$ — приращение величины x .

Прибавляя к значению переменной величины ее приращение, получим приращенное значение этой величины. Например, $x + \Delta x$ есть приращенное значение величины x .

Предположим, что y есть некоторая функция от аргумента x , т.е.
 $y = f(x)$. (4.19)

Дадим аргументу x приращение Δx , тогда y получит соответствующее приращение Δy :

$$y + \Delta y = f(x + \Delta x). \quad (4.20)$$

Из равенств (4.19) и (4.20) следует:

$$\Delta y = f(x + \Delta x) - f(x). \quad (4.21)$$

Приведем несколько определений непрерывности функции в соответствии с определениями предела функции в точке.

ОПРЕДЕЛЕНИЕ. Функция $f(x)$, определенная на множестве X , называется *непрерывной* при $x = x_1$ (или *непрерывной* в точке x_1), если:

- 1) функция определена при $x = x_1$ (т.е. $x_1 \in X$);
- 2) приращение функции в точке x_1 стремится к нулю, когда приращение аргумента $\Delta x_1 = x - x_1$ стремится к нулю, т.е.

$$\lim_{\Delta x_1 \rightarrow 0} [f(x_1 + \Delta x_1) - f(x_1)] = 0, \text{ где бесконечно малое приращение}$$

Δx_1 пробегает лишь те значения, для которых $f(x_1 + \Delta x_1)$ имеет смысл.

Другими словами, функция называется *непрерывной в данной точке*, если в этой точке бесконечно малому приращению аргумента соответствует бесконечно малое приращение функции. Если же это условие нарушено, то говорят, что при этом значении (или в этой точке) функция имеет *разрыв*, а эта точка называется *точкой разрыва* функции.

ОПРЕДЕЛЕНИЕ. Функция $f(x)$ называется *непрерывной* в точке x_1 , если для $\forall \varepsilon > 0 \exists \delta = \delta(\varepsilon, x_1) > 0$, такое, что $|f(x) - f(x_1)| = |f(x_1 + \Delta x_1) - f(x_1)| < \varepsilon$ при $|\Delta x_1| < \delta$.

ОПРЕДЕЛЕНИЕ. Функция $f(x)$ называется *непрерывной* в точке x_1 , если предел этой функции и ее значение в этой точке равны, т.е.

$$\lim_{x \rightarrow x_1} f(x) = f(x_1).$$

Так как $\lim_{x \rightarrow a} x = a$, то это равенство можно переписать в следующей форме: $\lim_{x \rightarrow x_1} f(x) = f(\lim_{x \rightarrow x_1} x)$.

УТВЕРЖДЕНИЕ. Приведенные выше определения непрерывности функции в точке x_1 эквивалентны.

ОПРЕДЕЛЕНИЕ. Функция $f(x)$ называется *непрерывной справа (слева)* в точке x_1 , если правый (левый) предел этой функции в точке x_1 равен значению функции в этой точке.

Символическая запись непрерывности функции справа и, соответственно, слева:

$$\lim_{x \rightarrow x_1+} f(x) = f(x_1), \text{ или } f(x_1 + 0) = f(x_1),$$

$$\lim_{x \rightarrow x_1-} f(x) = f(x_1), \text{ или } f(x_1 - 0) = f(x_1).$$

Если функция непрерывна в точке x_1 слева и справа, то она непрерывна в этой точке.

ОПРЕДЕЛЕНИЕ. Если функция $f(x)$ непрерывна в каждой точке промежутка X , то она называется **непрерывной** на всем этом промежутке X .

□ **ТЕОРЕМА 4.12.** Если множество значений монотонно возрастающей (убывающей) функции $f(x)$, которые она принимает, когда x изменяется в промежутке X , содержится в некотором промежутке Y и заполняет его сплошь, то функция $f(x)$ в промежутке X непрерывна. ■

□ **ТЕОРЕМА 4.13.** Если две функции $f(x)$ и $g(x)$ определены в одном и том же промежутке X и обе непрерывны в точке x_0 , то в той же точке будут непрерывны и функции $f(x) \pm g(x)$, $f(x) \cdot g(x)$,

$$\frac{f(x)}{g(x)}, (g(x_0) \neq 0). \quad \blacksquare$$

□ **ТЕОРЕМА 4.14.** Пусть функция $z = \varphi(x)$ непрерывна в точке x_0 , а функция $y = f(z)$ непрерывна в точке $z_0 = \varphi(x_0)$. Тогда сложная функция $y = f[\varphi(x)]$ непрерывна в точке x_0 . ■

Иначе говоря, непрерывная функция от непрерывной функции есть функция также непрерывная; сложная функция, состоящая из непрерывных функций, непрерывна.

В подпараграфе 6.3 главы 4 мы говорили о функции полезности. Теперь можно сформулировать условия, при которых функция полезности существует.

□ УТВЕРЖДЕНИЕ (теорема Дебре). *Если система предпочтений непрерывна, то существует непрерывная функция полезности.* ■

Надо отметить, что функция полезности, если она существует, не определяется единственным образом. Поэтому, если $u(X)$ — функция полезности, а $y = f(x)$ — произвольная строго возрастающая числовая функция на $\mathbf{R}$, то сложная функция $f(u(X))$ также есть функция полезности.

□ ТЕОРЕМА 4.15. *Если функция $y = f(z)$ непрерывна в точке z_0 и $z_0 = \lim_{z \rightarrow z_0} \varphi(x)$, то $\lim_{z \rightarrow z_0} f(z(x)) = f(\lim_{z \rightarrow z_0} \varphi(x)) = f(z_0)$.* ■

ОПРЕДЕЛЕНИЕ. Точка x_0 разрыва функции $f(x)$ называется **точкой разрыва I рода**, если существуют конечные односторонние пределы функции: $\lim_{x \rightarrow x_0 - 0} f(x) = f(x_0 - 0)$, $\lim_{x \rightarrow x_0 + 0} f(x) = f(x_0 + 0)$.

При этом функция $f(x)$ не обязательно должна быть определена в точке x_0 , т.е. $f(x_0)$ может не существовать.

Величина $\delta = f(x_0 + 0) - f(x_0 - 0)$ называется *скачком* функции $f(x)$ в точке x_0 .

ОПРЕДЕЛЕНИЕ. Точка x_0 называется **точкой разрыва II рода**, если в этой точке не существует хотя бы один из односторонних пределов функции $f(x)$ или хотя бы одно из односторонних предельных значений бесконечно.

На рис. 4.11 приведен случай, когда оба предела — бесконечность:

$$\lim_{x \rightarrow x_0 - 0} f(x) = +\infty; \quad \lim_{x \rightarrow x_0 + 0} f(x) = +\infty.$$

ОПРЕДЕЛЕНИЕ. **Кусочно-непрерывной функцией** называется такая функция, которая на любом конечном интервале имеет конечное число точек разрыва I рода.

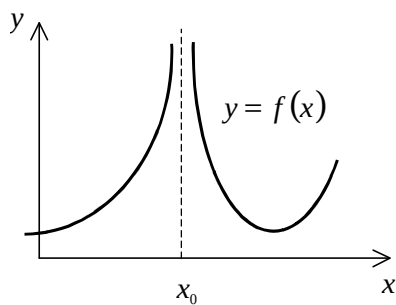


Рис. 4.11

Примеры кусочно-непрерывных функций приведены на рис. 4.11, 4.12.

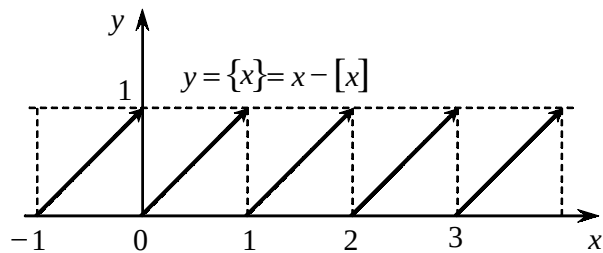


Рис. 4.12

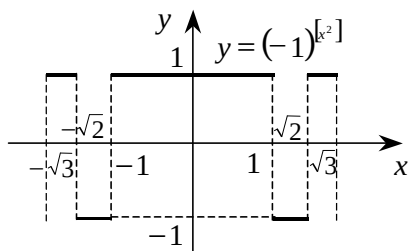


Рис. 4.13

Часто оказывается, что график функции неограниченно близко приближается к некоторой прямой l (рис. 4.14).

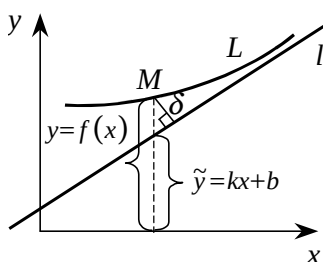


Рис. 4.14

ОПРЕДЕЛЕНИЕ. Прямая линия l называется **асимптотой** кривой L , если расстояние δ от точки M кривой L до прямой l (перпендикуляр, опущенный из произвольной точки графика на прямую) стремится к нулю.

В этом случае говорят, что кривая L асимптотически приближается к прямой l (рис. 4.14).

Различают асимптоты вертикальные, горизонтальные и наклонные.

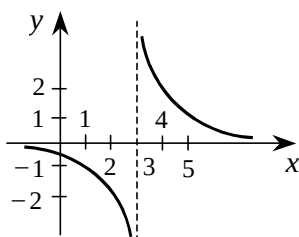


Рис. 4.15

Вертикальная асимптота. Из определения асимптоты ясно, что если график функции $y = f(x)$ имеет вертикальную асимптоту $x = a$, то при $x \rightarrow a$ хотя бы с одной из сторон (слева или справа) $y \rightarrow +\infty$ ($y \rightarrow -\infty$).

Очевидно и обратное: если при $x \rightarrow a$ хотя бы с одной из сторон $y \rightarrow +\infty$ ($y \rightarrow -\infty$), то прямая $x = a$ — вертикальная асимптота.

График $y = \frac{2}{x-3}$ (рис. 4.15) имеет вертикальную асимптоту $x = 3$,

так как при $x \rightarrow 3 - 0$ $y \rightarrow -\infty$, при $x \rightarrow 3 + 0$ $y \rightarrow +\infty$ и точка $x = 3$ является точкой разрыва II рода.

Горизонтальная асимптота. Если $\lim_{x \rightarrow \pm\infty} f(x) = b$, то прямая $y = b$ — горизонтальная асимптота для графика функции $f(x)$. Очевидно, что график функции на рис. 4.15 имеет горизонтальную асимптоту $y = 0$.

Наклонная асимптота. Пусть $f(x)$ можно представить в виде $f(x) = kx + b + \alpha(x)$, где $\alpha(x) \rightarrow 0$ при $x \rightarrow \pm\infty$. Тогда прямая $y = kx + b$ является наклонной асимптотой графика функции $f(x)$ при $x \rightarrow \pm\infty$ и

$$b = \lim_{x \rightarrow \pm\infty} [f(x) - kx], \quad k = \lim_{x \rightarrow \pm\infty} \frac{f(x)}{x}.$$

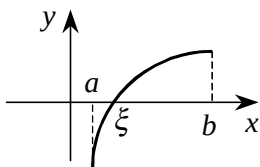


Рис. 4.16

□ **ТЕОРЕМА 4.16.** (о прохождении непрерывной функции через нуль при смене знаков). Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и пусть значения этой функции на концах отрезка $f(a)$ и $f(b)$ есть числа разных знаков. Тогда внутри отрезка $[a, b]$ найдется хотя бы одна такая точка ξ , значение функции

в которой равно нулю (рис. 4.16). (Теорема Больцано-Коши). ■

ОПРЕДЕЛЕНИЕ. Число M (или m) называется **точной верхней** (или **точной нижней**) **гранью** функции $f(x)$ на множестве $\{x\}$

$$M = \sup_{\{x\}} \{f(x)\} = \sup_{\{x\}} f(x) \quad (m = \inf_{\{x\}} \{f(x)\} = \inf_{\{x\}} f(x)),$$

если выполнены два требования:

- 1) для каждого значения x из множества $\{x\}$ справедливо неравенство $f(x) \leq M$ ($f(x) \geq m$);
- 2) для $\forall \varepsilon > 0 \exists x \in \{x\} : f(x) > M - \varepsilon$ ($f(x) < m + \varepsilon$).

□ **ТЕОРЕМА 4.17** (теорема Вейерштрасса). 1. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она ограничена на этом отрезке. 2. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она достигает на этом отрезке наименьшего значения m и наибольшего значения M . ■

10.2. Обратная функция

ОПРЕДЕЛЕНИЕ. Пусть на множестве X задана функция $f(x)$, причем Y — множество ее значений, т.е. задано множество пар чисел (x, y) ($x \in X, y \in Y$), в котором каждое число x входит лишь в одну пару, а каждое число y — по крайней мере в одну пару. Если в каждой паре этого множества поменять местами числа x и y , то получим множество пар чисел (y, x) , которое называется **обратной функцией** к функции $f(x)$.

Обозначим обратную функцию символом $x = \varphi(y)$. Обратная функция, определенная выше, вообще говоря, функцией не является. Можно сказать, что она является многозначной функцией,

поскольку теперь одному значению аргумента y может соответствовать несколько значений функции $x = \varphi(y)$. Например, для функции $y = x$ обратная функция $x = y$ однозначна, для функции $y = x^2$ обратная функция $x = \pm\sqrt{y}$ двужначна, для функции $y = \cos x$ обратная функция $x = \text{Arccos } y$ многозначна. Значит, для того чтобы обратная функция соответствовала определению функции, необходимо, чтобы исходная функция $f(x)$ была однозначной. Тогда и обратная функция будет однозначной.

□ ТЕОРЕМА 4.18. Если функция $f(x)$ непрерывна и строго монотонна на отрезке $[a, b]$, то существует однозначная обратная функция $x = \varphi(y)$, определенная на некотором промежутке, причем последняя также непрерывна и строго монотонна в том же смысле. ■

□ ТЕОРЕМА 4.19. Если функция $f(x)$ строго возрастает и непрерывна на отрезке $[a, b]$, $f(a) = A$, $f(b) = B$, то $f([a, b]) = [A, B]$ и обратная функция является однозначной строго возрастающей непрерывной на отрезке $[A, B]$ функцией. ■

□ ТЕОРЕМА 4.20. Если функция $f(x)$ строго возрастает и непрерывна на интервале (a, b) , $\lim_{x \rightarrow a} f(x) = A$, $\lim_{x \rightarrow b} f(x) = B$, то $f((a, b)) = (A, B)$ и обратная функция является однозначной строго возрастающей непрерывной на интервале (A, B) функцией, $A = \inf_{(a, b)} f(x)$,

$$B = \sup_{(a, b)} f(x). \blacksquare$$

Например, функцией, обратной к показательной функции $f(x) = e^x$, $X(f) \in \mathbb{R}$, является логарифмическая функция $\varphi(x) = \ln x$, $X(\varphi) = \{x | x > 0\}$.

§11. Комплексные числа. Функции комплексного переменного

11.1. Основные определения

ОПРЕДЕЛЕНИЕ. *Комплексным числом* z называют упорядоченную пару $z = (x, y)$ действительных чисел со следующими свойствами:

1. Два комплексных числа $z_1 = (x_1, y_1)$ и $z_2 = (x_2, y_2)$ равны тогда и только тогда, когда

$$x_1 = x_2 \text{ и } y_1 = y_2.$$

2. Сумма двух комплексных чисел $z_1 = (x_1, y_1)$ и $z_2 = (x_2, y_2)$ определяется следующим образом:

$$z_1 + z_2 = (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

3. Произведение двух комплексных чисел $z_1 = (x_1, y_1)$ и $z_2 = (x_2, y_2)$ определяется следующим образом:

$$z_1 \cdot z_2 = (x_1, y_1) \cdot (x_2, y_2) = (x_1 \cdot x_2 - y_1 \cdot y_2, x_1 \cdot y_2 + x_2 \cdot y_1). \quad (4.22)$$

4. Деление двух комплексных чисел $z_1 = (x_1, y_1)$ и $z_2 = (x_2, y_2)$ определяется как действие, обратное умножению:

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{(x_1, y_1)}{(x_2, y_2)} = \frac{(x_1 \cdot x_2 + y_1 \cdot y_2, x_2 \cdot y_1 - x_1 \cdot y_2)}{x_2^2 + y_2^2} = \\ &= \left(\frac{x_1 \cdot x_2 + y_1 \cdot y_2}{x_2^2 + y_2^2}, \frac{x_2 \cdot y_1 - x_1 \cdot y_2}{x_2^2 + y_2^2} \right) \text{ при } x_2^2 + y_2^2 \neq 0. \end{aligned} \quad (4.23)$$

Действительные числа содержатся в множестве комплексных чисел, а именно, все они являются парами вида $(x, 0)$; условимся в дальнейшем писать $(x, 0) = x$.

Пары вида $(0, y)$ называют (*чисто*) *мнимыми числами*.

ОПРЕДЕЛЕНИЕ. Пара $i = (0, 1)$ называется *мнимой единицей*:

$$i^2 = i \cdot i = -1. \quad (4.24)$$

(см. правило 3).

Запись $(0, y)$ для чисто мнимого числа эквивалентна записи iy , так как $(0, 1)(y, 0) = (0 \cdot y - 1 \cdot 0, 0 \cdot 0 + 1 \cdot y) = (0, y)$.

Каждое комплексное число $z = (x, y)$ можно записать в виде суммы действительного числа $x = (x, 0)$ и чисто мнимого числа $iy = (0, y)$:

$$z = (x, y) = (x, 0) + (0, y) = x + iy. \quad (4.25)$$

При этом $x = \operatorname{Re} z$ называется *действительной частью* комплексного числа $z = x + iy$, а $y = \operatorname{Im} z$ — *мнимой частью* $z = x + iy$.

Итак:

1°. два комплексных числа $z_1 = x_1 + iy_1$ и $z_2 = x_2 + iy_2$ равны друг другу тогда и только тогда, когда равны их действительные и мнимые части.

$$2^\circ. z_1 \pm z_2 = (x_1 + x_2) \pm i(y_1 + y_2). \quad (4.26)$$

$$3^\circ. z_1 \cdot z_2 = (x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1). \quad (4.27)$$

$$4^\circ. \frac{z_1}{z_2} = \frac{x_1 + iy_1}{x_2 + iy_2} = \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + i \frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2} \quad (4.28)$$

при $x_2^2 + y_2^2 \neq 0$.

ОПРЕДЕЛЕНИЕ. Множество всех элементов $x + iy$, в котором заданы операции (4.26)—(4.28), *называется множеством комплексных чисел* $\mathbb{C}$, а каждый его элемент — комплексным числом.

Каждому комплексному числу $x + iy$ соответствует упорядоченная пара действительных чисел (x, y) и, наоборот, паре действительных чисел (x, y) соответствует комплексное число $z = x + iy$. Эти соответствия взаимно однозначны.

Комплексное число $z = x + iy$ можно рассматривать как вектор на плоскости с координатами x и y (рис. 4.17).

Координатная плоскость, векторы $z = (x, y)$ которой интерпретируются как комплексные числа, называется *комплексной плоскостью*, ее ось X — действительной осью, Y — мнимой осью.

ОПРЕДЕЛЕНИЕ. Под комплексным числом, *сопряженным* к $z = x + iy$, понимают комплексное число $\bar{z} = x - iy$, отличающееся от z

только знаком мнимой части. При этом имеют место следующие соотношения:

$$\overline{z_1 \pm z_2} = \overline{z_1} \pm \overline{z_2}; \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}; \quad \overline{z_1 / z_2} = \overline{z_1} / \overline{z_2}; \quad \overline{z_1} \cdot z_1 = x^2 + y^2;$$

$$\overline{\overline{z}} = z; \quad \overline{\overline{z}} = z \text{ тогда и только тогда, когда } z \text{ — действительное число.}$$

Длина $|z|$ вектора $z = (x, y)$ называется модулем или абсолютной величиной комплексного числа $z = x + iy$:

$$|z| = \sqrt{x^2 + y^2}. \quad (4.29)$$

Если z — действительное число, то модуль совпадает с абсолютной величиной $|z|$. На модуль комплексных чисел могут быть перенесены различные неравенства для абсолютных величин. Так, для любых комплексных чисел $z, z_1, z_2, z_{1_k}, z_{2_k}$ верны соотношения:

$$|z| = \sqrt{z \cdot \overline{z}} = \sqrt{x^2 + y^2}$$

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2| \text{ (неравенство треугольника);}$$

$$\left| \sum_{k=1}^n z_{1_k} \cdot z_{2_k} \right|^2 \leq \sum_{k=1}^n |z_{1_k}|^2 \cdot \sum_{k=1}^n |z_{2_k}|^2 \text{ (неравенство Коши-Буняковского).}$$

11.2. Геометрическая интерпретация

Поставим в соответствие комплексному числу $z = x + iy$ на плоскости точку A , которая имеет абсциссу x и ординату y относительно прямоугольной декартовой системы координат: $A(x, y)$ (о числе z также говорят как о точке). Если $OA = (x, y)$ — радиус-вектор точки A , то он также может быть использован в качестве геометрической интерпретации числа $z = x + iy$ (рис. 4.18). Таким образом, множество точек плоскости, или множество радиус-векторов, соответствующих им, взаимно однозначно соответствует множеству комплексных чисел. Вся плоскость называется *комплексной плоскостью*. Точки оси абсцисс — *действительной оси* — соответствуют действительным, а точки оси ординат — *мнимой оси* — чисто мнимым комплексным числам.

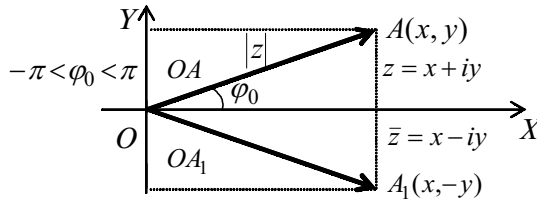


Рис. 4.18

Комплексное число $\bar{z} = x - iy$, сопряженное к $z = x + iy$, изображается точкой $A_1 = (x, -y)$ — зеркальным образом точки A относительно действительной оси.

11.3. Тригонометрическая форма комплексного числа

ОПРЕДЕЛЕНИЕ. *Аргументом* комплексного числа $z = x + iy$ ($\text{Arg } z$) называют угол φ_0 (в радианах) между радиус-вектором $OA = (x, y)$ и положительным направлением действительной оси (см. рис. 4.18), определяемый с точностью до слагаемого, кратного 2π . *Главным значением аргумента называют* $\arg z = \varphi$, где $-\pi < \varphi \leq \pi$. Тогда имеем: $\text{Arg } z = \arg z + 2\pi k, k \in \mathbf{Z}$.

Например, $\arg(1 + i) = \pi/4$.

Если обозначить длину вектора $|z| = r$, то справедливы равенства:

$$x = r \cos \varphi; y = r \sin \varphi.$$

Тогда

$$z = x + iy = r(\cos \varphi + i \sin \varphi), \quad (4.30)$$

$$r = |z| = \sqrt{(r \cos \varphi)^2 + (r \sin \varphi)^2}. \quad (4.31)$$

Тригонометрические и показательные функции связаны *формулой Эйлера*:

$$\cos \varphi + i \sin \varphi = e^{i\varphi},$$

значит,

$$z = |z| \cdot (\cos \varphi + i \sin \varphi) = r(\cos \varphi + i \sin \varphi) = re^{i\varphi}. \quad (4.32)$$

Таким образом, для каждого комплексного числа возможны три представления: алгебраическое, тригонометрическое и показательное. Если $a = |a| \cdot (\cos \varphi + i \sin \varphi)$, $b = |b| \cdot (\cos \psi + i \sin \psi)$, то

$$ab = |a| \cdot |b| \cdot [\cos(\varphi + \psi) + i\sin(\varphi + \psi)] = |a| \cdot |b| \cdot e^{i(\varphi + \psi)}, \quad (4.33)$$

$$\frac{a}{b} = \frac{|a|}{|b|} [\cos(\varphi - \psi) + i\sin(\varphi - \psi)] = \frac{|a|}{|b|} e^{i(\varphi - \psi)}. \quad (4.34)$$

11.4. Свойства показательной функции

Если z_1 и z_2 — комплексные числа, то верны следующие свойства:

1°. $e^{z_1 + z_2} = e^{z_1} e^{z_2}$.

□ В качестве примера докажем это свойство.

$$e^{z_1 + z_2} = e^{(x_1 + iy_1) + (x_2 + iy_2)} = e^{x_1} e^{x_2} [\cos(y_1 + y_2) + i\sin(y_1 + y_2)], e^{z_1} e^{z_2} = e^{(x_1 + iy_1)} \times \\ \times e^{(x_2 + iy_2)} = e^{x_1} (\cos y_1 + i\sin y_1) \cdot e^{x_2} (\cos y_2 + i\sin y_2) = e^{x_1} e^{x_2} (\cos y_1 \cos y_2 + \\ + i\cos y_1 \sin y_2 + i\sin y_1 \cos y_2 - \sin y_1 \sin y_2) = e^{x_1} e^{x_2} [\cos(y_1 + y_2) + i\sin(y_1 + y_2)].$$

Поскольку левые и правые части выражений равны, то $e^{z_1 + z_2} = e^{z_1} e^{z_2}$. ■

2°. $e^{z_1 - z_2} = \frac{e^{z_1}}{e^{z_2}}$.

3°. $(e^z)^m = e^{mz}$.

4°. $e^{z+2\pi k} = e^z, k \in \mathbf{Z}$.

Свойства 2°–4° предлагаем читателю доказать самостоятельно.

11.5. Степени и корни

1. Возведение комплексного числа $z = x + iy$ в n -ую степень (n -натуральное число) производится по формуле Муавра:

$$z^n = [|z| \cdot (\cos \varphi + i\sin \varphi)]^n = |z|^n (\cos \varphi + i\sin \varphi)^n = z^n e^{in\varphi}, n \in \mathbf{N}, \quad (4.35)$$

которую можно вывести из формулы умножения (4.33).

2. Если показатель степени — отрицательное целое число, полагаем

$z^n = 1/z^{-n}$ ($n < 0$). Тогда для z^n , как следует из формулы (4.34), может быть использована формула Муавра.

3. Пусть показатель степени имеет дробный вид, т.е. $n = \frac{1}{m}$, $m \in \mathbf{N}$, тогда возведение числа $z = x + iy$ в степень $n = \frac{1}{m}$ называют извлечением корня степени m из числа z , т.е. $z^{1/m}$.

ОПРЕДЕЛЕНИЕ. Каждое комплексное число W , такое что $W^m = z$, называют **корнем m -й степени из z** :

$$W = \sqrt[m]{z} = z^{1/m}. \quad (4.36)$$

Вычисление (4.36) также осуществляется по формуле Муавра, но теперь для $n = \frac{1}{m}$. Однако, в то время как все рассмотренные до сих пор вычислительные операции имели однозначные результаты, операция $\sqrt[m]{z}$ дает m различных результатов (корней):

$$W = \sqrt[m]{|z|} \left(\cos \frac{\varphi_0 + 2\pi k}{m} + i \sin \frac{\varphi_0 + 2\pi k}{m} \right), \quad (4.37)$$

т.е. $\sqrt[m]{z}$ дает m различных чисел: $\sqrt[m]{z} = z^{1/m} = \{W_k | k = 0, \dots, m-1\} = \{W_0, \dots, W_{m-1}\}$.

Решение

$$W_0 = \sqrt[m]{|z|} \left(\cos \frac{\varphi_0}{m} + i \sin \frac{\varphi_0}{m} \right)$$

уравнения $W^m = z$ называют главным значением корня $\sqrt[m]{z}$. Так как все W_k имеют одинаковый модуль $\sqrt[m]{|z|}$, то они в соответствии со значениями аргументов $\frac{\varphi_0 + 2\pi k}{m}$ ($k = 0, 1, \dots, m-1$) лежат в вершинах правильного многоугольника с центром в начале координат.

На рис. 4.19 изображено шесть значений $\sqrt[6]{z} = \{W_k; k = 0, 1, \dots, 5\}$.

Мы уже выяснили, что $e^{i\varphi} = \cos\varphi + i\sin\varphi$, тогда $e^{-i\varphi} = \cos\varphi - i\sin\varphi$, значит,

$$\cos\varphi = \frac{e^{i\varphi} + e^{-i\varphi}}{2}, \quad (4.38)$$

аналогично для синуса получим

$$\sin\varphi = \frac{e^{i\varphi} - e^{-i\varphi}}{2i}. \quad (4.39)$$

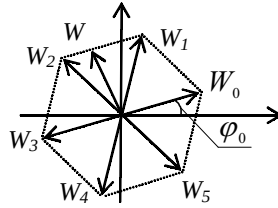


Рис. 4.19

Логарифм комплексного числа z определим как $\ln z = \ln r + i\varphi_0 + 2\pi k$, где $\ln r + i\varphi_0$ — главное значение логарифма.

Введем такие понятия, как *гиперболический синус* — $\operatorname{sh} x$ и *гиперболический косинус* — $\operatorname{ch} x$, для которых справедливы следующие соотношения:

$$\operatorname{sh} x = \frac{e^x - e^{-x}}{2}, \quad \operatorname{ch} x = \frac{e^x + e^{-x}}{2},$$

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = \frac{1}{4} \cdot (e^{2x} + 2 + e^{-2x} - e^{2x} + 2 - e^{-2x}) = 1,$$

$$2\operatorname{ch} x \cdot \operatorname{sh} x = 2 \frac{e^x + e^{-x}}{2} \cdot \frac{e^x - e^{-x}}{2} = \frac{e^{2x} - e^{-2x}}{2} = \operatorname{sh} 2x.$$

З а м е ч а н и е. Эпитет “гиперболический” объясняется тем, что уравнения $x = a \operatorname{cht}$, $y = a \operatorname{sht}$, $a > 0$, $-\infty < t < +\infty$ являются *параметрическими уравнениями* правой ветви гиперболы $x^2 - y^2 = a^2$ подобно тому, как уравнения $x = a \cos t$, $y = a \sin t$, $0 \leq t \leq 2\pi$ являются *параметрическими уравнениями* окружности $x^2 + y^2 = a^2$.

§ 12. Производная и дифференциал функции

12.1. Определение производной

Пусть функция $y = f(x)$ определена на промежутке X . Возьмем точку $x \in X$. Дадим значению x приращение $\Delta x \neq 0$ так, чтобы $x + \Delta x \in X$, тогда функция получит приращение $\Delta y = f(x + \Delta x) - f(x)$.

ОПРЕДЕЛЕНИЕ. *Производной функции* $y = f(x)$ называется предел отношения приращения функции к приращению аргумента при стремлении последнего к нулю (если этот предел существует):

$$y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \quad (4.40)$$

Производная функции имеет несколько обозначений: y' , $f'(x)$, $\frac{dy}{dx}$, $\frac{df(x)}{dx}$. Иногда в обозначении производной используется индекс, указывающий, по какой переменной взята производная, например, y'_x .

Нахождение производной функции называется *дифференцированием* этой функции.

УТВЕРЖДЕНИЕ. *Функция дифференцируема в некоторой точке в том и только том случае, когда она в этой точке имеет конечную производную.*

Если функция в точке x имеет конечную производную, то функция называется *дифференцируемой в этой точке*.

□ **ТЕОРЕМА 4.21.** *Если функция дифференцируема в некоторой точке, то она и непрерывна в этой точке.*

Д о к а з а т е л ь с т в о. Если функция дифференцируема в точке x_0 , значит, $\lim_{\Delta x \rightarrow 0} \Delta y = 0$, следовательно, она непрерывна в этой точке по определению. ■

Существуют функции непрерывные в некоторой точке, но не дифференцируемые.

Функция $y = |x|$ непрерывна в точке $x = 0$, так как $\Delta y = |\Delta x|$ и $\lim_{\Delta x \rightarrow 0} \Delta y = 0$, но $\lim_{\Delta x \rightarrow +0} \frac{\Delta y}{\Delta x} = 1$; $\lim_{\Delta x \rightarrow -0} \frac{\Delta y}{\Delta x} = -1$, значит, $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ не существует (см. рис. 4.20).

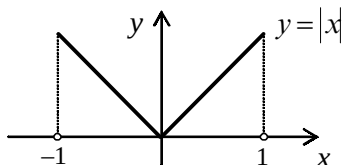


Рис. 4.20

ОПРЕДЕЛЕНИЕ. Функция, дифференцируемая во всех точках промежутка X , называется *дифференцируемой на этом промежутке*.

Геометрический смысл производной: производная $f'(x_0)$ есть *угловой коэффициент* (тангенс угла наклона) *касательной*, проведенной к кривой $y = f(x)$ в точке x_0 , т.е. $k = \operatorname{tg} \alpha = f'(x_0)$ (рис. 4.21).

Тогда уравнение касательной к кривой $y = f(x)$ в точке x_0 примет вид

$$y - f(x_0) = f'(x_0)(x - x_0). \quad (4.41)$$

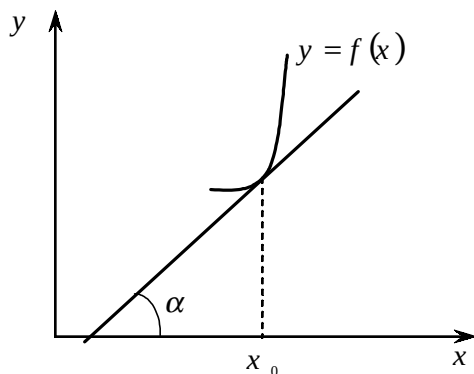


Рис. 4.21

12.2. Основные правила дифференцирования

Производная функции $y = f(x)$ может быть найдена по следующей схеме:

1. Дать аргументу x приращение $\Delta x \neq 0$ и найти наращенное значение функции $y + \Delta y = f(x + \Delta x)$.

2. Найти приращение функции $\Delta y = f(x + \Delta x) - f(x)$.

3. Составить отношение $\frac{\Delta y}{\Delta x}$.

4. Найти предел этого отношения при $\Delta x \rightarrow 0$, т.е. $y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ (если этот предел существует).

▷ ПРИМЕР 4.19. Найти производную функции $y = x^3$.

Решение: 1. Дадим аргументу x приращение $\Delta x \neq 0$ и найдем наращенное значение функции $y + \Delta y = (x + \Delta x)^3$.

2. Найдем приращение функции $\Delta y = (x + \Delta x)^3 - x^3 = x^3 + 3x^2\Delta x + 3x\Delta x^2 + \Delta x^3 - x^3 = \Delta x(3x^2 + 3x\Delta x + \Delta x^2)$.

3. Составим отношение $\frac{\Delta y}{\Delta x} = 3x^2 + 3x\Delta x + \Delta x^2$.

4. Найдем предел

$$y' = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} (3x^2 + 3x\Delta x + \Delta x^2) = 3x^2. \blacktriangleright$$

Итак, мы получили, что $(x^3)' = 3x^2$. Можно доказать, что для любого (не только натурального) n

$$(x^n)' = n x^{n-1}. \quad (4.42)$$

Полезно знать частные случаи этой формулы при $n = \frac{1}{2}$, $n = -1$:

$$(\sqrt{x})' = \frac{1}{2\sqrt{x}}, \quad (4.43)$$

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}. \quad (4.44)$$

▷ ПРИМЕР 4.20. Найти производную функции $y = x^2 \cdot \sqrt[4]{x^3}$.

Р е ш е н и е. Представим функцию в виде $y = x^2 \cdot x^{\frac{3}{4}} = x^{\frac{11}{4}}$.

Теперь по формуле (4.42) найдем $y' = \frac{11}{4} x^{\frac{7}{4}}$. ►

▷ ПРИМЕР 4.21. Составить уравнение касательной к кривой

$$y = \frac{1}{x} \text{ в точке } x = 1.$$

Р е ш е н и е. В соответствии с (4.41) уравнение касательной к кривой $y = f(x) = \frac{1}{x}$ в точке $x = 1$ имеет вид $y - f(1) = f'(1)(x - 1)$.

По формуле (4.44) найдем производную $f'(x) = -\frac{1}{x^2}$. При $x = 1$

значение $f'(1) = -1$. Значение функции в этой точке $f(1) = 1$. Уравнение касательной $y - 1 = -1(x - 1)$ или $x + y - 2 = 0$. ►

Правила дифференцирования.

1. Производная постоянной равна нулю, т.е.

$$C' = 0.$$

Правило очевидно, так как любое приращение постоянной функции $y = C$ равно нулю.

2. Производная аргумента равна 1, т.е.

$$x' = 1.$$

Правило следует из формулы (4.42) при $n = 1$.

В следующих правилах будем полагать, что $u = u(x)$, $v = v(x)$ — дифференцируемые функции.

3. Производная алгебраической суммы конечного числа дифференцируемых функций равна такой же сумме производных этих функций, т.е.

$$(u + v)' = u' + v' \quad (4.45)$$

4. Производная произведения двух дифференцируемых функций равна произведению производной первого сомножителя на второй плюс произведение первого сомножителя на производную второго, т.е.

$$(uv)' = u'v + uv'. \quad (4.46)$$

СЛЕДСТВИЕ 1. *Постоянный множитель можно выносить за знак производной:*

$$(Cu)' = Cu'. \quad (4.47)$$

СЛЕДСТВИЕ 2. *Производная произведения нескольких дифференцируемых функций равна сумме произведений производной каждого из сомножителей на все остальные, например:*

$$(uvw)' = u'vw + uv'w + uvw'. \quad (4.48)$$

5. *Производная частного двух дифференцируемых функций может быть найдена по формуле*

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} \quad (4.49)$$

(при условии, что $v \neq 0$).

□ В качестве примера докажем правило 4, т.е. формулу (4.46). Пусть $u = u(x)$ и $v = v(x)$ — дифференцируемые функции. Найдем производную функции $y = uv$.

1. Дадим аргументу x приращение $\Delta x \neq 0$. Тогда функции u и v получают наращенные значения $u + \Delta u$ и $v + \Delta v$, а функции y — значение $y + \Delta y = (u + \Delta u)(v + \Delta v)$.

2. Найдем приращение функции

$$\Delta y = (u + \Delta u)(v + \Delta v) - uv = uv + \Delta uv + u\Delta v + \Delta u\Delta v - uv = \Delta uv + u\Delta v + \Delta u\Delta v.$$

3. Составим отношение $\frac{\Delta y}{\Delta x}$, которое представим в виде

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} \cdot v + u \cdot \frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta x} \frac{\Delta v}{\Delta x} \Delta x.$$

4. Найдем предел этого отношения при $\Delta x \rightarrow 0$, используя теорему о пределах (см. подпараграф 7.3 главы 4):

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} \cdot v + u \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} + \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} \cdot \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} \cdot \lim_{\Delta x \rightarrow 0} \Delta x.$$

На основании определения производной получаем, что $y' = u'v + uv' + u'v' \cdot 0$, или $y' = u'v + uv'$. ■

▷ П Р И М Е Р 4.22. Найти производную функции $y = f(x)$ и вычислить ее значение в точке $x = 1$:

а) $y = x^3(\sqrt[4]{x} + 1)$; б) $y = 15(x^4 - 1)$; в) $y = \frac{x^3 - 1}{\sqrt{x}}$.

Р е ш е н и е. а) По формулам (4.46), (4.45) и (4.42)

$$\begin{aligned} y' &= (x^3)'(\sqrt[4]{x} + 1) + x^3(\sqrt[4]{x} + 1)' = 3x^2(\sqrt[4]{x} + 1) + x^3\left(\frac{1}{4}x^{-3/4} + 0\right) = \\ &= x^2\left(\frac{13}{4}\sqrt[4]{x} + 3\right) \end{aligned}$$

Вычислить производную можно проще, если сначала раскрыть скобки, затем применить формулы (4.45), (4.42):

$$y = x^{\frac{13}{4}} + x^3, \quad y' = \frac{13}{4}x^{\frac{9}{4}} + 3x^2 = x^2\left(\frac{13}{4}\sqrt[4]{x} + 3\right)$$

Значение производной в точке $x = 1$ есть

$$y'(1) = 1 \cdot \left(\frac{13}{4} \cdot 1 + 3\right) = 4,25.$$

б) Сначала вынесем постоянный множитель за знак производной:

$$y' = 15(x^4 - 1)' = 15 \cdot 4x^3 = 60x^3; \quad y'(1) = 60.$$

в) По формуле (4.49)

$$\begin{aligned} y' &= \frac{(x^3 - 1)' \sqrt{x} - (x^3 - 1)(\sqrt{x})'}{(\sqrt{x})^2} = \frac{3x^2 \sqrt{x} - (x^3 - 1) \cdot \frac{1}{2\sqrt{x}}}{x} = \frac{5x^3 + 1}{2x\sqrt{x}}; \\ y'(1) &= 3. \end{aligned}$$

Как и в примере 4.22, вычислим производную более простым способом, разделив предварительно каждое слагаемое числителя на знаменатель. Получим

$$y = x^{\frac{5}{2}} - x^{-\frac{1}{2}}, \quad y' = \frac{5}{2}x^{\frac{3}{2}} + \frac{1}{2}x^{-\frac{3}{2}}, \quad y'(1) = \frac{5}{2} + \frac{1}{2} = 3. \blacktriangleright$$

12.3. Производная сложной и обратной функций

Пусть переменная y есть функция от переменной u ($y = f(u)$), а переменная u , в свою очередь, есть функция от независимой переменной x , т.е. $u = \varphi(x)$, значит, задана сложная функция $y = f(\varphi(x))$.

□ ТЕОРЕМА 4.22. Если $y = f(u)$ и $\varphi(x)$ — дифференцируемые функции от своих аргументов, то производная сложной функции существует и равна производной данной функции по промежуточному аргументу u , умноженной на производную самого промежуточного аргумента по независимой переменной x , т.е.

$$y' = f'(u) \cdot u'. \blacksquare \quad (4.50)$$

▷ ПРИМЕР 4.23. Найти производные функций:

$$\text{а) } y = (\sqrt{x} + 5)^3; \text{ б) } y = \sqrt[3]{\frac{x^2 - 1}{x^2 + 1}}; \text{ в) } y = \frac{12}{x^2 + x + 1}.$$

Р е ш е н и е. а) Функцию можно представить в виде $y = u^3$, где $u = \sqrt{x} + 5$. Поэтому на основании формулы

$$(u^n)' = nu^{n-1}u'. \quad (4.51)$$

$$y' = 3u^2 \cdot u' = 3(\sqrt{x} + 5)^2 (\sqrt{x} + 5)' = \frac{3(\sqrt{x} + 5)^2}{2\sqrt{x}}.$$

б) Имеем $y = \sqrt[3]{u}$, где $u = \frac{x^2 - 1}{x^2 + 1}$, поэтому по формулам (4.50)

$$\begin{aligned} y' &= \frac{1}{3}u^{-\frac{2}{3}} \cdot u' = \frac{1}{3} \cdot \left(\frac{x^2 - 1}{x^2 + 1} \right)^{-\frac{2}{3}} \left(\frac{x^2 - 1}{x^2 + 1} \right)' = \\ &= \frac{1}{3} \sqrt[3]{\left(\frac{x^2 + 1}{x^2 - 1} \right)^2} \frac{2x(x^2 + 1) - (x^2 - 1) \cdot 2x}{(x^2 + 1)^2} = \\ &= \frac{4x}{3(x^2 + 1)\sqrt[3]{(x^2 + 1)(x^2 - 1)^2}}. \end{aligned}$$

в) Вынося постоянный множитель 12 за знак производной и используя формулу $\left(\frac{1}{u}\right)' = -\frac{1}{u^2} \cdot u'$, где $u = x^2 + x + 1$, получим

$$y' = 12 \left(\frac{1}{x^2 + x + 1} \right)' = 12 \left(-\frac{1}{(x^2 + x + 1)^2} \right) (x^2 + x + 1)' = \frac{-12(2x + 1)}{(x^2 + x + 1)^2}. \blacktriangleright$$

□ ТЕОРЕМА 4.23. Для дифференцируемой функции с производной, не равной нулю, производная обратной функции равна обратной величине производной данной функции, т.е.

$$x'_y = \frac{1}{y'_x}. \blacksquare \quad (4.52)$$

12.4. Дифференциал функции

ОПРЕДЕЛЕНИЕ. Функция $y = f(x)$ называется дифференцируемой в точке x_0 , если ее приращение $\Delta y = (x_0 + \Delta x) - f(x_0)$ в этой точке представимо в виде:

$$\Delta y = A \cdot \Delta x + \varepsilon(\Delta x) \cdot \Delta x, \quad (4.53)$$

где A — некоторое число, не зависящее от Δx , а $\varepsilon(\Delta x)$ — бесконечно малая функция при $\Delta x \rightarrow 0$ и $\lim_{\Delta x \rightarrow 0} \varepsilon(\Delta x) = 0$ (см. §9 главы 4).

ОПРЕДЕЛЕНИЕ. Линейная функция $A \cdot \Delta x$ относительно приращения Δx называется **дифференциалом функции** $f(x)$ в точке x_0 и обозначается $df(x_0)$ или $dy = A \cdot \Delta x$, где $A = f'(x_0)$; обозначив Δx как dx , получим

$$dy = A \cdot dx. \quad (4.54)$$

Дифференциал dy функции $f(x)$ записывается в виде:

$$dy = f'(x_0) dx, \quad (4.55)$$

Из равенства (4.55) производную $f'(x_0)$ в любой точке x можно вычислить как отношение дифференциала функции dy к дифференциалу независимой переменной dx :

$$f'(x) = \frac{dy}{dx}. \quad (4.56)$$

Таблица дифференциалов

$y = c; \quad dy = 0.$	$y = \sin x; \quad dy = \cos x \, dx$
$y = x^\mu; \quad dy = \mu \cdot x^{\mu-1} dx$	$y = \cos x; \quad dy = -\sin x \, dx$
$y = \frac{1}{x}; \quad dy = -\frac{dx}{x^2}$	$y = \operatorname{tg} x; \quad dy = \sec^2 x \, dx = \frac{dx}{\cos^2 x}$
$y = \sqrt{x}; \quad dy = \frac{dx}{2\sqrt{x}}$	$y = \operatorname{ctg} x; \quad dy = -\operatorname{cosec}^2 x \, dx = -\frac{dx}{\sin^2 x}$
$y = a^x; \quad dy = a^x \ln a \, dx$	$y = \arcsin x; \quad dy = \frac{dx}{\sqrt{1-x^2}}$
$y = e^x; \quad dy = e^x \, dx$	$y = \arccos x; \quad dy = -\frac{dx}{\sqrt{1-x^2}}$
$y = \log_a x; \quad dy = \frac{\log_a e}{x} dx = \frac{1}{x \ln a} dx$	$y = \operatorname{arctg} x; \quad dy = \frac{dx}{1+x^2}$
$y = \ln x; \quad dy = \frac{dx}{x}$	$y = \operatorname{arccctg} x; \quad dy = -\frac{dx}{1+x^2}$
$d(cu) = c \cdot du, \quad d(u \pm v) = du \pm dv$	$d(uv) = v \cdot du + u \cdot dv, \quad d\left(\frac{u}{v}\right) = \frac{v \cdot du - u \cdot dv}{v^2}$

12.5. Геометрический смысл дифференциала функции

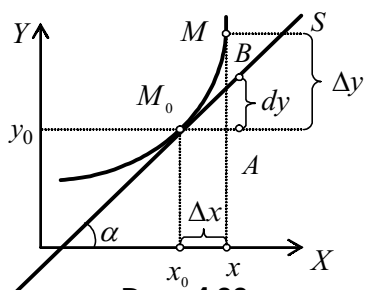


Рис. 4.22

Пусть $f(x)$ (рис. 4.22) определена в некоторой окрестности точки x_0 , непрерывна в точке x_0 , $y_0 = f(x_0)$, $M_0 = (x_0, y_0)$. Зафиксируем приращение аргумента Δx . Пусть приращение функции $\Delta y = f(x_0 + \Delta x) - f(x_0)$, тогда точка $M = (x_0 + \Delta x, y_0 + \Delta y)$.

Здесь M_0S — касательная к кривой $f(x)$ в точке $M_0 = (x_0, y_0)$, α — угол

между касательной и осью Ox . Тогда M_0A — приращение аргумента, AM — соответствующее приращение функции. Рассматривая треугольник ABM_0 , получаем $AB = \Delta x \operatorname{tg} \alpha = f'(x_0) \Delta x = dy$, так как $f'(x_0) = \operatorname{tg} \alpha$, т.е. это главная по порядку величины Δx и линейная относительно нее часть приращения функции Δy .

12.6. Производные комплекснозначных функций действительного аргумента

В §11 главы 4 мы ввели понятия гиперболического синуса $\operatorname{sh} x$ и гиперболического косинуса $\operatorname{ch} x$:

$$\operatorname{sh} x = \frac{e^x - e^{-x}}{2}, \quad \operatorname{ch} x = \frac{e^x + e^{-x}}{2}.$$

Вычислим их производные:

$$(\operatorname{sh} x)' = \left(\frac{e^x - e^{-x}}{2} \right)' = \frac{e^x + e^{-x}}{2} = \operatorname{ch} x, \quad (4.57)$$

$$(\operatorname{ch} x)' = \left(\frac{e^x + e^{-x}}{2} \right)' = \frac{e^x - e^{-x}}{2} = \operatorname{sh} x. \quad (4.58)$$

Иными словами, производная от гиперболического синуса — гиперболический косинус и — наоборот.

Если функция $f(x)$ задана в некоторой окрестности точки x_0 числовой оси и принимает комплексные значения, т.е. имеет вид:

$$f(x) = u(x) + iv(x), \quad u(x) \in \mathbf{R}, \quad v(x) \in \mathbf{R}, \quad x \in \mathbf{C},$$

то ее производная в точке x_0 определяется равенством:

$$f'(x_0) = u'(x_0) + iv'(x_0),$$

если $\exists u'(x), v'(x)$ в точке x_0 .

▷ ПРИМЕР 4.24. Найти производную функции $f(x) = \cos \alpha x + i \sin \alpha x$.

Решение. Воспользуемся основными правилами дифференцирования (подпараграф 12.2). Тогда

$$f'(x_0) = -\alpha \sin \alpha x + i \alpha \cos \alpha x = i \alpha (\cos \alpha x + i \sin \alpha x) = i \alpha \cdot f(x). \blacktriangleright$$

З а м е ч а н и е. $(\lambda_1 f_1 + \lambda_2 f_2)' = \lambda_1 f_1' + \lambda_2 f_2', \lambda_1 \in \mathbf{C}, \lambda_2 \in \mathbf{C}$.

§ 13. Некоторые приложения производной

13.1. Производные высших порядков

Пусть функция $y = f(x)$ имеет производную $y' = f'(x)$ во всех точках некоторой окрестности точки x . Если функция $f'(x)$ в свою очередь имеет в точке x производную $[f'(x)]'$, то она называется *второй производной* (или *производной второго порядка*) функции $f(x)$ в точке x и обозначается y'' , или $f''(x)$, или $f^{(2)}(x)$, или $\frac{d^2 y}{dx^2}$, или $D^2 y$.

Аналогично, если функция $y = f(x)$ имеет конечную вторую производную $f''(x)$ во всех точках некоторой окрестности точки x и функция $f''(x)$ дифференцируема, то ее производная называется *производной третьего порядка* от функции $y = f(x)$ и обозначается y''' , или $f'''(x)$, или $y^{(3)}f(x)$, или $\frac{d^3 y}{dx^3}$, или $D^3 y$.

Подобным образом от третьей производной переходим к четвертой и так далее. Если предположить, что понятие $(n-1)$ -й производной уже определено и что $(n-1)$ -я производная существует и конечна в некоторой окрестности точки x , а также является дифференцируемой функцией, то ее производная в некоторой точке x этого промежутка называется *производной n -го порядка*: $y^{(n)}$, $f^{(n)}(x)$, $\frac{d^n y}{dx^n}$, $D^n y$.

Итак, для того чтобы вычислить n -ю производную от какой-либо функции, нужно вычислить производные всех предшествующих порядков.

Будем считать, что $y^{(0)}(x) = y(x)$.

Правила дифференцирования:

$$(Cu)^{(n)} = Cu^{(n)}, \quad (4.59)$$

$$(u \pm v)^{(n)} = u^{(n)} \pm v^{(n)}. \quad (4.60)$$

Приведем примеры:

1. Если $y = a^x$, $a > 0$, то $y' = a^x \ln a$, $y'' = a^x \ln^2 a, \dots$. Общий закон: $y^{(n)} = a^x \ln^n a$, $n = 0, 1, 2, \dots$. Если $y = e^x$, то $(e^x)^{(n)} = e^x$. (4.61)

2. $y = x^\mu$, μ — любая степенная функция, то $y' = \mu \cdot x^{\mu-1}$, $y'' = \mu \cdot (\mu-1) \cdot x^{\mu-2}$, ... Общий закон: $y^{(n)} = \mu \cdot (\mu-1) \cdot (\mu-n+1) \cdot x^{\mu-n}$. Если $\mu = -1$, то

$$\left(\frac{1}{x}\right)^{(n)} = (-1) \cdot (-2) \cdots (-n) \cdot x^{-1-n} = \frac{(-1)^n \cdot n!}{x^{n+1}}. \quad (4.62)$$

При расчете производных высших порядков применяется формула **бинома Ньютона**:

$$(x+a)^n = x^n + C_n^1 a x^{n-1} + C_n^2 a^2 x^{n-2} + \dots + C_n^{n-1} a^{n-1} x + a^n, \quad (4.63)$$

$$(x+a)^n = \sum_{k=0}^n C_n^k a^k x^{n-k}, \quad (4.64)$$

$$C_n^k + C_n^{k-1} = C_{n+1}^k. \quad (4.65)$$

Биномиальные коэффициенты

$$C_n^k = \frac{n!}{k! \cdot (n-k)!}. \quad (4.66)$$

Например:

$$C_4^0 = \frac{4!}{0! \cdot (4-0)!} = 1, \quad C_4^1 = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot (1 \cdot 2 \cdot 3)} = 4, \quad C_4^2 = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot (1 \cdot 2)} = 6,$$

$$C_4^3 = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 3 \cdot (1)} = 4, \quad C_4^4 = \frac{1 \cdot 2 \cdot 3 \cdot 4}{1 \cdot 2 \cdot 3 \cdot 4 \cdot (0!)} = 1.$$

(Заметим, что $0! = 1$.)

Биномиальные коэффициенты можно найти с помощью следующей треугольной таблицы:

$$\begin{array}{ccccccc} & & & & 1 & & & \\ & & & & 1 & 2 & 1 & \\ & & & 1 & 3 & 3 & 1 & \\ & & 1 & 4 & 6 & 4 & 1 & \\ & 1 & 5 & 10 & 10 & 5 & 1 & \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

называемой *треугольником Паскаля*, в которой первые и последние числа во всех строчках равны единице и начиная с третьей каждое число в строке, отличное от первого и последнего, получается сложением двух ближайших к нему чисел предшествующей строки. В n -й строке будут стоять числа $C_n^0, C_n^1, C_n^2, \dots, C_n^n$; при этом $C_n^0 = C_n^n = 1, C_n^k = C_n^{n-k}$.

□ ТЕОРЕМА 4.24. Если функции $y_1 = f(x_1)$ и $y_2 = f(x_2)$ имеют в точке x_0 производные порядка $n \in \mathbf{N}$, то любая их линейная комбинация $\lambda_1 y_1 + \lambda_2 y_2, \lambda_1 \in \mathbf{R}, \lambda_2 \in \mathbf{R}$ и их произведение $y_1 \cdot y_2$ имеют в точке x_0 производные порядка $n \in \mathbf{N}$, причем:

$$(\lambda_1 y_1 + \lambda_2 y_2)^{(n)} = \lambda_1 y_1^{(n)} + \lambda_2 y_2^{(n)}, \quad (4.67)$$

$$(y_1 \cdot y_2)^{(n)} = \sum_{k=0}^n C_n^k y_1^{(n-k)} y_2^{(k)}, \quad n \in \mathbf{N}. \quad \blacksquare \quad (4.68)$$

Иными словами, справедлива формула:

$$\begin{aligned} (y_1 \cdot y_2)^{(n)} &= y_1^{(n)} \cdot y_2 + C_n^1 y_1^{(n-1)} \cdot y_2' + C_n^2 y_1^{(n-2)} \cdot y_2^{(2)} + \\ &+ C_n^3 y_1^{(n-3)} \cdot y_2^{(3)} + \dots + y_1 \cdot y_2^{(n)}. \end{aligned} \quad (4.69)$$

Например, вычислим “ n ”-ю производную функции $y = x^2 e^x$, полагая $y_1 = e^x, y_2 = x^2, y_1^{(k)} = e^x; y_2' = 2x, y_2^{(2)} = 2, y_2^{(3)} = y_2^{(4)} = \dots = 0$.

$$(x^2 e^x)^{(n)} = e^x x^2 + C_n^1 e^x 2x + C_n^2 e^x 2 = e^x [x^2 + 2nx + n(n-1)]$$

Формула (4.69) называется формулой Лейбница. Она напоминает формулу бинома Ньютона (4.63), только вместо степени берется производная соответствующего порядка.

13.2. Дифференциальные теоремы о среднем

□ ТЕОРЕМА ФЕРМА. Пусть функция $f(x)$ определена в некотором промежутке X и во внутренней точке c этого промежутка принимает наибольшее или наименьшее значение. Если в этой точке существует конечная производная $f'(c)$, то производная функции в этой точке равна нулю, т.е. $f'(c) = 0$.

Д о к а з а т е л ь с т в о. Пусть для определенности $f(x)$ принимает в точке c наибольшее значение: для всех $x \in X$ выполняется условие $f(x) \leq f(c)$. По определению производной:

$$f'(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}, \text{ причем предел этот не зависит от того,}$$

будет ли x приближаться к c справа или слева. Но при $x > c$

$$\frac{f(x) - f(c)}{x - c} \leq 0, \text{ так что в пределе при } x \rightarrow c + 0 \text{ получим}$$

$$f'(c) \leq 0. \quad (4.70)$$

Если же $x \rightarrow c$, то $\frac{f(x) - f(c)}{x - c} \geq 0$, и, переходя к пределу при $x \rightarrow c - 0$, получим

$$f'(c) \geq 0. \quad (4.71)$$

Из (4.70) и (4.71) следует, что $f'(c) = 0$. ■

Геометрический смысл теоремы Ферма очевиден: *в точке наибольшего или наименьшего значения, достигаемого внутри промежутка X , касательная к графику функции параллельна оси абсцисс.*

□ **ТЕОРЕМА РОЛЛЯ.** Пусть выполняются следующие условия:

- 1) функция $f(x)$ определена и непрерывна на отрезке $[a, b]$;
- 2) функция дифференцируема, по крайней мере, в открытом промежутке (a, b) ;
- 3) на концах отрезка функция принимает равные значения: $f(a) = f(b)$.

Тогда между a и b найдется такая точка c : $a < c < b$: $f'(c) = 0$.

Д о к а з а т е л ь с т в о. Функция $f(x)$ непрерывна на отрезке $[a, b]$, и по теореме Вейерштрасса (теорема 4.17) принимает в этом промежутке как свое наибольшее значение M , так и свое наименьшее значение m .

Рассмотрим два случая.

1. $M = m$. Тогда $f(x)$ на отрезке $[a, b]$ сохраняет постоянное значение. В самом деле неравенство $m \leq f(x) \leq M$ в этом случае дает равенство $f(x) = M = m$ при всех x , поэтому $f'(x) = 0$ во всем промежутке, значит, в качестве c можно взять любую точку из (a, b) .

2. $M > m$. Мы знаем, что оба эти значения функцией достигаются, но, так как $f(a) = f(b)$, то оба они не могут достигаться на концах промежутка, и хоть одно из них достигается в некоторой точке c : $a < c < b$, то по теореме Ферма $f'(c) = 0$. ■

Геометрический смысл теоремы Ролля

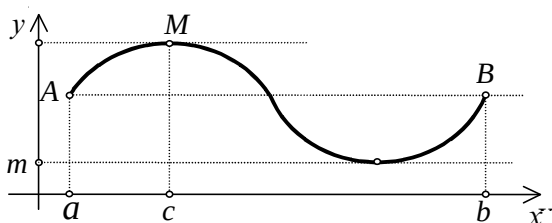


Рис. 4.23

Если крайние ординаты кривой $y = f(x)$ (см. рис. 4.23) равны, то на кривой найдется хотя бы одна точка, в которой касательная к графику функции будет параллельна оси X .

В этой точке производная и будет равна нулю (заметим, что на рис. 4.23 таких точек две).

Все условия теоремы Ролля очень важны, и невыполнение хотя бы одного из них может привести к неверному заключению теоремы. Так, для функций, приведенных на рис. 4.24, нарушено только одно условие — дифференцируемость на интервале (a, b) . В результате не существует такой точки c : $f'(c) = 0$.

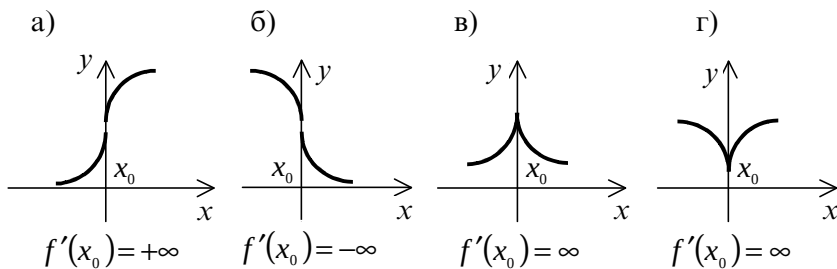


Рис. 4.24

□ ТЕОРЕМА ЛАГРАНЖА. Пусть выполняются следующие условия:

- 1) функция $f(x)$ определена и непрерывна на отрезке $[a, b]$;
- 2) функция дифференцируема, по крайней мере, в открытом промежутке (a, b) ;

Тогда между a и b найдется такая точка c ($a < c < b$), что для нее выполняется равенство (рис. 4.25):

$$\frac{f(b) - f(a)}{b - a} = f'(c). \quad (4.72)$$

До к а з а т е л ь с т в о. Введем вспомогательную функцию, определив ее на отрезке $[a, b]$ равенством: $F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a)$. Эта функция удовлетворяет всем условиям теоремы Ролля. В самом деле, она непрерывна в $[a, b]$, так как представляет собой разность между непрерывной функцией $f(x)$ и линейной функцией. В промежутке (a, b) она имеет определенную конечную производную, равную $F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$.

Наконец непосредственной подстановкой убеждаемся в том, что $F(a) = F(b) = 0$, т.е. $F(x)$ принимает равные значения на концах промежутка.

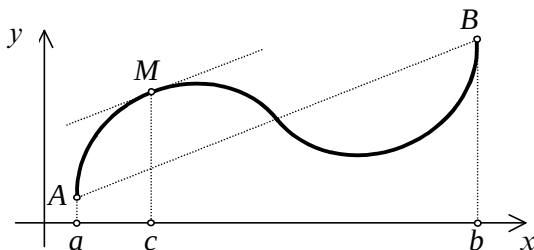


Рис. 4.25

Следовательно, к функции $F(x)$ можно применить теорему Ролля и утверждать существование в (a, b) такой точки c , что $F'(c) = 0$.

Таким образом, $f'(c) - \frac{f(b) - f(a)}{b - a} = 0$, или

$$f(b) - f(a) = f'(c)(b - a). \blacksquare \quad (4.73)$$

СЛЕДСТВИЕ. Если функция $f(x)$ непрерывна на промежутке X и во всех внутренних точках этого промежутка имеет производную, равную нулю, то функция тождественно постоянна на X .

Доказательство. Возьмем на рассматриваемом промежутке X отрезок $[a, x]$. Согласно теореме Лагранжа $f(x) - f(a) = f'(c)(x - a)$, где $a < c < b$. По условию $f'(c) = 0$, следовательно, $f(x) - f(a) = 0$, т.е. $f(x) = f(a) = \text{const.}$ $\blacksquare$

13.3. Формула Тейлора для многочлена

Если $p(x)$ есть целый многочлен степени n :

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_nx^n, \quad (4.74)$$

то, последовательно дифференцируя его n раз, получим:

$$p'(x) = a_1 + 2a_2x + 3a_3x^2 + \dots + na_nx^{n-1};$$

$$p''(x) = 1 \cdot 2a_2 + 2 \cdot 3a_3x + \dots + (n-1) \cdot na_nx^{n-2};$$

$$p'''(x) = 1 \cdot 2 \cdot 3a_3 + \dots + (n-2) \cdot (n-1) \cdot na_nx^{n-3};$$

.....

$$p^{(n)}(x) = 1 \cdot 2 \cdot 3 \cdots n \cdot a_n.$$

Полагая во всех этих формулах $x = 0$, найдем выражения коэффициентов многочлена через значения самого многочлена и его производных при $x = 0$:

$$a_0 = p(0); \quad a_1 = \frac{p'(0)}{1!}; \quad a_2 = \frac{p''(0)}{2!}; \quad a_3 = \frac{p'''(0)}{3!}; \quad \dots;$$

$$a_n = \frac{p^{(n)}(0)}{n!}.$$

Подставим эти значения коэффициентов в (4.74):

$$p(x) = p(0) + \frac{p'(0)}{1!}x + \frac{p''(0)}{2!}x^2 + \frac{p'''(0)}{3!}x^3 + \dots + \frac{p^{(n)}(0)}{n!}x^n. \quad (4.75)$$

Вместо того чтобы разлагать многочлен по степеням x , можно было бы взять его разложение по степеням $x - x_0$, где x_0 есть некоторое постоянное частное значение x :

$$p(x) = A_0 + A_1(x - x_0) + A_2(x - x_0)^2 + \dots + A_n(x - x_0)^n.$$

Тогда:

$$\begin{aligned} p(x) = & p(x_0) + \frac{p'(x_0)}{1!}(x - x_0) + \frac{p''(x_0)}{2!}(x - x_0)^2 + \dots + \\ & + \frac{p^{(n)}(x_0)}{n!}(x - x_0)^n. \end{aligned} \quad (4.76)$$

Формула (4.76) так же, как и ее частный случай (4.75), называется формулой Тейлора.

З а м е ч а н и е. Если многочлен $p(x)$ представлен в виде:

$$p(x) = c_0 + \frac{c_1}{1!}(x - x_0) + \frac{c_2}{2!}(x - x_0)^2 + \dots + \frac{c_n}{n!}(x - x_0)^n,$$

то $p(0) = c_0$, $p'(0) = c_1$, $p''(0) = c_2$, ..., $p^{(n)}(0) = c_n$.

13.4. Разложение произвольной функции

Рассмотрим произвольную функцию $f(x)$, не являющуюся целым многочленом, но определенную в некотором промежутке X . Предположим, что для нее в точке x_0 существуют производные всех порядков до n -го включительно (точнее, до $n-1$ -го: $f'(x)$, $f''(x)$, ..., $f^{(n-1)}(x)$) в некоторой окрестности точки x_0 , и есть производная n -го порядка $f^{(n)}(x_0)$ в самой точке x_0 . Тогда по формуле (4.76) для функции $f(x)$ может быть составлен многочлен:

$$\begin{aligned} p_n(x) = & f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \\ & + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \end{aligned} \quad (4.77)$$

Этот многочлен и его производные в точке x_0 имеют те же значения, что и функция $f(x)$.

Но если сама функция $f(x)$ не есть целый многочлен n -й степени, уже нельзя утверждать равенства $f(x) = p_n(x)$. Многочлен $p_n(x)$ дает лишь некоторое приближение к функции $f(x)$, с помощью которого она и может быть вычислена с некоторой степенью точности.

Если к многочлену (4.77) добавить очередной $(n + 1)$ -й член формулы Тейлора, но в точке, средней между x_0 и x — точке c , тогда выражение

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \\ + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1} -$$

— **формула Тейлора** с дополнительным членом в форме Лагранжа. Если перенести в ней $f(x_0)$ налево и положить $x - x_0 = \Delta x$, то она переписывается так:

$$\Delta f(x_0) = f(x_0 + \Delta x) - f(x_0) = \frac{f'(x_0)}{1!}\Delta x + \frac{f''(x_0)}{2!}\Delta x^2 + \dots + \\ + \frac{f^{(n)}(x_0)}{n!}\Delta x^n + \frac{f^{(n+1)}(c)}{(n+1)!}\Delta x^{n+1}. \quad (4.78)$$

Формула (4.78) является прямым обобщением формулы конечных приращений: $\Delta f(x_0) = f(x_0 + \Delta x) - f(x_0) = f'(c)\Delta x$, которая отвечает $n = 0$.

Проще всего выглядит формула Тейлора, если $x_0 = 0$ (**формула Тейлора-Маклорена**):

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + r_n(x). \quad (4.79)$$

Например, разложим функцию $f(x) = e^x$ по формуле Тейлора-Маклорена. Получим:

$$f(x) = e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + r_n(x).$$

Позже, когда мы будем говорить о рядах (§ 23 главы 4), мы рассмотрим этот пример (пример 4.61) более подробно.

§ 14. Построение графиков функций

14.1. Правило Лопиталя

□ ТЕОРЕМА 4.25. Предел отношения двух бесконечно малых или бесконечно больших функций равен пределу отношения их производных (конечному или бесконечному), если последний существует в указанном смысле. ■

Итак, если имеется неопределенность вида $\frac{0}{0}$ или $\frac{\infty}{\infty}$, то

$$\lim_{\substack{x \rightarrow 0 \\ (x \rightarrow \infty)}} \frac{f(x)}{g(x)} = \lim_{\substack{x \rightarrow x_0 \\ (x \rightarrow \infty)}} \frac{f'(x)}{g'(x)}. \quad (4.80)$$

▷ ПРИМЕР 4.25. Найти:

$$\text{а) } \lim_{x \rightarrow +\infty} \frac{x}{e^x}; \text{ б) } \lim_{x \rightarrow +\infty} \frac{x^k}{a^x} \text{ при } a > 1; \text{ в) } \lim_{x \rightarrow +\infty} \frac{\log_a x}{x^k} \text{ при } a > 1.$$

Р е ш е н и е. а) Имеем неопределенность вида $\frac{\infty}{\infty}$. Применяя правило Лопиталя, получим:

$$\lim_{x \rightarrow +\infty} \frac{x}{e^x} = \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow +\infty} \frac{x'}{(e^x)'} = \lim_{x \rightarrow +\infty} \frac{1}{e^x} = 0.$$

б) Имеем неопределенность вида $\frac{\infty}{\infty}$. Применим правило Лопиталя $[k] + 1$ раз, где $[k]$ — целая часть числа k :

$$\begin{aligned} \lim_{x \rightarrow +\infty} \frac{x^k}{a^x} &= \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow +\infty} \frac{kx^{k-1}}{a^x \ln a} = \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow +\infty} \frac{k(k-1)x^{k-2}}{a^x \ln^2 a} = \\ &= \left[\frac{\infty}{\infty} \right] = \dots = \lim_{x \rightarrow +\infty} \frac{k(k-1) \dots (k-[k])x^{k-[k]-1}}{a^x \ln^{[k]+1} a}. \end{aligned}$$

При каждом применении правила Лопиталья степень числителя будет уменьшаться на единицу и через $[k] + 1$ раз станет отрицательной, т.е. числитель обратится в бесконечно малую величину, если k — не целое число (если k — целое, то в постоянную величину). Знаменатель же будет оставаться бесконечно большой

величиной. Таким образом, $\lim_{x \rightarrow +\infty} \frac{x^k}{a^k} = 0$.

$$\begin{aligned} \text{в) } \lim_{x \rightarrow +\infty} \frac{\log_a x}{x^k} &= \left[\frac{\infty}{\infty} \right] = \lim_{x \rightarrow +\infty} \frac{(\log_a x)'}{(x^k)'} = \lim_{x \rightarrow +\infty} \frac{1}{kx^{k-1}} = \\ &= \frac{1}{k \ln a} \lim_{x \rightarrow +\infty} \frac{1}{x^k} = 0. \blacktriangleright \end{aligned}$$

Правило Лопиталья позволяет сравнивать бесконечно большие величины: степенная функция x^n — бесконечно большая более высокого порядка, чем логарифмическая $\log_a x$, а показательная a^x , $a > 1$ — бесконечно большая более высокого порядка, чем степенная x^n ; это означает, что $\lim_{x \rightarrow +\infty} \frac{x^n}{\log_a x} = \infty$, $\lim_{x \rightarrow +\infty} \frac{a^x}{x^n} = \infty$, где $a > 1$.

14.2. Возрастание и убывание функций

В подпараграфе 6.2 мы уже дали определения возрастания и убывания, монотонности функций.

Геометрическая интерпретация условия монотонности функции приведена на рис. 4.26.

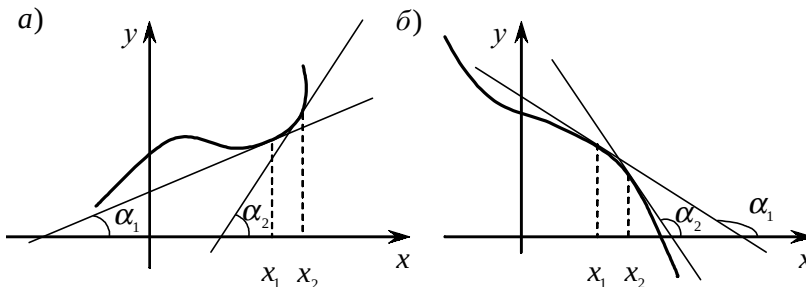


Рис. 4.26

Если касательные к кривой в некотором промежутке направлены под острыми углами к оси абсцисс (см. рис. 4.26, а), то функция возрастает, если под тупыми (см. рис. 4.26, б), то убывает.

Если функция имеет горизонтальную касательную в точке, то производная в этой точке равна нулю (подпараграф 13.2).

□ ТЕОРЕМА 4.26. (достаточное условие возрастания (убывания) функции). Если производная дифференцируемой функции положительна (отрицательна) внутри некоторого промежутка X , то функция возрастает (убывает) на этом промежутке. ■

Заметим, что необходимое условие монотонности более слабое.

□ ТЕОРЕМА 4.27. (необходимое условие возрастания (убывания) функции). Дифференцируемая на некотором промежутке X функция $f(x)$ возрастает (убывает) на этом промежутке, если производная этой функции $f'(x)$ положительна (отрицательна) на этом промежутке за исключением конечного числа точек, в которых производная равна нулю. ■

▷ ПРИМЕР 4.26. Найти интервалы монотонности функции:
а) $y = x^2 - 4x + 3$; б) $y = x^3$.

Решение. а) Имеем $y' = 2x - 4$. Очевидно, $y' > 0$ при $x > 2$ и $y' < 0$ при $x < 2$, т.е. функция убывает на интервале $(-\infty; 2)$ и возрастает на интервале $(2; +\infty)$, при этом $x = 2$ — абсцисса вершины параболы.

б) Найдем производную $y' = 3x^2$. Очевидно, что $y' > 0$ при $x \neq 0$. При $x = 0$ производная обращается в нуль. Функция же монотонно возрастает на всей числовой оси. ►

14.3. Экстремум функции

ОПРЕДЕЛЕНИЕ. Точка x_0 называется **точкой максимума** функции $f(x)$, если в некоторой окрестности этой точки при $x \neq x_0$ выполняется неравенство $f(x) < f(x_0)$.

ОПРЕДЕЛЕНИЕ. Точка x_1 называется **точкой минимума** функции $f(x)$, если в некоторой окрестности этой точки при $x \neq x_1$ выполняется неравенство $f(x) > f(x_1)$.

Значения функции в точках x_0 и x_1 называются соответственно **максимумом** и **минимумом функции**. Понятия максимума и минимума функции объединяются общим понятием **экстремума функции**.

□ **ТЕОРЕМА 4.28.** (первое достаточное условие экстремума). Если при переходе через точку x_0 производная дифференцируемой функции $y = f(x)$ меняет свой знак с плюса на минус, то точка x_0 есть точка максимума функции $y = f(x)$, а если с минуса на плюс, — то точка минимума. ■

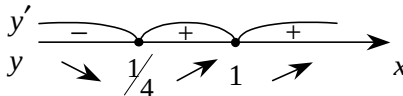
ОПРЕДЕЛЕНИЕ. Если функция $f(x)$ определена в некоторой окрестности точки x_0 , и в этой точке производная либо существует и равна нулю, либо не существует, то точка x_0 называется **критической точкой** этой функции.

▷ **ПРИМЕР 4.27.** Исследовать на экстремум функцию $y = x(x - 1)^3$.

Решение.

1. Производная функции $y' = (x - 1)^3 + 3x(x - 1)^2 = (x - 1)^2(4x - 1)$.

2. Приравнявая производную к нулю, находим **критические точки** функции $x_1 = 1/4$; $x_2 = 1$. (Точек, в которых производная не существует, у данной функции нет: $f'(x)$ определена на всей числовой оси).



3. Нанесем критические точки на числовую прямую (рис. 4.27).

Рис. 4.27

Для определения знака производной слева и справа от критической точки $x_1 = 1/4$ выберем, например, значения $x = 0$ и $x = 1/2$ и найдем $f'(0) = -1 < 0$ и $f'(1/2) = 1/4 > 0$; следовательно, $f'(x) < 0$ при всех $x < 1/4$ и $f'(x) > 0$ на интервале $(1/4, 1)$.

Аналогично устанавливаем, что $f'(x) > 0$ и на интервале $(1, +\infty)$.

Согласно достаточному условию $x_1 = 1/4$ — точка минимума данной функции. В точке $x_2 = 1$ экстремума нет.

4. Находим $f_{\min}\left(\frac{1}{4}\right) = \frac{1}{4}\left(\frac{1}{4} - 1\right)^3 = -\frac{27}{256}$. ►

14.4. Выпуклость функции. Точки перегиба

Пусть функция $f(x)$ задана на интервале (a, b) и $a < x_1 < x_2 < b$. Проведем прямую через точки $A = (x_1, f(x_1))$ и $B = (x_2, f(x_2))$, лежащие на графике функции $y = f(x)$. Уравнение этой прямой можно записать в виде:

$$y = \frac{f(x_2) \cdot (x - x_1) + f(x_1) \cdot (x_2 - x)}{x_2 - x_1}.$$

Обозначим $y = l(x)$.

ОПРЕДЕЛЕНИЕ. Функция $f(x)$ называется **выпуклой вверх** (рис. 4.28) на интервале (a, b) , если, каковы бы ни были точки x_1 и x_2 : $a < x_1 < x_2 < b$, для любого $x \in (x_1, x_2)$ выполняется неравенство: $l(x) \leq f(x)$. Если для всех точек $x \in (x_1, x_2)$ выполняется противоположное неравенство $l(x) \geq f(x)$, то функция $f(x)$ называется **выпуклой вниз** на интервале (a, b) (рис. 4.29).

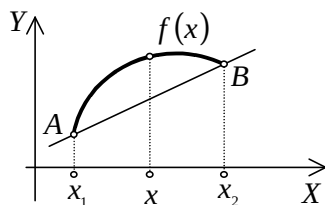


Рис. 4.28

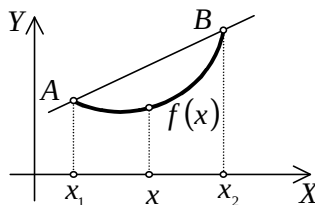


Рис. 4.29

То есть, любая точка хорды AB на рис. 4.29 расположена не ниже графика функции $y = f(x)$.

□ **ТЕОРЕМА 4.29.** Если вторая производная дифференцируемой функции положительна (отрицательна) внутри некоторого промежутка X , то функция выпукла вниз (вверх) на этом промежутке. ■

ОПРЕДЕЛЕНИЕ. *Точкой перегиба* графика непрерывной функции называется точка, в которой график имеет касательную и $\exists$ такая окрестность точки x_0 , в пределах которой график функции имеет разные направления выпуклости.

Если x_0 — точка перегиба функции (рис. 4.30), то точка $(x_0, f(x_0))$ точка перегиба графика функции $f(x)$. В точке $(x_0, f(x_0))$ график функции $f(x)$ переходит с одной стороны наклонной касательной на другую сторону, т.е. в окрестности точки x_0 — разные направления выпуклости.

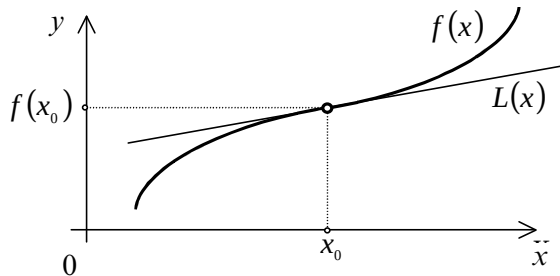


Рис. 4.30

□ **ТЕОРЕМА 4.30** (необходимое условие перегиба). *Вторая производная $f''(x)$ дважды дифференцируемой функции в точке перегиба x_0 равна нулю, т.е. $f''(x_0) = 0$.* ■

□ **ТЕОРЕМА 4.31** (достаточное условие перегиба). *Если вторая производная $f''(x)$ дважды дифференцируемой функции при переходе через некоторую точку x_0 меняет свой знак на противоположный, то точка с координатами $(x_0, f(x_0))$ есть точка перегиба ее графика.* ■

Теорема верна и для случая, когда $f''(x)$ существует в некоторой окрестности точки x_0 , за исключением самой точки x_0 , и существует касательная к графику функции в точке $(x_0, f(x_0))$. Например, функция $f(x) = x^{1/3}$ в точке $x = 0$ имеет бесконечные производные; в точке 0 касательная совпадает с осью Oy . Однако, график этой функции имеет перегиб в начале координат, поскольку вторая производная имеет разные знаки слева и справа от точки $x = 0$.

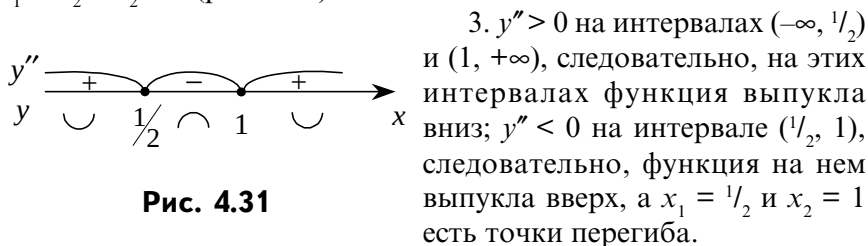
Например, рассмотрим функцию $f(x) = x^3$. Поскольку $f''(x) = 6x$, то $f''(x) < 0$ для любого $x < 0$, $f''(x) > 0$ для любого $x > 0$, значит, функция $f(x) = x^3$ выпукла вниз на интервале $(0, +\infty)$, выпукла вверх на интервале $(-\infty, 0)$. Уравнение касательной к графику в точке $(0, 0)$ имеет вид: $y = 0$. Так как при $x < 0$ $f(x) < 0$, а при $x > 0$ $f(x) > 0$, то точка $x = 0$ является точкой перегиба графика функции $f(x) = x^3$.

▷ П Р И М Е Р 4.28. Найти интервалы выпуклости и точки перегиба графика функции $y = x(x - 1)^3$.

Р е ш е н и е. 1. $y' = (x - 1)^2(4x - 1)$ (см. пример 4.27).

$$2. y'' = 2(x - 1)(4x - 1) + (x - 1)^2 \cdot 4 = 12 \left(x - \frac{1}{2} \right) (x - 1); y'' = 0 \text{ при}$$

$x_1 = 1/2$ и $x_2 = 1$ (рис. 4.31).



4. Значения функции в точках перегиба $f(1/2) = -1/16$, $f(1) = 0$. ►

14.5. Общая схема исследования функций и построение их графиков

При исследовании функций и построении их графиков рекомендуется использовать следующую схему:

1. Найти область определения функции.
2. Исследовать функцию на четность — нечетность, периодичность.
3. Найти вертикальные асимптоты, если они есть.
4. Исследовать поведение функции в бесконечности, найти горизонтальные или наклонные асимптоты, если функция определена на бесконечном интервале.
5. Найти экстремумы и интервалы монотонности функции.

6. Найти интервалы выпуклости функции и точки перегиба.

7. Найти точки пересечения с осями координат и, возможно, некоторые дополнительные точки, уточняющие график.

Заметим, что исследование функции проводится одновременно с построением ее графика.

▷ ПРИМЕР 4.29. Исследовать функцию $y = \frac{1+x^2}{1-x^2}$ и построить ее график.

Решение. 1. Область определения $(-\infty, -1) \cup (-1, 1) \cup (1, +\infty)$, т.е. $x \neq \pm 1$.

2. Функция четная, так как $f(-x) = f(x)$, и ее график симметричен относительно оси ординат. Функция не периодическая.

3. Вертикальные асимптоты могут пересекать ось абсцисс в точках $x = \pm 1$. Так как пределы функции при $x \rightarrow 1 - 0$ (слева) и при $x \rightarrow 1 + 0$ (справа) бесконечны, точнее, $\lim_{x \rightarrow 1+0} \frac{1+x^2}{1-x^2} = -\infty$ и

$\lim_{x \rightarrow 1-0} \frac{1+x^2}{1-x^2} = +\infty$, то прямая $x = 1$ есть вертикальная асимптота.

В силу симметрии графика функции $y = f(x)$ прямая $x = -1$ также является вертикальной асимптотой.

4. Вычислим $\lim_{x \rightarrow +\infty} \frac{1+x^2}{1-x^2} = -1$. В силу четности функции имеем также

$\lim_{x \rightarrow -\infty} \frac{1+x^2}{1-x^2} = -1$, т.е. прямая $y = -1$ есть горизонтальная асимптота.

5. Найдем $y' = \frac{2x(1-x^2) - (1+x^2)(-2x)}{(1-x^2)^2} = \frac{4x}{(1-x^2)^2}$ в области

определения функции; $y' = 0$ при $x = 0$.

Поскольку при $x < 0$ ($x \neq -1$) $f'(x) < 0$, а при $x > 0$ ($x \neq 1$) $f'(x) > 0$ (рис. 4.32), то $x = 0$ — точка минимума и $f_{\min} = f(0) = 1$ — минимум функции. На интервалах $(-\infty, -1)$ и $(-1, 0)$ функция убывает, на интервалах $(0, 1)$ и $(1, +\infty)$ — возрастает.

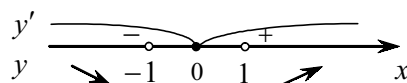


Рис. 4.32

6. Найдем $y'' = \frac{4(1-x^2)^2 - 4x \cdot 2(1-x^2)(-2x)}{(1-x^2)^4} = \frac{4(1+3x^2)}{(1-x^2)^3}$.

Очевидно, что $y'' > 0$ на интервале $(-1, 1)$ и функция здесь выпукла вниз; $y'' < 0$ на интервалах $(-\infty, -1)$, $(1, +\infty)$, и на этих интервалах функция выпукла вверх. Точек перегиба нет.

7. Значение функции $f(0) = 1$, т.е. точка пересечения с осью ординат $(0, 1)$. Уравнение $f(x) = 0$ решений не имеет, следовательно, график функции не пересекает ось абсцисс.

График функции изображен на рис. 4.33. ►

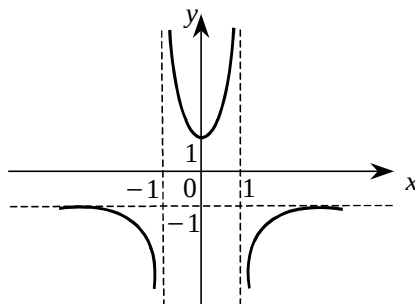


Рис. 4.33

14.6. Теорема о пересечении предельных и средних издержек

Общими издержками называются все затраты производителя на создание продукции, выраженные в денежной форме. Общие издержки подразделяются на постоянные и переменные.

К *постоянным* относятся издержки, которые не меняются на краткосрочном временном интервале и не зависят от количества выпускаемой продукции (например, плата за аренду производственного помещения); обозначаются постоянные издержки $\bar{C}$.

Переменные издержки зависят от объема выпускаемой продукции x . (например, плата за сырье, заработная плата и т.д.); обозначаются $\tilde{C}(x)$.

Общие издержки $C(x) = \bar{C} + \tilde{C}(x)$.

Установлено, что общие издержки могут быть выражены кубической функцией от x (рис. 4.34).

Кроме общих издержек существуют также средние издержки и предельные издержки.

Средние издержки на производство товаров объема x (обозначаются $AVC(x)$) — это общие издержки в расчете на единицу выпуска.

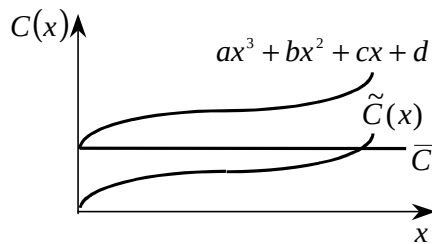


Рис. 4.34

$$AVC(x) = \frac{C(x)}{x}.$$

Предельные издержки (обозначаются $MC(x)$) — это прирост общих издержек при увеличении производства на единицу: $MC(x) = C'(x)$.

Например, пусть зависимость издержек производства от объема выпускаемой продукции выражается формулой $C(x) = 40x - 0,03x^3$.

Определим средние и предельные издержки при объеме продукции $x = 15$ ден. ед.

Функция средних издержек на единицу продукции в нашем случае

$AVC(x) = 40 - 0,03x^2$, откуда $AVC(15) = 40 - 0,03 \cdot 225 = 33,25$ ден. ед.

Предельные издержки определяются как $MC(x) = C'(x) = 40 - 0,09x^2$.

Иными словами, при средних издержках на производство единицы продукции в 33, 25 ден. ед. дополнительные затраты на производство единицы дополнительной продукции составят 19,75 ден. ед. и не превысят средних издержек.

□ ТЕОРЕМА 4.32. *Кривая предельных издержек всегда пересекает кривую средних издержек в точке минимума функции средних издержек.*

Доказательство. Пусть x^* — объем выпускаемой продукции, при котором достигается минимум функции средних издержек (рис. 4.34). По необходимому условию экстремума:

$$\left(\frac{C(x)}{x} \right)' \Big|_{x=x^*} = 0.$$

$$\text{Получаем: } \frac{C'(x^*) \cdot x^* - C(x^*)}{(x^*)^2} = 0 \Rightarrow C'(x^*) = \frac{C(x^*)}{x^*}, \text{ т.е. пре-}$$

дельные издержки в точке минимума функции средних издержек $(x^*, C(x^*))$ равны средним. ■

§ 15. Модели экономического взаимодействия на простейших рынках

15.1. Паутинообразная модель рынка

Мы начали изучать теорию спроса и предложения в подпараграфе 2.4 главы 2, когда знакомились с построением кривых на плоскости. Но спрос и предложение — это прежде всего функции. Об этом мы здесь и поговорим.

1. Спрос. Экономистами спрос изображается в виде графика, показывающего количество продукта, которое потребители готовы и в состоянии купить по некоторой цене из возможных в течение определенного периода времени цен. Речь, таким образом, идет о функции спроса, зависимости количества покупаемого товара от цены. Важно отметить, что речь идет не о какой-то одноразовой покупке, а о потоке товара, уносимого с рынка покупателями. Допустим, на этой неделе цена товара p и ежедневно покупатели покупают D единиц товара, на следующей неделе цена станет p_1 и ежедневно покупатели будут уносить с рынка D_1 единиц товара.

Обозначим $D(p)$ — количество товара, покупаемого на данном рынке за единицу времени при цене p за единицу товара (по-английски “demand” — спрос). Функция $D(p)$ называется *функцией спроса*, или просто *спросом*. Согласно этой функции величина спроса разная при разных ценах. Цены товаров всегда считаем положительными. Фундаментальное свойство функции спроса выражает следующая аксиома.

АКСИОМА СПРОСА. *Функция спроса является убывающей: при увеличении цены величина спроса на товар уменьшается до нуля, при уменьшении цены величина спроса увеличивается.*

Экономисты предлагают разные объяснения этой аксиоме, для нас же достаточно ее самой.

В дальнейшем будем использовать в качестве примеров для иллюстраций следующие функции спроса: а) линейно убывающую

$D(p) = a - bp$, $0 < p < a/b$, $a, b > 0$; б) обратную зависимость $D(p) = 1/p$, $p > 0$; в) логарифмическую $D(p) = \ln[(1 + p)/p]$, $p > 0$. Схематически их графики приведены на рис. 4.35.

При изменении условий на рынке или вне него функция, спроса может измениться, тогда говорят об изменении спроса. Изменение спроса надо отличать от изменения величины спроса при передвижении по графику данной функции спроса. Например, при повышении цен на бензин вполне может повыситься спрос на велосипеды. Это означает, что вся кривая спроса (рис. 4.35) (ее график) передвинется вправо.

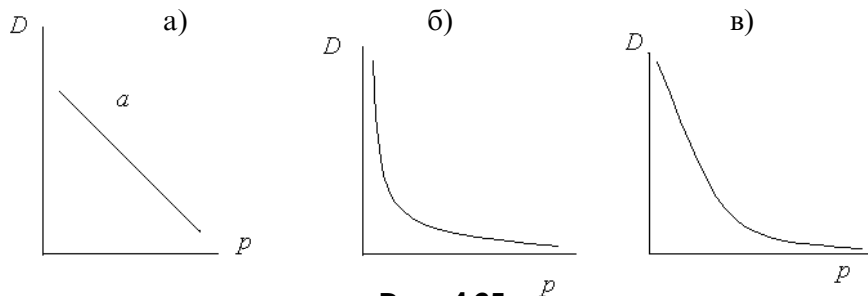


Рис. 4.35

Рассмотрим математические характеристики кривой спроса и их экономические иллюстрации.

Производная функции спроса по цене $D'(p) = dD/dp$ показывает (приблизительно), на сколько изменится величина спроса при изменении цены товара p на 1 единицу. Так как функция спроса предполагается убывающей, то $D'(p) < 0$.

Эластичность спроса по цене показывает, на сколько процентов изменится величина спроса при изменении цены товара p на 1%. Обозначается эластичность E_p^D , она равна $D'(p)/(D(p)/p)$.

▷ ПРИМЕР 4.30. Найти эластичность спроса по цене для каждой из функций, приведенных выше.

Р е ш е н и е.

а) $D'(p) = -b$, так что $E_p^D = -b / [(a - bp)/p] = -bp/(a - bp)$;

б) $D'(p) = -1/p^2$, так что $E_p^D = (-1/p^2)/[(1/p)/p] = -1$;

в) $D'(p) = 1/(1 + p) - 1/p = -1/[p(1 + p)]$, так что

$E_p^D = \{-1/[p(1 + p)]\} / \{\ln[(1 + p)/p]/p\} = -1 / \{(1 + p)\ln[(1 + p)/p]\}$. ►

2. Предложение. Как и в случае со спросом, под предложением товара понимается функция — зависимость количества поставляемого на рынок товара от цены, сложившейся на рынке. Важно отметить, что речь идет не о какой-то одноразовой поставке, а о потоке товара, поставляемого на рынок продавцами. На этой неделе цена товара p , и ежедневно продавцы поставляют S единиц товара. На следующей неделе цена станет p_1 , и ежедневно продавцы будут поставлять на рынок S_1 единиц товара.

Обозначим $S(p)$ количество товара, поставляемого на данный рынок за единицу времени при цене p за единицу товара (по-английски “supply” — предложение). Функция $S(p)$ называется *функцией предложения*, или просто предложением. Согласно этой функции величина предложения разная при разных ценах.

Фундаментальное свойство функции предложения выражает следующая аксиома.

АКСИОМА ПРЕДЛОЖЕНИЯ. *Функция предложения является возрастающей: при увеличении цены величина предложения товара неограниченно увеличивается, при уменьшении цены величина предложения уменьшается, приближаясь к нулю.*

Так же, как и в случае со спросом, существуют объяснения этой аксиомы, для нас же достаточно ее самой.

В дальнейшем будем использовать следующие функции предложения: а) линейно возрастающую $S(p) = -c + lp$, $c/l < p$, $c, l > 0$; б) степенную $S(p) = p^\alpha$, $\alpha > 0$, $p > 0$; в) логарифмическую $S(p) = \ln(1 + p)$. Схематически их графики приведены на рис. 4.36.

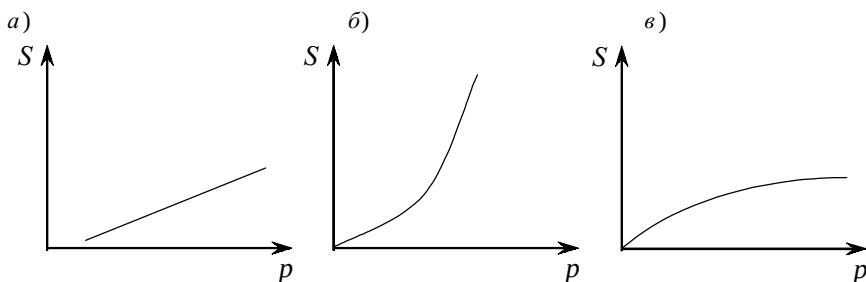


Рис. 4.36

При изменении условий на рынке или вне него функция предложения может измениться, тогда говорят об изменении предложения. Например, при открытии поблизости месторождения алмазов может увеличиться предложение необработанных алмазов, а, возможно, через некоторое время — ювелирных украшений.

Рассмотрим математические характеристики кривой предложения и их экономические иллюстрации.

Производная функции предложения по цене $S'(p) = dS/dp$ показывает приблизительно, насколько изменится величина предложения при изменении цены товара p на 1 единицу. Так как функция предложения товара предполагается возрастающей, то $S'(p) > 0$.

Эластичность предложения по цене показывает, на сколько процентов изменится величина предложения при изменении цены товара на 1%. Обозначается эластичность E_p^S и равна $S'(p)/(S(p)/p)$.

▷ П Р И М Е Р 4.31. Найти эластичность предложения $S(p) = \sqrt{p}$ по цене в точке $p = 4$.

Р е ш е н и е. Имеем $E_p^S(4) = [1/(2\sqrt{p})] / (\sqrt{p}/p) = 1/2$. Для этой функции предложения эластичность оказалась постоянной величиной. ►

3. Равновесие на рынке одного товара. Состояние рынка, при котором спрос равен предложению, мы назвали в подпараграфе 2.4 главы 2 равновесным, а цену, при которой достигается равенство спроса и предложения, — равновесной ценой.

□ У Т В Е Р Ж Д Е Н И Е. Пусть функции спроса и предложения непрерывны и $D(p_0) > S(p_0)$ при некоторой цене p_0 ; тогда существует состояние равновесия.

Д о к а з а т е л ь с т в о. Так как при $p \rightarrow \infty$ $D(p)$ убывает к нулю, а $S(p)$ неограниченно возрастает, то $D(p_1) < S(p_1)$ при некотором $p_1 > p_0$. Величина *неудовлетворенного спроса* $Z(p) = D(p) - S(p)$ принимает на концах отрезка $[p_0, p_1]$ значения разного знака, а так как она непрерывна, то по теореме Больцано-Коши (теорема 4.16) найдется точка p^* отрезка $[p_0, p_1]$, в которой она равна нулю, т.е. $D(p^*) = S(p^*)$. Параметры равновесия снабдим знаком “*”: $p^*, D^* = D(p^*) = S(p^*) = S^*$. Обычно саму тройку (D^*, p^*, S^*) также называют равновесием. ■

▷ П Р И М Е Р 4.32. Найти состояние равновесия для линейных функций спроса $D(p) = a - bp$ и предложения $S(p) = -c + lp$.

Р е ш е н и е. Приравнивая функции, получим $a - bp = -c + lp$, $p^* = (a + c)/(b + l)$, $D^* = (al - bc)/(b + l) = S^*$.

Графическое решение показано на рис. 4.37, а. ►

4. Паутинообразная модель рынка. В реальности нахождение равновесной цены происходит опытным путем, посредством последовательных приближений. Эта процедура называется *паутинообразной моделью рынка*.

Считаем, что функции спроса и предложения удовлетворяют необходимым предположениям о них (см. пункты 1–2). В силу условий, наложенных на функции, уравнение $D(p) = S(p)$ имеет единственное решение p^* и тройка (D^*, p^*, S^*) , где $D^* = D(p^*) = S^* = S(p^*)$, есть единственное состояние равновесия.

Процесс отыскания этого равновесия называют “нащупыванием” (рис. 4.37, б).

Пусть в начальный момент цена на товар была назначена p_0 . Так как спрос больше предложения, т.е. $D_0 > S_0$, то цена увеличивается до p_1 , так чтобы $D_1 = S_0$, т.е. чтобы спрос в следующем периоде понизился до величины предложения в предыдущем периоде. Если спрос меньше предложения, т.е. $D_1 < S_1$, то цена уменьшается до p_2 и т.д.

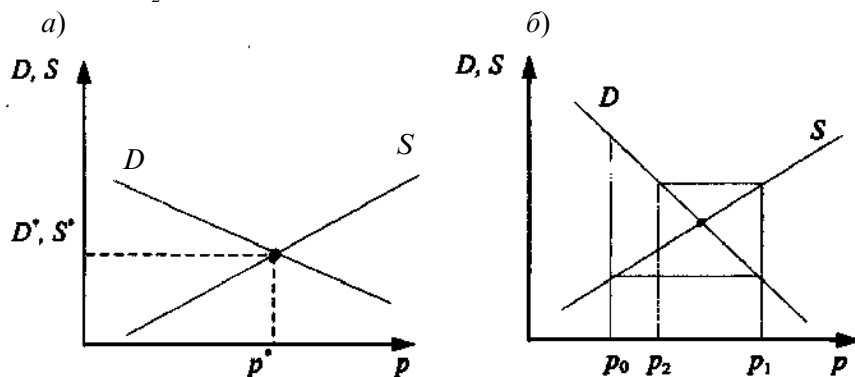


Рис. 4.37

З а м е ч а н и е. Не все так просто. Для сходимости этого процесса к состоянию равновесия необходимо соблюдение некоторых условий. Пусть, например, функции спроса и предложения линейны: $D(p) = a - bp$, $S(p) = -c + lp$. Тогда для сходимости необходимо и достаточно выполнения условия $l < b$, где l и b — угловые коэффициенты, т.е. линия спроса должна быть более наклонной, чем линия предложения. Исходя из этого в общем случае, т.е. для нелинейных функций спроса и предложения, хотя бы вблизи равновесной цены p^* , должно быть выполнено аналогичное условие $S'(p^*) < D'(p^*)$.

Если в некоторой окрестности равновесной цены процесс сходится к состоянию равновесия при любом начальном значении цены из этой окрестности, то состояние равновесия называется *устойчивым*. Если же, наоборот, существует окрестность равновесной цены такая, что при любом начальном значении цены из этой окрестности, отличном от равновесного, процесс не сходится к состоянию равновесия, то такое состояние равновесия называется *неустойчивым*. При линейных функциях спроса и предложения состояние равновесия устойчивое, если $l < b$, и неустойчивое, если $b \leq l$.

15.2. Модель Солоу экономического роста

В качестве еще одного примера использования в экономике аппарата дифференциального исчисления рассмотрим **модель Солоу**.

Предложение товаров, или *национальный доход*, пусть описывается производственной функцией (см. подпараграф 6.3 главы 4) $Y = F(K, L)$, где K, L — запасы труда и капитала.

Под экономическим ростом понимаем увеличение во времени национального дохода в расчете на душу населения (работающего).

Модель Солоу предполагает, что $F(K, L)$ обладает постоянной отдачей от масштаба, т.е. увеличение труда и капитала в Z раз вызывает рост национального дохода тоже в Z раз.

Можно доказать, что национальный доход в расчете на душу населения является функцией величины K/L , которая называется *капиталовооруженностью* и обозначается k :

$$k = K/L.$$

Обозначим через y национальный доход на душу населения, тогда $y = f(k)$ (рис. 4.38).

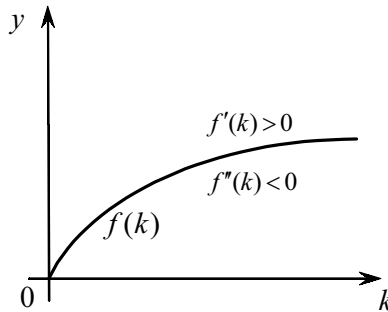


Рис. 4.38

Производную $f'(k)$ назовем *предельной производительностью капитала*, которая характеризует прирост национального дохода на душу населения при единичном увеличении капиталовооруженности k .

Отрицательное значение $f''(k)$ показывает, что по мере роста количества капитала на одного человека каждая дополнительная единица капитала приносит все меньший эффект.

В модели Солоу *спрос на товары* описывается так (правительственные расходы не рассматриваются):

$$Y = C + I,$$

где C — общественное потребление, а I — инвестиции.

В расчете на душу населения $y = c + i$, где $c = C/L$, а $i = I/L$.

В то же время известно, что общественное потребление $C = (1 - s)Y$, где s — склонность к сбережению, причем $s = 1 - c$, а Y — национальный доход (общий). Тогда в расчете на душу населения $c = (1 - s)y$, $1 - s = \text{const}$.

Подставим полученное выражение в $y = c + i$, тогда

$$y = (1 - s)y + i \Rightarrow i = sy. \quad (4.81)$$

Формула (4.81) означает, что инвестиции на душу населения возрастают с ростом национального дохода.

Запасы капитала могут измениться по двум причинам:

1) новые инвестиции (i) вызывают увеличение запасов капитала;

2) амортизация, или выбытие капитала, вызывает снижение запасов.

Через δ обозначается процент капитала, который выбывает ежегодно, тогда изменение k можно записать так:

$$\Delta k = i - \delta k, \quad (4.82)$$

где i — инвестиции, сделанные за это же время, δk — количество выбывшего капитала на душу населения, т.е. выбытие капитала.

Формулу (4.82) можно записать в виде $\Delta k = s f(k) - \delta k$ [используя формулу (4.81)].

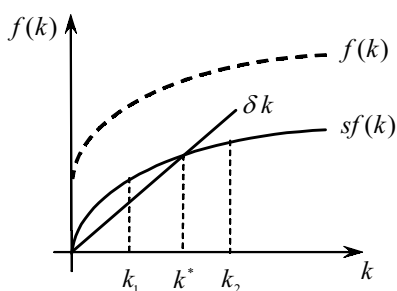


Рис. 4.39

На рис. 4.39 точка k^* — точка равновесия, в которой инвестиции $i = s f(k)$ равны выбытию капитала δk . Если экономика достигнет точки k^* , то уровень k перестанет изменяться, т.е. $\Delta k = 0$. Это состояние называют *устойчивой капиталовооруженностью*.

Уровень k^* — устойчивое равновесие (в математическом понимании), т.е. отклонение от него приводит к возврату.

Предположим, что в некоторый момент времени уровень капиталовооруженности k_1 ниже k^* . В этом случае инвестиции превышают выбытие, наблюдается рост k до состояния k^* . Если же, напротив, начальный уровень капиталовооруженности k_2 выше, чем k^* , то выбытие превышает инвестиции и уровень k снижается до k^* .

§ 16. Понятие функции нескольких переменных. Модель фирмы

До сих пор мы изучали совместное изменение двух переменных, из которых одна зависела от другой. Нередки, однако, случаи, когда независимых переменных оказывается несколько.

Пусть производственная фирма выпускает один вид продукции или много видов. Тогда годовой выпуск фирмы в натурально-вещественной форме Y — это число единиц продукции одного вида или чисто многономенклатурных агрегатов.

Для производства продукции такая фирма использует *настоящий* труд L (среднее число занятых в год либо отработанные за год человекочасы) и *прошлый* труд K в виде средств труда (основные производственные фонды) и предметов труда (затраченное за год топливо, энергия, сырье, материалы, комплектующие и т.п.).

Каждый из этих двух агрегированных видов ресурсов (труд L , фонды и материалы K) имеет определенное число разновидностей (труд разной квалификации, оборудование различного вида и т.п.). Обозначим вектор-столбец возможных объемов затрат различных видов ресурсов через $x = (x_1, x_2, \dots, x_m)$. Тогда технология фирмы определяется ее производственной функцией $Y = F(K, L)$.

16.1. Функция m переменных и область ее определения

Рассмотрим арифметическое m -мерное пространство (см. подпараграф 1.2 главы 4). Говоря об изменении m независимых переменных $x_1, \dots, x_m$, мы должны всякий раз указывать, какие m значений $(x_1, \dots, x_m)$ они могут принимать совместно. Множество μ таких точек и будет областью определения переменных $x_1, \dots, x_m$.

ОПРЕДЕЛЕНИЕ. Переменная z (с областью изменения Z) называется *функцией независимых переменных* $x_1, \dots, x_m$ на множестве μ , если каждой точке $(x_1, \dots, x_m)$ из μ по некоторому правилу (или закону) ставится в соответствие одно определенное значение z (из Z).

Множество μ — область определения функции; сами переменные $x_1, \dots, x_m$ по отношению к их функции z называются ее аргументами. Функциональная зависимость будет иметь вид: $z = f(x_1, \dots, x_m)$.

16.2. Предел функции нескольких переменных

Рассмотрим в m -мерном пространстве последовательность точек

$$\{M_n(x_1^{(n)}, x_2^{(n)}, \dots, x_m^{(n)})\} \quad n=1, 2, 3, \dots$$

Будем говорить, что эта последовательность сходится к предельной точке $M_0(a_1, a_2, \dots, a_m)$, если координаты точки M_n порознь стремятся к соответствующим координатам точки M_0 , т.е. если при $n \rightarrow \infty$ $x_1^{(n)} \rightarrow a_1, x_2^{(n)} \rightarrow a_2, \dots, x_m^{(n)} \rightarrow a_m$.

Говорят, что функция $f(x_1, x_2, \dots, x_m) = f(M)$ имеет пределом число A при стремлении переменных $x_1, x_2, \dots, x_m$ к $a_1, a_2, \dots, a_m$, если, какую бы ни извлечь из множества μ последовательность (4.83) отличных от $M_0(a_1, a_2, \dots, a_m)$ точек, сходящуюся к M_0 , числовая последовательность $\{f(x_1^{(n)}, \dots, x_m^{(n)})\} = \{f(M_n)\}$, состоящая из соответствующих значений функции, всегда сходится к A :

$$A = \lim_{\substack{x_1 \rightarrow a_1 \\ \dots \\ x_m \rightarrow a_m}} f(x_1, x_2, \dots, x_m) \text{ или } A = \lim_{M \rightarrow M_0} f(M).$$

Говорят, что функция $f(x_1, x_2, \dots, x_m)$ имеет пределом число A при стремлении $x_1, x_2, \dots, x_m$ соответственно к $a_1, a_2, \dots, a_m$, если для каждого числа $\varepsilon > 0$ найдется такое число $\delta > 0$, что $|f(x_1, x_2, \dots, x_m) - A| < \varepsilon$, лишь только $|x_1 - a_1| < \delta, \dots, |x_m - a_m| < \delta$. При этом предполагается, что точка $(x_1, x_2, \dots, x_m)$ взята из μ и отлична от $(a_1, a_2, \dots, a_m)$.

Введем для точек $(x_1, x_2, \dots, x_m)$ и $(a_1, a_2, \dots, a_m)$ обозначения M и M_0 . Тогда в геометрических терминах можно перефразировать сказанное так: число A называется *пределом функции* $f(M)$ при стремлении точки M к M_0 , если для каждого числа $\varepsilon > 0$ существует такое $r > 0$: $|f(M) - A| < \varepsilon$, лишь только расстояние $\overline{M_0 M} < r$.

§ 17. Дифференцирование функций нескольких переменных. Производная по направлению градиента

17.1. Частные производные

Пусть в некоторой (открытой) области $D \in \mathbf{R}^3$ имеем функцию $u = f(x, y, z)$. Возьмем точку $M_0 = (x_0, y_0, z_0)$ в этой области. Если зафиксируем $y = y_0, z = z_0$ и будем изменять x , то u будет функцией от одной переменной x в окрестности x_0 .

Можно вычислить производную функции $u = f(x, y_0, z_0)$ в точке x_0 . Дадим этому значению x_0 приращение Δx , тогда функция получит приращение $\Delta_x u = f(x_0 + \Delta x, y_0, z_0) - f(x_0, y_0, z_0)$, которое можно было бы назвать ее частным приращением (по x). По определению производной находим (если существует) предел:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta_x u}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x, y_0, z_0) - f(x_0, y_0, z_0)}{\Delta x}. \quad (4.83)$$

Этот предел называется *частной производной* функции $f(x, y, z)$ по x в точке (x_0, y_0, z_0) . Частную производную будем обозначать:

$$\frac{\partial u}{\partial x}, \quad \frac{\partial f(x_0, y_0, z_0)}{\partial x}, \quad u'_x, \quad f'_x, \quad f'_x(M_0).$$

Если x и z — постоянны, y — переменная, тогда

$$\lim_{\Delta y \rightarrow 0} \frac{\Delta_y u}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{f(x_0, y_0 + \Delta y, z_0) - f(x_0, y_0, z_0)}{\Delta y} = \frac{\partial u}{\partial y}.$$

Если x и y — постоянны, z — переменная, тогда

$$\lim_{\Delta z \rightarrow 0} \frac{\Delta_z u}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{f(x_0, y_0, z_0 + \Delta z) - f(x_0, y_0, z_0)}{\Delta z} = \frac{\partial u}{\partial z}.$$

17.2. Полное приращение функции

Если исходя из значений $x = x_0, y = y_0, z = z_0$ независимых переменных дать всем трем некоторые приращения $\Delta x, \Delta y, \Delta z$, то функ-

ция $u = f(x, y, z)$ получит приращение $\Delta u = \Delta f(x_0, y_0, z_0) = f(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) - f(x_0, y_0, z_0)$, которое называется *полным приращением функции*.

Для функции $y = f(x)$ от одной переменной в предположении существования в точке x_0 конечной производной $f'(x_0)$ приращение функции

$$\Delta y = \Delta f(x_0) = f'(x_0) \cdot \Delta x + \alpha \cdot \Delta x, \alpha \rightarrow 0 \text{ при } \Delta x \rightarrow 0.$$

Для функции $u = f(x, y, z)$ аналогично получается:

$$\begin{aligned} \Delta u = \Delta f(x_0, y_0, z_0) &= f'_x(x_0, y_0, z_0)\Delta x + f'_y(x_0, y_0, z_0)\Delta y + \\ &+ f'_z(x_0, y_0, z_0)\Delta z + \alpha\Delta x + \beta\Delta y + \gamma\Delta z, \end{aligned} \quad (4.84)$$

где $\alpha, \beta, \gamma \rightarrow 0$ при $\Delta x, \Delta y, \Delta z \rightarrow 0$.

Если частные производные $f'_x(x, y, z)$, $f'_y(x, y, z)$, $f'_z(x, y, z)$ существуют не только в точке (x_0, y_0, z_0) , но и в некоторой ее окрестности и, кроме того, непрерывны (как функции от x, y, z) в этой точке, то имеет место формула (4.84).

17.3. Полный дифференциал

Пусть

$$\begin{aligned} \Delta u = \Delta f(x_0, y_0, z_0) &= f'_x(x_0, y_0, z_0)\Delta x + f'_y(x_0, y_0, z_0)\Delta y + \\ &+ f'_z(x_0, y_0, z_0)\Delta z + o(\rho), \end{aligned}$$

где $o(\rho)$ — бесконечно малая функция более высокого порядка по сравнению с $\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$, тогда функция $f(x, y, z)$ называется *дифференцируемой в точке* (x_0, y_0, z_0) и выражение

$$\begin{aligned} u'_x\Delta x + u'_y\Delta y + u'_z\Delta z &= f'_x(x_0, y_0, z_0)\Delta x + f'_y(x_0, y_0, z_0)\Delta y + \\ &+ f'_z(x_0, y_0, z_0)\Delta z \end{aligned}$$

называется ее *полным дифференциалом* и обозначается du или $df(x_0, y_0, z_0)$:

$$du = u'_x dx + u'_y dy + u'_z dz,$$

$$df(x_0, y_0, z_0) = f'_x(x_0, y_0, z_0)dx + f'_y(x_0, y_0, z_0)dy + f'_z(x_0, y_0, z_0)dz.$$

Очевидно, что если $u = f(x)$ — функция одной переменной, то $du = u'_x dx$ — ее дифференциал.

Вернемся к *модели фирмы*. Технология фирмы определяется ее *производственной функцией* (см. подпараграф 6.3 главы 4), выражающей связь между затратами ресурсов (K и L) и выпуском: $Y = F(K, L)$. В общем случае $Y = F(X)$, где $X = (x_1, \dots, x_m)$.

Мы уже говорили, что производственная функция F удовлетворяет двум аксиомам (см. подпараграф 6.3).

Первая из них утверждает, что существует подмножество пространства затрат, называемое экономической областью E , в которой увеличение любого вида затрат не приводит к уменьшению выпуска. Таким образом, если X_1, X_2 — две точки этой области, то $X_1 \geq X_2$ влечет $F(X_1) \geq F(X_2)$.

В дифференциальной форме это выражается в том, что в этой области все первые частные производные функции F неотрицательны: $\partial F / \partial x_i \geq 0$. Эти производные называются *предельными продуктами*, а вектор $\partial F / \partial x = (\partial F / \partial x_1, \dots, \partial F / \partial x_m)$ — *вектором предельных продуктов*.

Вторая аксиома утверждает, что существует выпуклое подмножество S экономической области, для которой подмножества $\{X \in S: F(X) \geq a\}$ выпуклы для всех $a \geq 0$. В этом подмножестве S должны быть отрицательны все вторые производные функции $\partial^2 F / \partial x_i^2 < 0$ для любого $i = 1, \dots, m$.

Остановимся на экономическом содержании этих аксиом.

Первая аксиома утверждает, что производственная функция не какая-то совершенно абстрактная функция, придуманная теоретиком-математиком. Она, пусть и не на всей своей области определения, а только лишь на ее части, отражает экономически важное, бесспорное и в то же время тривиальное утверждение: в любой разумной экономике увеличение затрат не может привести к уменьшению выпуска.

Из второй аксиомы поясним только экономический смысл требования, чтобы производная $\partial^2 F / \partial x_i^2$ была меньше нуля для каждого вида затрат. Это свойство называется в экономике *зако-*

ном убывающей отдачи или убывающей доходности: по мере увеличения затрат начиная с некоторого момента (при входе в область S) уменьшается предельный продукт.

Классическим примером этого закона является добавление все большего и большего количества труда в производство зерна на фиксированном участке земли.

В дальнейшем подразумевается, что производственная функция рассматривается на области S , в которой обе аксиомы справедливы.

17.4. Поверхности уровня

Пусть в пространстве (x, y, z) имеется область D , в которой задана функция

$$u = u(x, y, z)$$

В этом случае говорят, что в области D задано скалярное поле. Если, например, $u(x, y, z)$ обозначает температуру в точке $M(x, y, z)$, то говорят, что задано скалярное поле температур; если область D заполнена жидкостью или газом и $u(x, y, z)$ обозначает давление, то имеется скалярное поле давлений и т.д.

Рассмотрим точки области D , в которых функция $u(x, y, z)$ имеет постоянное значение c :

$$u(x, y, z) = c.$$

Совокупность этих точек образует некоторую поверхность. Такие поверхности называются *поверхностями уровня*.

Пусть задано скалярное поле $u(x, y, z) = \frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{16}$. Здесь поверхностями уровня будут поверхности $\frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{16} = c$, которые называются эллипсоидами с полуосями $2\sqrt{c}$, $3\sqrt{c}$, $4\sqrt{c}$.

Если функция u есть функция двух переменных: $u = u(x, y)$, то “поверхностями” уровня будут линии на плоскости Oxy :

$$u(x, y) = c,$$

которые называются *линиями уровня*.

Если значения u мы будем откладывать по оси Oz : $z = u(x, y)$, то линиями уровня на плоскости Oxy будут проекции линий, которые получаются в пересечении поверхности $z = u(x, y)$ с плоскостями $z = c$.

Линиями уровня функции $z = 1 - x^2 - y^2$ будут линии, задаваемые уравнениями $1 - x^2 - y^2 = c$. Это – окружности с радиусом $\sqrt{1 - c}$.

В частности, при $c = 0$ получаем окружность $x^2 + y^2 = 1$.

17.5. Производная по направлению

Рассмотрим в области D функцию $u = u(x, y, z)$ и точку $M(x, y, z)$. Проведем из точки M вектор S , направляющие косинусы которого – $\cos\alpha, \cos\beta, \cos\gamma$. На векторе S на расстоянии ΔS от его начала рассмотрим точку $M_1(x + \Delta x, y + \Delta y, z + \Delta z)$ (рис. 4.40).

Таким образом, $\Delta S = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$.

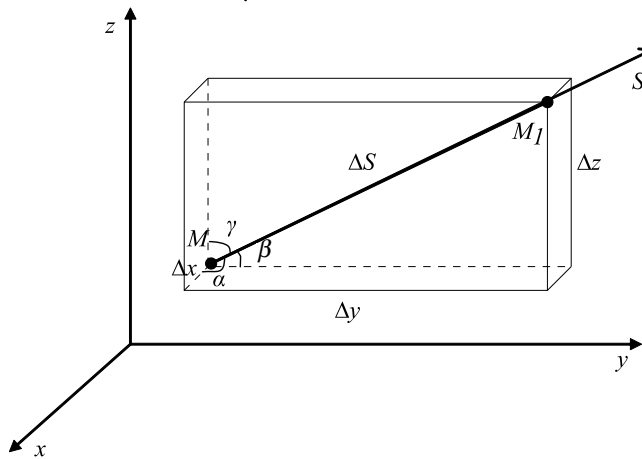


Рис. 4.40

Будем предполагать, что функция $u(x, y, z)$ непрерывна и имеет непрерывные производные по своим аргументам в области D .

Полное приращение функции представим так:

$$\Delta u = \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y + \frac{\partial u}{\partial z} \Delta z + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y + \varepsilon_3 \Delta z,$$

где $\varepsilon_1, \varepsilon_2, \varepsilon_3 \rightarrow 0$ при $\Delta S \rightarrow 0$.

Разделим все члены полученного равенства на ΔS :

$$\frac{\Delta u}{\Delta S} = \frac{\partial u}{\partial x} \frac{\Delta x}{\Delta S} + \frac{\partial u}{\partial y} \frac{\Delta y}{\Delta S} + \frac{\partial u}{\partial z} \frac{\Delta z}{\Delta S} + \varepsilon_1 \frac{\Delta x}{\Delta S} + \varepsilon_2 \frac{\Delta y}{\Delta S} + \varepsilon_3 \frac{\Delta z}{\Delta S}.$$

Получим

$$\frac{\Delta x}{\Delta S} = \cos \alpha, \quad \frac{\Delta y}{\Delta S} = \cos \beta, \quad \frac{\Delta z}{\Delta S} = \cos \gamma.$$

Значит можно записать так:

$$\begin{aligned} \frac{\Delta u}{\Delta S} &= \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma + \varepsilon_1 \cos \alpha + \\ &+ \varepsilon_2 \cos \beta + \varepsilon_3 \cos \gamma. \end{aligned} \quad (4.85)$$

ОПРЕДЕЛЕНИЕ. Предел отношения $\frac{\Delta u}{\Delta S}$ при $\Delta S \rightarrow 0$ называется *производной от функции $u = u(x, y, z)$ в точке (x, y, z) по направлению вектора S* и обозначается $\frac{\partial u}{\partial S}$, т.е. $\lim_{\Delta S \rightarrow 0} \frac{\Delta u}{\Delta S} = \frac{\partial u}{\partial S}$.

Таким образом, переходя к пределу в равенстве (4.85), получим:

$$\frac{\partial u}{\partial S} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma.$$

Из приведенной формулы следует, что, зная частные производные, легко найти производную по любому направлению S . Сами частные производные являются частным случаем производной по направлению.

Так, например, при $\alpha = 0$, $\beta = \frac{\pi}{2}$, $\gamma = \frac{\pi}{2}$, получаем:

$$\frac{\partial u}{\partial S} = \frac{\partial u}{\partial x} \cos 0 + \frac{\partial u}{\partial y} \cos \frac{\pi}{2} + \frac{\partial u}{\partial z} \cos \frac{\pi}{2} = \frac{\partial u}{\partial x}.$$

▷ **ПРИМЕР 4.33.** Для заданной функции $u = x^2 + y^2 + z^2$ найти производную $\frac{\partial u}{\partial S}$ в точке $M(1, 1, 1)$ в двух направлениях:

а) в направлении вектора $S_1 = 2\bar{i} + \bar{j} + 3\bar{k}$;

б) в направлении вектора $S_2 = \bar{i} + \bar{j} + \bar{k}$.

Р е ш е н и е.

а) Находим направляющие косинусы вектора S_1 :

$$\cos \alpha = \frac{2}{\sqrt{4+1+9}} = \frac{2}{\sqrt{14}}; \quad \cos \beta = \frac{1}{\sqrt{14}}; \quad \cos \gamma = \frac{3}{\sqrt{14}}.$$

Тогда

$$\frac{\partial u}{\partial S_1} = \frac{\partial u}{\partial x} \frac{2}{\sqrt{14}} + \frac{\partial u}{\partial y} \frac{1}{\sqrt{14}} + \frac{\partial u}{\partial z} \frac{3}{\sqrt{14}};$$

$$\frac{\partial u}{\partial x} = 2x; \quad \frac{\partial u}{\partial y} = 2y; \quad \frac{\partial u}{\partial z} = 2z.$$

В точке $M(1, 1, 1)$ найдем значения частных производных:

$$\left. \frac{\partial u}{\partial x} \right|_M = 2; \quad \left. \frac{\partial u}{\partial y} \right|_M = 2; \quad \left. \frac{\partial u}{\partial z} \right|_M = 2.$$

$$\text{Следовательно, } \frac{\partial u}{\partial S_1} = 2 \frac{2}{\sqrt{14}} + 2 \frac{1}{\sqrt{14}} + 2 \frac{3}{\sqrt{14}} = \frac{12}{\sqrt{14}}.$$

б) Находим направляющие косинусы вектора S_2 :

$$\cos \alpha = \frac{1}{\sqrt{3}}; \quad \cos \beta = \frac{1}{\sqrt{3}}; \quad \cos \gamma = \frac{1}{\sqrt{3}}.$$

Значит, частная производная в направлении вектора S_2 имеет вид

$$\frac{\partial u}{\partial S_2} = 2 \frac{1}{\sqrt{3}} + 2 \frac{1}{\sqrt{3}} + 2 \frac{1}{\sqrt{3}} = \frac{6}{\sqrt{3}} = 2\sqrt{3}. \quad \blacktriangleright$$

17.6. Градиент

В каждой точке области D , в которой задана функция $u = u(x, y, z)$, определим вектор, проекциями которого на оси координат явля-

ются значения частных производных $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial u}{\partial z}$ этой функции в соответствующей точке:

$$\text{grad } u = \frac{\partial u}{\partial x} \bar{i} + \frac{\partial u}{\partial y} \bar{j} + \frac{\partial u}{\partial z} \bar{k}. \quad (4.86)$$

ОПРЕДЕЛЕНИЕ. Вектор вида (4.86) называется *градиентом* функции $u(x, y, z)$.

Говорят, что в области D определено векторное поле градиентов.

□ **ТЕОРЕМА 4.34.** Пусть дано скалярное поле $u = u(x, y, z)$, и в этом скалярном поле определено поле градиентов

$\text{grad } u = \frac{\partial u}{\partial x} \bar{i} + \frac{\partial u}{\partial y} \bar{j} + \frac{\partial u}{\partial z} \bar{k}$. Тогда производная $\frac{\partial u}{\partial S}$ по направлению некоторого вектора S равняется проекции вектора $\text{grad } u$ на вектор S . ■

Предположим, задана функция $u = x^2 + y^2 + z^2$. Необходимо:

- а) определить градиент в точке $M(1, 1, 1)$;
- б) определить производную от функции $u(x, y, z)$ в точке $M(1, 1, 1)$ в направлении градиента.

Решение. а) Используя формулу (4.86), получим $\text{grad } u = 2x\bar{i} + 2y\bar{j} + 2z\bar{k}$.

Значит, в точке M $\text{grad } u|_M = 2\bar{i} + 2\bar{j} + 2\bar{k}$. Тогда

$$|\text{grad } u|_M = 2\sqrt{3}.$$

Для определения производной от функции $u(x, y, z)$ в точке $M(1, 1, 1)$ в направлении градиента сначала найдем направляющие косинусы градиента:

$$\cos \alpha = \frac{2}{\sqrt{2^2 + 2^2 + 2^2}} = \frac{1}{\sqrt{3}}; \quad \cos \beta = \frac{1}{\sqrt{3}}; \quad \cos \gamma = \frac{1}{\sqrt{3}}.$$

Тогда

$$\frac{\partial u}{\partial S} = 2 \frac{1}{\sqrt{3}} + 2 \frac{1}{\sqrt{3}} + 2 \frac{1}{\sqrt{3}} = 2\sqrt{3}.$$

Мы получили равенство $\frac{\partial u}{\partial S} = |\text{grad } u|$, показывающее, что производная в данной точке по направлению вектора S совпадает с направлением градиента; это наибольшее значение производной равно $|\text{grad } u|$. Можно доказать, что производная по направлению вектора, перпендикулярного вектору $\text{grad } u$, равна нулю.

17.7. Производственная функция Кобба-Дугласа

В качестве примера рассмотрим одну из наиболее распространенных производственных функций — функцию Кобба-Дугласа: $Y = AK^\alpha L^\beta$, где $A, \alpha, \beta > 0$ — константы, $\alpha + \beta < 1$; K — объем фондов либо в стоимостном выражении, либо в натуральном количестве, скажем, — число станков; L — объем трудовых ресурсов также либо в стоимостном выражении, либо в натуральном количестве — число рабочих, число человеко-дней и т.п. и, наконец, Y — выпуск продукции в стоимостном или натуральном выражении.

Проверим, выполняются ли для нее основные требования к производственным функциям.

1. Положительность предельных продуктов имеем, так как $\partial Y / \partial K = A\alpha K^{\alpha-1} L^\beta > 0$, $\partial Y / \partial L = A\beta K^\alpha L^{\beta-1} > 0$.

2. Отрицательность вторых частных производных, т.е. убывание предельных продуктов также имеет место:

$$\partial^2 Y / \partial K^2 = A\alpha(\alpha - 1)K^{\alpha-2} L^\beta < 0, \quad \partial^2 Y / \partial L^2 = A\beta(\beta - 1)K^\alpha L^{\beta-2} < 0.$$

Перейдем к основным экономико-математическим характеристикам производственной функции Кобба-Дугласа.

ОПРЕДЕЛЕНИЕ. *Средняя производительность труда* $y = Y/L$ — отношение объема произведенного продукта к количеству затраченного труда; *средняя фондоотдача* $k = Y/K$ — это отношение объема произведенного продукта к величине фондов.

Для функции Кобба-Дугласа средняя производительность труда $y = AK^\alpha L^{\beta-1}$ является убывающей функцией L , т.е. с увеличени-

ем затрат труда средняя производительность труда падает. Этот вывод допускает естественное объяснение — поскольку величина второго фактора K остается неизменной, то, значит, вновь привлекаемая рабочая сила не обеспечивается дополнительными средствами производства, что и приводит к снижению производительности труда (это справедливо и в самом общем случае — на уровне производственных множеств, см. подпараграф 2.2 главы 4).

Предельная производительность труда есть частная производная функции по труду $\partial Y/\partial L = A\beta K^\alpha L^{\beta-1}$, откуда видно, что для функции Кобба-Дугласа предельная производительность труда пропорциональна средней производительности и меньше нее. Аналогично определяются средняя и предельная фондоотдачи. Для них также справедливо указанное соотношение — предельная фондоотдача пропорциональна средней фондоотдаче и меньше нее.

Найдем теперь эластичность продукции по труду $(\partial Y/\partial L) : (Y/L) = (\partial Y/\partial L)L/Y = A\beta K^\alpha L^{\beta-1}L/(AK^\alpha L^\beta) = \beta$.

Теперь нам ясен смысл параметра β — это *эластичность продукции по труду*. Аналогичный смысл имеет параметр α — это *эластичность продукции по фондам*.

И еще одно замечание представляется интересным. Пусть $\alpha + \beta = 1$. Легко проверить, что $Y = (\partial Y/\partial K)K + (\partial Y/\partial L)L$. Будем считать, что общество состоит только из рабочих и предпринимателей. Тогда доход X распадается на две части — доход рабочих и доход предпринимателей. Поскольку при оптимальном размере фирмы величина $\partial Y/\partial L$ — предельный продукт по труду — совпадает с заработной платой, то $(\partial Y/\partial L)L$ представляет собой доход рабочих. Аналогично величина $\partial Y/\partial K$ есть *предельная фондоотдача*, экономический смысл которой есть норма прибыли, следовательно, $(\partial Y/\partial K)K$ представляет доход предпринимателей.

Вообще, понятие предельных показателей, связанных с нахождением частных производных, в экономической теории играет особую роль, так как имеет строго определенный экономический смысл.

17.8. Теория фирмы

Теперь введем в рассмотрение цены. Пусть P — вектор цен. Если $T = (X, Y)$ (см. подпараграф 2.2 главы 4) — технология, т.е. вектор “затраты-выпуск”, X — затраты, Y — выпуск, то их ска-

лярное произведение $PT = PX + PY$ есть прибыль от использования технологии T (затраты — отрицательные величины).

Задача производителя состоит в том, что он выбирает, ищет технологию из своего производственного множества, стремясь *максимизировать* прибыль.

Итак, производитель решает следующую задачу:

$$PT \rightarrow \max, T \in \tau. \quad (4.87)$$

Это резко упрощает ситуацию выбора. Так, если цены положительны, что естественно, то компонента “выпуск” решения этой задачи автоматически будет лежать на кривой производственных возможностей.

Действительно, пусть $T = (X, Y)$ — какое-нибудь решение задачи (4.87). Тогда существует $Z \in K_x$, где K_x — *кривая производственных возможностей* — множество лучших выпусков при данных затратах X , $Z \geq Y$, следовательно, $P(X, Z) \geq P(X, Y)$, значит, точка (X, Z) также есть решение задачи производителя.

Для случая двух видов затрат задачу (4.87) можно решить графически (рис. 4.41), так как для любого вектора затрат X все наилучшие выпуски лежат на кривой производственных возможностей. Для этого надо “двигать” прямую линию, перпендикулярную вектору P , в направлении, куда он показывает; тогда последняя точка, когда эта прямая линия еще пересекает производственное множество, и будет оптимальным решением (на рис. 4.41 это точка T).

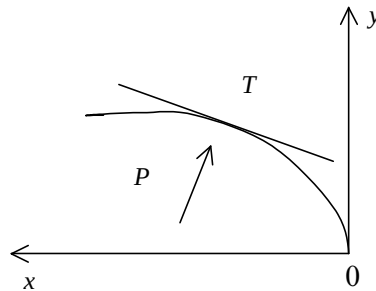


Рис. 4.41

Мы видим, что строгая выпуклость нужной части производственного множества гарантирует единственность решения.

Примерно такие же рассуждения действуют и в общем случае, для большего числа видов затрат и выпуска.

Однако мы не пойдем по этому пути, а используем аппарат теории производственных функций и производителя.

Итак, выпуск фирмы можно охарактеризовать одной величиной — либо объемом выпуска, если выпускается один товар, либо суммарной стоимостью всего выпуска. Пространство затрат m -мерно, вектор затрат $X = (x_1, \dots, x_m)$. Затраты однозначно определяют выпуск Y , а эта связь и есть производственная функция $Y = F(X)$.

В данной ситуации обозначим через P вектор цен на товары — затраты, и пусть v — цена единицы выпускаемого товара. Следовательно, прибыль W , являющаяся в итоге функцией X (и цен, но они считаются постоянными), есть

$$W(X) = vY - PX = vF(X) - PX.$$

Приходим к задаче фирмы: $W(X) = vF(X) - PX \rightarrow \max, X \geq 0$.

Приравнивая частные производные функции W нулю, получим: $v(\partial F / \partial x_j) = p_j$ для $j = 1, \dots, m$ или $v(\partial F / \partial X) = P$. (4.88)

Будем предполагать, что все затраты строго положительны (нулевые можно просто исключить из рассмотрения). Тогда точка, определяемая соотношением (4.88), оказывается внутренней, т.е. точкой экстремума. И поскольку еще предполагается отрицательность вторых производных функции F (см. требования к производственным функциям подпараграфе 17.2), то это точка максимума.

Итак, при естественных предположениях на производственную функцию (эти предположения выполняются для производителя со здравым смыслом и в разумной экономике) соотношение (4.88) дает решение задачи фирмы, т.е. определяет объем X^* перерабатываемых ресурсов, в результате чего получается выпуск $Y^* = F(X^*)$. Точку X^* , или $(X^*, F(X^*))$, назовем *оптимальным решением фирмы*. Остановимся на экономическом смысле соотношения (4.88).

Как говорилось ранее, но в наших обозначениях, $\partial F / \partial K = (\partial F / \partial x_1, \dots, \partial F / \partial x_m)$ называется предельным вектором-продуктом, или вектором предельных продуктов, а $\partial F / \partial x_i$ называется i -м предельным продуктом, или откликом выпуска на изменение i -го то-

вара затрат. Следовательно, $v (\partial F / \partial x_i) dx_i$ — это стоимость i -го предельного продукта, дополнительно полученного из dx_i единиц i -го ресурса. Однако стоимость dx_i единиц i -го ресурса равна $p_i dx_i$, т.е. получилось равновесие: можно вовлечь в производство дополнительно dx_i единиц i -го ресурса, потратив на его закупку $p_i dx_i$, но выигрыша не будет, так как после переработки продукции получим ровно такую же сумму, какую и затратили. Итак, оптимальная точка, определяемая соотношением (4.88), является точкой равновесия — уже невозможно выжать из товаров-ресурсов больше, чем затрачено на их покупку.

Предположим, объем добычи щебня Y (т/ч) зависит от количества вложенного труда (чел. $\times$ ч) так: $Y = 6\sqrt{x}$. Цена щебня $v = 40$ руб./т, зарплата работника $p = 30$ руб./ч. Кроме зарплат другие издержки не учитываются. Требуется найти оптимальное количество x вложенного труда.

Прибыль найдем по формуле $W = vY - px = 240\sqrt{x} - 30x$. Для максимизации прибыли найдем ее производную и приравняем нулю. Получим $W' = 240/(2\sqrt{x}) - 30 = 0$, откуда оптимальное решение $x^* = 60$.

Для поиска решения можно использовать и соотношение (4.88). В данном случае предельное соотношение будет выглядеть так: $40 \cdot 6/(2\sqrt{x}) = 30$, что опять же дает ответ: $x^* = 60$.

Для наглядной интерпретации можно на графике производственной функции $Y = 6\sqrt{x}$ найти точку A , в которой касательная будет иметь угол наклона с тангенсом $p/v = 30/40 = 3/4$.

§ 18. Неопределенный интеграл

Ранее мы рассматривали такую задачу: дана функция $F(x)$; требуется найти ее производную, т.е. функцию $f(x) = F'(x)$.

Теперь мы будем рассматривать обратную задачу: дана функция $f(x)$, требуется найти такую функцию $F(x)$, производная которой равна $f(x)$, т.е. $F'(x) = f(x)$.

ОПРЕДЕЛЕНИЕ. $F(x)$ называется *первообразной* для функции $f(x)$ на отрезке $[a, b]$, если во всех точках этого отрезка выполняется равенство: $F'(x) = f(x)$.

Найдем первообразную, например, для функции $f(x) = x^2$. Из определения первообразной следует, что функция $F(x) = \frac{x^3}{3}$ яв-

ляется первообразной, так как $\left(\frac{x^3}{3}\right)' = x^2$. Легко видеть, что если

для данной функции $f(x)$ существует первообразная, то эта первообразная не является единственной. В качестве первообразных можно было взять также следующие функции:

$$F(x) = \frac{x^3}{3} + 1; \quad F(x) = \frac{x^3}{3} - 7; \quad F(x) = \frac{x^3}{3} + C, \quad C \text{ — произволь-}$$

ная постоянная, так как $\left(\frac{x^3}{3} + C\right)' = x^2$. С другой стороны, можно

доказать, что функциями вида $\frac{x^3}{3} + C$ исчерпываются все первообразные для функции x^2 . Это вытекает из следующей теоремы.

□ **ТЕОРЕМА 4.35.** Если $F_1(x)$ и $F_2(x)$ — две первообразные от функции $f(x)$ на отрезке $[a, b]$, то разность между ними равна постоянному числу. ■

ОПРЕДЕЛЕНИЕ. Если функция $F(x)$ является некоторой первообразной для $f(x)$, то выражение $F(x) + C$ называется **неопределенным интегралом** от функции $f(x)$ и обозначается $\int f(x)dx$.

Таким образом, $\int f(x)dx = F(x) + C$, если $F'(x) = f(x)$. При этом функция $f(x)$ называется *подынтегральной функцией*, $f(x)dx$ — *подынтегральным выражением*.

Неопределенный интеграл представляет собой семейство функций $y = F(x) + C$ (семейство кривых, сдвинутых вверх или вниз).

Для всякой ли функции $f(x)$ существуют первообразные (и, следовательно, неопределенный интеграл)? Нет, не для всякой.

Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то для этой функции существует первообразная (а значит, и неопределенный интеграл).

Нахождение первообразной для данной функции $f(x)$ называется *интегрированием* функции $f(x)$.

Из определения неопределенного интеграла вытекают следующие утверждения:

1) производная от неопределенного интеграла равна подынтегральной функции, т.е. если $F'(x) = f(x)$, то и

$(\int f(x)dx)' = (F(x) + C)' = f(x)$, значит, дифференциал от неопределенного интеграла равен подынтегральному выражению

$$d(\int f(x)dx) = f(x)dx;$$

2) неопределенный интеграл от дифференциала некоторой функции равен этой функции плюс произвольная постоянная

$\int dF(x) = F(x) + C$ (легко проверить дифференцированием обеих частей).

Таблица интегралов

$$1. \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C, \quad \alpha \neq -1; \quad 2. \int \frac{dx}{x} = \ln|x| + C;$$

$$\begin{array}{ll}
3. \int \sin x dx = -\cos x + C; & 4. \int \cos x dx = \sin x + C; \\
5. \int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C; & 6. \int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C; \\
7. \int \operatorname{tg} x dx = -\ln |\cos x| + C; & 8. \int \operatorname{ctg} x dx = \ln |\sin x| + C; \\
9. \int e^x dx = e^x + C; & 10. \int a^x dx = \frac{a^x}{\ln a} + C; \\
11. \int \frac{dx}{1+x^2} = \operatorname{arctg} x + C; & 11'. \int \frac{dx}{a^2+x^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C; \\
12. \int \frac{dx}{a^2-x^2} = \frac{1}{2a} \ln \left| \frac{a+x}{a-x} \right| + C; & 13. \int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C; \\
13'. \int \frac{dx}{\sqrt{a^2-x^2}} = \arcsin \frac{x}{a} + C; & 14. \int \frac{dx}{\sqrt{x^2 \pm a}} = \\
& = \ln \left| x + \sqrt{x^2 \pm a} \right| + C.
\end{array}$$

§ 19. Некоторые свойства неопределенного интеграла

19.1. Основные правила вычисления неопределенного интеграла

□ ТЕОРЕМА 4.36. *Неопределенный интеграл от алгебраической суммы двух или нескольких функций равен алгебраической сумме их интегралов:*

$$\int [f_1(x) + f_2(x)]dx = \int f_1(x)dx + \int f_2(x)dx. \blacksquare \quad (4.89)$$

□ ТЕОРЕМА 4.36. *Постоянный множитель можно выносить за знак интеграла, т.е. если $a = \text{const}$, то*

$$\int af(x)dx = a \int f(x)dx. \blacksquare$$

Если $\int f(x)dx = F(x) + C$, справедливы следующие правила:

$$1. \int f(ax)dx = \frac{1}{a}F(ax) + C. \quad (4.90)$$

$$2. \int f(x+b)dx = F(x+b) + C. \quad (4.91)$$

$$3. \int f(ax+b)dx = \frac{1}{a}F(ax+b) + C. \quad (4.92)$$

□ В качестве примера докажем правило 1.

Действительно, дифференцируя левую и правую части (4.90), получим

$$\left(\int f(ax)dx \right)' = f(ax);$$

$$\left(\frac{1}{a}F(ax) \right)' = \frac{1}{a}(F(ax))' = \frac{1}{a}F'(ax)a = F'(ax) = f(ax). \blacksquare$$

▷ П Р И М Е Р 4.34. Найти интеграл $\int (2x^3 - 3\sin x + 5\sqrt{x})dx$

Р е ш е н и е.

$$\begin{aligned}\int (2x^3 - 3\sin x + 5\sqrt{x})dx &= \int 2x^3 dx - \int 3\sin x dx + \int 5\sqrt{x} dx = \\ &= 2\int x^3 dx - 3\int \sin x dx + 5\int \sqrt{x} dx = 2\frac{x^{3+1}}{3+1} - 3(-\cos x) + \\ &+ 5\frac{x^{\frac{1}{2}+1}}{\frac{1}{2}+1} + C = \frac{1}{2}x^4 + 3\cos x + \frac{10}{3}x\sqrt{x} + C. \blacktriangleright\end{aligned}$$

19.2. Интегрирование методом замены переменной

Пусть требуется найти интеграл $\int f(x)dx$, причем непосредственно подобрать первообразную для $f(x)$ мы не можем. Сделаем замену переменной в подынтегральном выражении, положив

$$x = \varphi(t), \quad (4.93)$$

где $\varphi(t)$ — непрерывная функция с непрерывной производной, имеющая обратную функцию. Тогда $dx = \varphi'(t) dt$. Докажем, что

$$\int f(x)dx = \int f[\varphi(t)]\varphi'(t)dt. \quad (4.94)$$

Найдем производную от левой части равенства (4.94):

$$\left(\int f(x)dx\right)' = f(x).$$

Правую часть (4.94) будем дифференцировать по x как сложную функцию, где t — промежуточный аргумент. Зависимость

t от x выражается равенством (4.93), при этом $\frac{dx}{dt} = \varphi'(t)$ и по пра-

вилу дифференцирования обратной функции $\frac{dt}{dx} = \frac{1}{\varphi'(t)}$. Итак,

$$\begin{aligned} \left(\int f[\varphi(t)]\varphi'(t)dt \right)'_x &= \left(\int f[\varphi(t)]\varphi'(t)dt \right)'_t \frac{dt}{dx} = f[\varphi(t)]\varphi'(t) \frac{1}{\varphi'(t)} = \\ &= f[\varphi(t)] = f(x). \end{aligned}$$

Значит производные по x от правой и левой частей равенства (4.94) равны.

19.3. Интегрирование по частям

Пусть u и v — две дифференцируемые функции от x . Тогда $d(uv) = u dv + v du, \Rightarrow uv = \int u dv + \int v du, \Rightarrow \int u dv = uv - \int v du$ — (4.95) — формула интегрирования по частям.

▷ ПРИМЕР 4.35. Интегрированием по частям вычислить интегралы: а) $\int x^2 e^x dx$; б) $\int \sqrt{a^2 - x^2} dx$.

Р е ш е н и е.

$$\text{а) } \int x^2 e^x dx = \left. \begin{array}{l} u = x^2 \\ dv = e^x dx \\ du = 2x dx \\ v = e^x \end{array} \right| = x^2 e^x - 2 \int x e^x dx = \left. \begin{array}{l} u_1 = x \\ dv_1 = e^x dx \\ du_1 = dx \\ v_1 = e^x \end{array} \right| =$$

$$= x^2 e^x - 2x e^x + 2 \int e^x dx = x^2 e^x - 2(x e^x - e^x) + C = e^x (x^2 - 2x + 2) + C;$$

$$\text{б) } \int \sqrt{a^2 - x^2} dx = \int \frac{a^2 - x^2}{\sqrt{a^2 - x^2}} dx = a^2 \int \frac{dx}{\sqrt{a^2 - x^2}} - \int \frac{x^2}{\sqrt{a^2 - x^2}} dx =$$

$$= a^2 \arcsin \frac{x}{a} - \int \frac{x^2}{\sqrt{a^2 - x^2}} dx.$$

Найдем

$$\int \frac{x^2}{\sqrt{a^2 - x^2}} dx = \int x \frac{xdx}{\sqrt{a^2 - x^2}} = \left| \begin{array}{l} u = x \\ dv = \frac{xdx}{\sqrt{a^2 - x^2}} \\ du = dx \\ v = -\sqrt{a^2 - x^2} \end{array} \right| = -x\sqrt{a^2 - x^2} +$$

$$+ \int \sqrt{a^2 - x^2} dx,$$

тогда

$$\int \sqrt{a^2 - x^2} dx = a^2 \arcsin \frac{x}{a} + x\sqrt{a^2 - x^2} - \int \sqrt{a^2 - x^2} dx.$$

Откуда

$$\int \sqrt{a^2 - x^2} dx = \frac{a^2}{2} \arcsin \frac{x}{a} + \frac{x}{2} \sqrt{a^2 - x^2} + C. \blacktriangleright$$

19.4. Метод неопределенных коэффициентов

Пусть $P(x)$ и $Q(x)$ — целые многочлены, причем степень $P(x)$ меньше степени $Q(x)$, и многочлен $Q(x)$ имеет только вещественные корни $a, \dots, l$, тогда дробь $\frac{P(x)}{Q(x)}$ можно представить в следующем виде:

$$\frac{P(x)}{Q(x)} = \frac{A_1}{(x-a)} + \frac{A_2}{(x-a)^2} + \dots + \frac{A_\alpha}{(x-a)^\alpha} + \dots + \frac{L_1}{(x-l)} +$$

$$+ \frac{L_2}{(x-l)^2} + \dots + \frac{L_\lambda}{(x-l)^\lambda}.$$

Неопределенные коэффициенты $A_1, \dots, A_\alpha, \dots, L_1, \dots, L_\lambda$ можно найти, приравнявая коэффициенты при одинаковых степенях переменной x в числителях слева и справа.

▷ ПРИМЕР 4.36. Вычислить методом неопределенных коэф-

фициентов $\int \frac{xdx}{(x-1)(x+1)^2} = I.$

Р е ш е н и е. Разложим подынтегральное выражение следующим образом:

$$\frac{x}{(x-1)(x+1)^2} = \frac{A}{x-1} + \frac{B_1}{x+1} + \frac{B_2}{(x+1)^2}.$$

Откуда

$$x \equiv A(x+1)^2 + B_1(x-1)(x+1) + B_2(x-1); \quad (4.96)$$

$$x \equiv (A + B_1)x^2 + (2A + B_2)x + (A - B_1 - B_2).$$

Тогда

$$0 \cdot x^2 + 1 \cdot x + 0 = (A + B_1)x^2 + (2A + B_2)x + (A - B_1 - B_2).$$

Приравняв коэффициенты при одинаковых степенях x , получим

$$0 = A + B_1; \quad 1 = 2A + B_2; \quad 0 = A - B_1 - B_2.$$

$$\text{Отсюда: } A = \frac{1}{4}; \quad B_1 = -\frac{1}{4}; \quad B_2 = \frac{1}{2}. \text{ Следовательно,}$$

$$\begin{aligned} I &= \frac{1}{4} \int \frac{dx}{x-1} - \frac{1}{4} \int \frac{dx}{x+1} + \frac{1}{2} \int \frac{dx}{(x+1)^2} = \frac{1}{4} \ln|x-1| - \frac{1}{4} \ln|x+1| + \\ &+ \frac{-1}{2(x+1)} + C = \frac{-1}{2(x+1)} + \frac{1}{4} \ln \left| \frac{x-1}{x+1} \right| + C. \end{aligned}$$

Этот пример можно решить по-другому. Так как равенство (4.96) — тождество, то, придавая переменной x последовательно значения 0, -1 и 1 , получим

$$\begin{cases} 0 = A - B_1 - B_2; \\ -1 = -2B_2; \\ 1 = 4A. \end{cases}$$

$$\text{Отсюда найдем } A = \frac{1}{4}; \quad B_1 = -\frac{1}{4}; \quad B_2 = \frac{1}{2}. \blacktriangleright$$

§20. Определенный интеграл

20.1. Определение определенного интеграла

Мощным средством исследования в математике, физике, экономике, статистике, социологии и других науках является определенный интеграл — одно из основных понятий математического анализа.

Вычисление площадей, ограниченных кривыми, сводится к вычислению определенного интеграла.

Пусть на отрезке $[a, b]$ задана непрерывная функция $y = f(x)$ (рис. 4.42). Обозначим через m и M ее наименьшее и наибольшее значения на этом отрезке. Разобьем отрезок $[a, b]$ на n частей точками деления $a = x_0$, $x_1, x_2, \dots, x_{n-1}, x_n = b$, при-

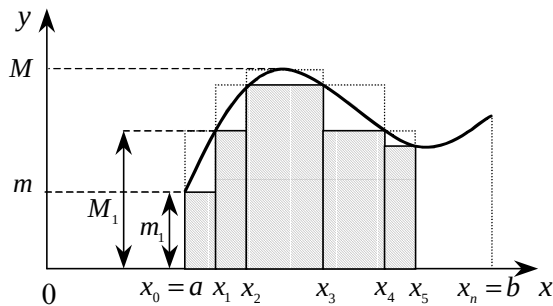


Рис. 4.42

чем $x_0 < x_1 < x_2 < \dots < x_n$, и положим $x_1 - x_0 = \Delta x_1$, $x_2 - x_1 = \Delta x_2$, ..., $x_n - x_{n-1} = \Delta x_n$. Обозначим далее наименьшее и наибольшее значения функции $f(x)$ на отрезке $[x_0, x_1]$ через m_1 и M_1 , на отрезке $[x_1, x_2]$ через m_2 и M_2 , ..., на отрезке $[x_{n-1}, x_n]$ через m_n и M_n . Составим суммы

$$\underline{S}_n = m_1 \Delta x_1 + m_2 \Delta x_2 + \dots + m_n \Delta x_n = \sum_{i=1}^n m_i \Delta x_i;$$

$$\overline{S}_n = M_1 \Delta x_1 + M_2 \Delta x_2 + \dots + M_n \Delta x_n = \sum_{i=1}^n M_i \Delta x_i.$$

Сумму $\underline{S}_n$ называют *нижней интегральной суммой*, а сумму $\overline{S}_n$ — *верхней интегральной суммой*. При $f(x) \geq 0$ площадь заштрихованной фигуры содержится между числами $\underline{S}_n$ и $\overline{S}_n$.

В каждом из отрезков $[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$ возьмем по произвольной точке, которые обозначим $\xi_1, \xi_2, \dots, \xi_n$; $x_0 < \xi_1 < x_1$, $x_1 < \xi_2 < x_2$, $x_{n-1} < \xi_n < x_n$ (рис. 4.43).

В каждой из этих точек вычислим значение функции $f(\xi_1), f(\xi_2), \dots, f(\xi_n)$. Составим сумму

$$S_n = f(\xi_1)\Delta x_1 + \dots + f(\xi_n)\Delta x_n = \sum_{i=1}^n f(\xi_i)\Delta x_i.$$

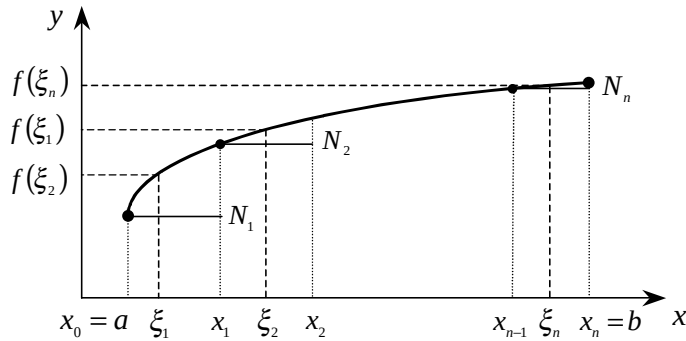


Рис. 4.43

Эта сумма называется *интегральной суммой* для функции $f(x)$ на отрезке $[a, b]$.

Рассмотрим некоторую последовательность разбиений, у которых $\max \Delta x_i \rightarrow 0$ при $n \rightarrow +\infty$. При каждом разбиении выбираем произвольно точки ξ_i^* , получим последовательность интегральных сумм. Предположим, что эта последовательность интегральных сумм S_n^* стремится к некоторому пределу

$$\lim_{\max \Delta x_i \rightarrow 0} S_n^* = \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f(\xi_i^*) \Delta x_i = S.$$

ОПРЕДЕЛЕНИЕ. Если при любых разбиениях отрезка $[a, b]$ таких, что при $n \rightarrow +\infty$ $\max \Delta x_i \rightarrow 0$, и при любом выборе точек ξ_i на

отрезках $[x_{i-1}, x_i]$ интегральная сумма $S_n = \sum_{i=1}^n f(\xi_i)\Delta x_i$ стремится к

одному и тому же пределу S , то этот предел называется **определенным интегралом** от функции $f(x)$ на отрезке $[a, b]$ и обозначается

$$\int_a^b f(x)dx.$$

При $f(x) \geq 0$ значение интеграла $\int_a^b f(x)dx$ равно площади заштрихованной фигуры (криволинейной трапеции, рис. 4.42). Таким образом, по определению

$$\lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i = \int_a^b f(x)dx. \quad (4.97)$$

Число a называется *нижним пределом* интеграла, b — *верхним пределом*. Отрезок $[a, b]$ называется *отрезком интегрирования*, x — *переменной интегрирования*, $f(x)$ — *подынтегральной функцией*, а $f(x)dx$ — *подынтегральным выражением*.

ОПРЕДЕЛЕНИЕ. Если для функции $f(x)$ предел (4.97) существует, то функцию называют **интегрируемой на отрезке $[a, b]$** .

20.2. Вычисление определенного интеграла. Формула Ньютона-Лейбница

Если функция $f(x)$ интегрируема на отрезке $[a, b]$, то она интегрируема и на отрезке $[a, x]$, где $x \in [a, b]$. Рассмотрим функцию аргумента x

$$\Phi(x) = \int_a^x f(t)dt. \quad (4.98)$$

(Здесь переменная интегрирования обозначена буквой t , чтобы избежать путаницы с верхним переменным пределом x).

ОПРЕДЕЛЕНИЕ. Функция $\Phi(x)$ (4.98) называется **интегралом с переменным верхним пределом**.

Если $f(t)$ — неотрицательная функция, то величина $\Phi(x)$ численно равна площади криволинейной трапеции $aAXx$. Очевидно, что эта площадь изменяется в зависимости от x (см. рис. 4.44).

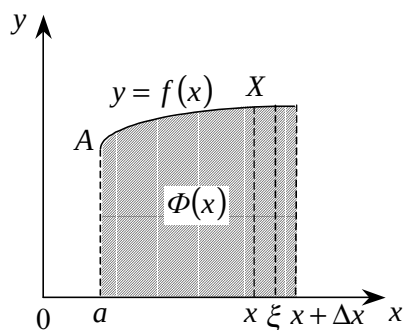


Рис. 4.44

□ **ТЕОРЕМА 4.37.** Если $f(x)$ — непрерывная функция и $\Phi(x) = \int_a^x f(t)dt$, то $\Phi'(x) = f(x)$. ■

Иными словами, производная от определенного интеграла по верхнему пределу равна подынтегральной функции, в которой вместо переменной интегрирования подставлено значение верхнего предела (при условии, что подынтегральная функция непрерывна) и функция $\Phi(x)$ является одной из первообразных функций $f(x)$.

□ **ТЕОРЕМА 4.38.** Если $F(x)$ есть какая-либо первообразная от непрерывной функции $f(x)$, то справедлива формула

$$\int_a^b f(x)dx = F(b) - F(a). \quad \blacksquare \quad (4.99)$$

Формула (4.99) называется *основной формулой интегрального исчисления*, или *формулой Ньютона-Лейбница*.

Отметим, что разность $F(b) - F(a)$ не зависит от выбора первообразной F , так как все первообразные отличаются на постоянную величину, которая при вычитании уничтожается.

Если ввести обозначение $F(b) - F(a) = F(x) \Big|_a^b$, то формулу (4.99) можно переписать так:

$$\int_a^b f(x)dx = F(x) \Big|_a^b = F(b) - F(a).$$

▷ ПРИМЕР 4.37. Вычислить определенный интеграл $\int_a^b x dx$.

Р е ш е н и е. По формуле Ньютона — Лейбница имеем:

$$\int_a^b x dx = \left. \frac{x^2}{2} \right|_a^b = \frac{b^2 - a^2}{2}. \blacktriangleright$$

При вычислении определенных интегралов методика для вычисления неопределенных интегралов остается верной с той лишь разницей, что теперь необходимо учитывать пределы интегрирования.

§ 21. Некоторые приложения интегрального исчисления

21.1. Классическая модель Вильсона управления запасами

Пусть $y(t)$ — величина запаса некоторого товара на складе в момент времени $t, t \geq 0$. Дефицит на складе не допускается, т.е. $y(t) \geq 0$ при $t \geq 0$. Товар пользуется равномерным спросом с интенсивностью μ , т.е. за время Δt со склада извлекается часть запаса величиной $\mu\Delta t$. С другой стороны, в моменты времени $t_0 = 0, t_1, t_2, \dots$ на склад приходят поставки величиной $Q_0, Q_1, Q_2, \dots$ соответственно. Таким образом, изменение запаса $y(t)$ товара на складе может быть изображено зубчатой ломаной (рис. 4.45), наклонные отрезки которой параллельны, при этом $\operatorname{tg} \alpha = \mu$.

Примем следующие обозначения. Если t_2 — момент очередной поставки, то величину запаса на складе в момент непосредственно перед поставкой будем обозначать $y(t_2^-)$, а величину запаса в момент поступления партии размера Q_2 будем обозначать $y(t_2^+)$ (см. рис. 4.45). Таким образом, $y(t_2^+) = y(t_2^-) + Q_2$.

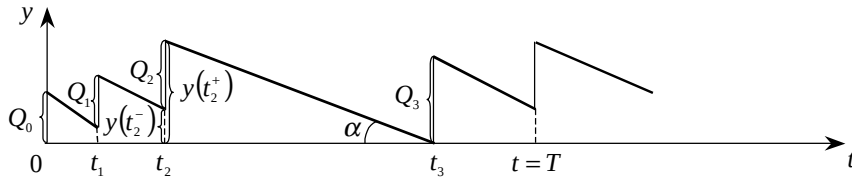


Рис. 4.45

Пусть s — плата за хранение единицы продукта в течение единицы времени, g — плата за доставку одной партии (не зависящая от размера поставки). Тогда средние издержки за время T

$$f_T(y) = f_T([y(t)], 0 \leq t \leq T) = \frac{1}{T} \left(s \int_0^T y(t) dt + gn(T) \right), \quad (4.100)$$

где $n(T)$ — количество поставок в интервале $[0, T]$.

Запись в левой части формулы (4.100) означает, что средние издержки зависят от значений функции $y(t)$ при $\forall 0 \leq t \leq T$. Условно функция обозначена как y . Область определения $f_T(y)$ — не множество чисел, а множество функций.

Для оптимизации системы управления запасами необходимо выбрать моменты поставки $t_1, t_2, t_3, \dots$ и величины поставок $Q_0, Q_1, Q_2, Q_3, \dots$ так, чтобы минимизировать значение $f_T(y)$ при фиксированном T . При этом период времени T будем называть *горизонтом планирования*.

Задание моментов приходов поставок и их величин полностью определяет функцию $y(t)$, $0 \leq t \leq T$, и наоборот: функция $y(t)$ полностью задает моменты прихода поставок и их величины. И то и другое будем называть *планом*.

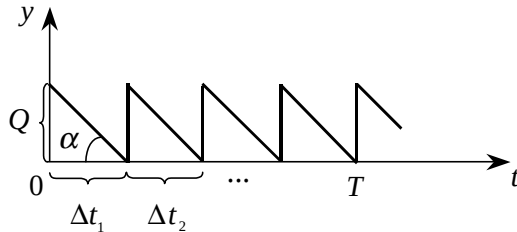


Рис. 4.46

ОПРЕДЕЛЕНИЕ. *Планом Вильсона*, соответствующим величине поставки Q , называется план, в котором все поставки равны, т.е. $Q_0 = Q_1 = Q_2 = \dots = Q$, и все временные интервалы равны одному и тому же значению $\Delta t_i = Q/\mu$, где $i = 1, 2, 3, \dots$ (рис. 4.46).

□ **ТЕОРЕМА 4.39.** Для любого периода времени T существует оптимальный план, на котором $f_T(y)$ достигает минимума. Этим планом является план Вильсона. ■

Пусть Q — размер поставки в плане Вильсона на отрезке $[0, T]$, соответствующий числу поставок $n(T)$, т.е. $\mu T = Qn(T)$. Тогда

$$f_T(y) = \frac{1}{T} \left(\frac{s\mu}{2} \frac{T^2}{n(T)} + gn(T) \right) = \frac{sQ}{2} + \frac{g\mu}{Q}. \quad (4.101)$$

Выражение в правой части формулы (4.101) достигает минимума при $Q = Q_0$, где

$$Q_0 = \sqrt{2\mu g / s}. \quad (4.102)$$

В том случае, если $\mu T/Q_0$ является целым числом, Q_0 определяет оптимальный размер поставки, и, следовательно, соответствующий план является планом Вильсона. Формула (4.2) — знаменитая “формула квадратного корня”, приводимая во всех учебниках по теории управления запасами. Минимальное значение выражения (4.1) равно $Q_0 = \sqrt{2\mu g/s}$.

21.2. Несобственные интегралы

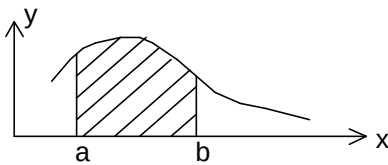


Рис. 4.47

Интегралы с бесконечными пределами. Пусть функция $f(x)$ определена и непрерывна при всех значениях x , таких, что $a \leq x < +\infty$. Рассмотрим интеграл: $I(b) = \int_a^b f(x)dx$.

Этот интеграл имеет смысл при любом $b > a$. При изменении b интеграл изменяется, он является непрерывной функцией b .

Рассмотрим вопрос о поведении этого интеграла при $b \rightarrow +\infty$ (см. рис. 4.47).

ОПРЕДЕЛЕНИЕ. Если существует конечный предел $\lim_{b \rightarrow +\infty} \int_a^b f(x)dx$, то этот предел называют **несобственным интегралом** от функции $f(x)$ на интервале $[a, +\infty)$ и обозначают так:

$$\int_a^{+\infty} f(x)dx.$$

Следовательно, по определению имеем:

$$\int_a^{+\infty} f(x)dx = \lim_{b \rightarrow +\infty} \int_a^b f(x)dx.$$

Говорят, что в этом случае несобственный интеграл $\int_a^{+\infty} f(x)dx$ существует или сходится. Если $\int_a^b f(x)dx$ при $b \rightarrow +\infty$ не имеет конечного предела, то говорят, что интеграл

$$\int_a^{+\infty} f(x)dx$$

не существует или расходится.

Геометрический смысл несобственного интеграла в случае, когда $f(x) \geq 0$: Если интеграл $\int_a^b f(x)dx$ выражает площадь области, ограниченной кривой $y = f(x)$, осью абсцисс и ординатами $x = a$, $x = b$, то будем считать, что несобственный интеграл $\int_a^{+\infty} f(x)dx$ выражает площадь неограниченной (бесконечной) области, заключенной между линиями $y = f(x)$, $x = a$ и осью абсцисс.

Аналогичным образом определяются несобственные интегралы и для других бесконечных интервалов:

$$\int_{-\infty}^b f(x)dx = \lim_{a \rightarrow -\infty} \int_a^b f(x)dx,$$

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^c f(x)dx + \int_c^{+\infty} f(x)dx.$$

Последнее равенство следует понимать так: если каждый из несобственных интегралов, стоящих справа, существует, то существует (сходится) и интеграл, стоящий слева.

▷ ПРИМЕР 4.38. Вычислить интеграл $\int_0^{+\infty} \frac{dx}{1+x^2}$.

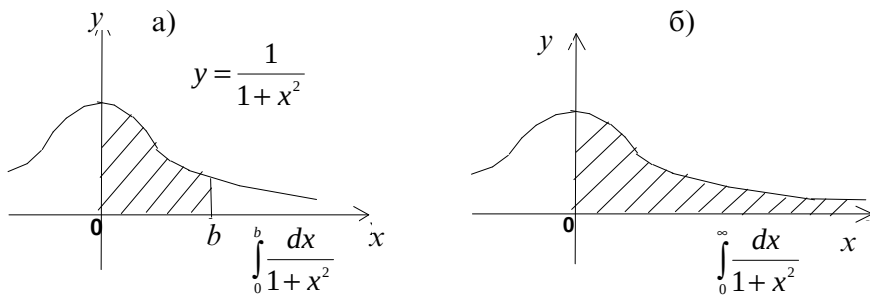


Рис. 4.48

Решение.

По определению несобственного интеграла и с учетом рис. 4.48, a, b :

$$\int_0^{+\infty} \frac{dx}{1+x^2} = \lim_{b \rightarrow +\infty} \int_0^b \frac{dx}{1+x^2} = \lim_{b \rightarrow +\infty} \arctg(x) \Big|_0^b = \lim_{b \rightarrow +\infty} \arctg(b) = \frac{\pi}{2}. \blacktriangleright$$

□ ТЕОРЕМА 4.40. Если для всех значений x ($x \geq a$) выполняется неравенство $0 \leq f(x) \leq \varphi(x)$ и если $\int_a^{+\infty} \varphi(x) dx$ сходится, то $\int_a^{+\infty} f(x) dx$ также сходится, при этом: $\int_a^{+\infty} f(x) dx \leq \int_a^{+\infty} \varphi(x) dx$. Если же выполняется неравенство $0 \leq \varphi(x) \leq f(x)$, причем интеграл $\int_a^{+\infty} \varphi(x) dx$ расходится, то расходится и интеграл $\int_a^{+\infty} f(x) dx$. ■

Для случая функции $f(x)$, меняющей знак в бесконечном интервале, имеет место следующая теорема.

□ ТЕОРЕМА 4.41. Если интеграл $\int_a^{+\infty} |f(x)| dx$ сходится, то сходится и интеграл $\int_a^{+\infty} f(x) dx$. В этом случае последний интеграл называется **абсолютно сходящимся**. ■

Интеграл от разрывной функции. Пусть функция $f(x)$ определена и непрерывна при $a \leq x < c$, а при $x = c$ функция либо не определена, либо терпит разрыв. В этом случае нельзя говорить об интеграле $\int_a^c f(x) dx$ как о пределе интегральных сумм, так как $f(x)$ не является непрерывной на отрезке $[a, c]$, и поэтому этот предел может и не существовать.

Интеграл $\int_a^c f(x) dx$ от функции $f(x)$, разрывной в точке c , определяется следующим образом: $\int_a^c f(x) dx = \lim_{b \rightarrow c-0} \int_a^b f(x) dx$. Если предел, стоящий справа, существует, то интеграл называют **несобственным сходящимся** интегралом, в противном случае интеграл называют **расходящимся**.

Если функция $f(x)$ имеет разрыв в левом конце отрезка $[a, c]$ (т.е. при $x = a$), то, по определению, $\int_a^c f(x) dx = \int_a^{x_0} f(x) dx + \int_{x_0}^c f(x) dx$.

Если функция $f(x)$ имеет разрыв в некоторой точке $x = x_0$ внутри отрезка $[a, c]$, то полагают $\int_a^c f(x)dx = \int_a^{x_0} f(x)dx + \int_{x_0}^c f(x)dx$, если оба несобственных интеграла, стоящих в правой части равенства, существуют.

Замечание. Если функция $f(x)$, определенная на $[a, b]$, имеет внутри этого отрезка конечное число точек разрыва $a_1, a_2, \dots, a_n$, то интеграл от функции $f(x)$ на отрезке $[a, b]$ определяется следующим образом:

$$\int_a^b f(x)dx = \int_a^{a_1} f(x)dx + \int_{a_1}^{a_2} f(x)dx + \dots + \int_{a_n}^b f(x)dx,$$

если каждый из несобственных интегралов в правой части равенства *сходится*.

Если же хотя бы один из этих интегралов *расходится*, то и $\int_a^b f(x)dx$ называют *расходящимся*.

21.3. Примеры вычисления площадей плоских фигур

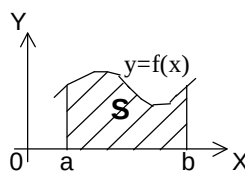


Рис. 4.49

Если непрерывная кривая задана в прямоугольных координатах уравнением $y = f(x)$ ($f(x) \geq 0$), то площадь криволинейной трапеции, ограниченной этой кривой, двумя вертикалями в точках $x = a$ и $x = b$ и отрезком оси абсцисс $a \leq x \leq b$ (см. рис. 4.49), определяется формулой:

$$S = \int_a^b f(x)dx$$

▷ ПРИМЕР 4.39. Вычислить площадь,

ограниченную параболой $y = \frac{x^2}{2}$, прямыми $x = 1$, $x = 3$ и осью абсцисс (см. рис. 4.50).

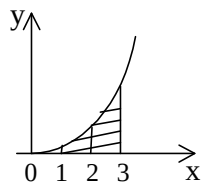


Рис. 4.50

Решение. Искомая площадь выражается интегралом:

$$S = \int_1^3 \frac{x^2}{2} dx = 4\frac{1}{3}. \blacktriangleright$$

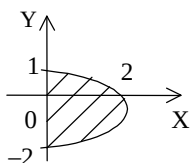


Рис. 4.51

▷ ПРИМЕР 4.40. Вычислить площадь, ограниченную кривой $x = 2 - y - y^2$ и осью ординат (см. рис. 4.51).

Решение. Здесь изменены роли осей координат, и потому искомая площадь выражается интегралом:

$$S = \int_{-2}^1 (2 - y - y^2) dy = 4\frac{1}{2},$$

где пределы интегрирования $y_1 = -2$ и $y_2 = 1$ найдены как ординаты точек пересечения данной кривой с осью ординат. ►

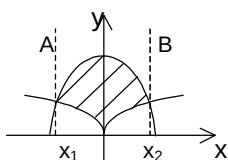


Рис. 4.52

▷ ПРИМЕР 4.41. Вычислить площадь S , заключенную между кривыми (см. рис. 4.52):

$$y = 2 - x^2, y^3 = x^2. \quad (4.103)$$

Решение. Решая совместно систему уравнений (4.103), находим пределы интегрирования $x_1 = -1$ и $x_2 = 1$. Используя геометрические соображения и определение интеграла, получим:

$$S = \int_{-1}^1 (2 - x^2 - x^{2/3}) dx = \left(2x - \frac{x^3}{3} - \frac{3}{5} x^{5/3} \right) \Big|_{-1}^1 = 2\frac{2}{15}. \quad \blacktriangleright$$

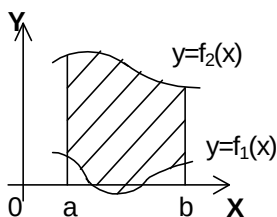


Рис. 4.53

В более общем случае, если площадь S ограничена двумя непрерывными кривыми $y = f_1(x)$ и $y = f_2(x)$ и двумя вертикалями $x = a$ и $x = b$, где $f_1(x) \leq f_2(x)$ при $a \leq x \leq b$ (см. рис. 4.53), будем иметь:

$$S = \int_a^b [f_2(x) - f_1(x)] dx. \quad (4.104)$$

21.4. Приближенное вычисление определенных интегралов

Известно, что не для всякой непрерывной функции ее первообразная выражается через элементарные функции. В этих случаях вычисление определенных интегралов по формуле Ньютона-Лейбница затруднительно, и на практике применяются различные методы *приближенного* вычисления определенных интегралов. Изложим несколько способов приближенного интегрирования, исходя из понятия определенного интеграла как предела суммы.

Формула прямоугольников. Пусть на отрезке $[a, b]$ задана непрерывная функция $y = f(x)$. Требуется вычислить определенный интеграл $\int_a^b f(x)dx$. Разделим отрезок $[a, b]$ точками $x_0 = a, x_1, x_2, \dots,$

$x_n = b$ на n равных частей длины Δx : $\Delta x = \frac{b-a}{n}$. Обозначим далее через $y_0, y_1, y_2, \dots, y_n$ значения функции $f(x)$ в точках $x_0, x_1, x_2, \dots, x_n$, т.е. $y_0 = f(x_0), y_1 = f(x_1), \dots, y_n = f(x_n)$.

Составим суммы: $y_0\Delta x + y_1\Delta x + \dots + y_{n-1}\Delta x, y_1\Delta x + y_2\Delta x + \dots + y_n\Delta x$.

Каждая из этих сумм является интегральной суммой для $f(x)$ на отрезке $[a, b]$ и потому приближенно выражает интеграл:

$$\int_a^b f(x)dx \approx \frac{b-a}{n}(y_0 + y_1 + y_2 + \dots + y_{n-1}), \text{ или} \quad (4.105)$$

$$\int_a^b f(x)dx \approx \frac{b-a}{n}(y_1 + y_2 + y_3 + \dots + y_n). \quad (4.106)$$

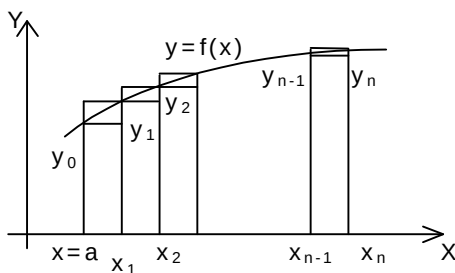


Рис. 4.54

Это и есть *формула прямоугольников*. Из рис. 4.54 видно, что формула (4.105) выражает площадь ступенчатой фигуры, составленной из прямоугольников под кривой, а формула (4.106) — площадь ступенчатой фигуры, состоящей из прямоугольников над кривой.

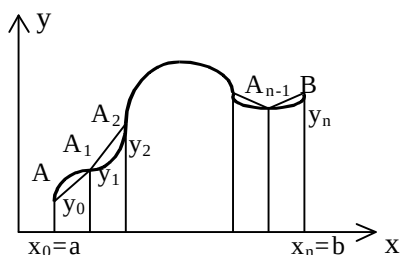


Рис. 4.55

молинейных трапеций, ограниченных сверху хордами $AA_1, A_1A_2, \dots,$

$A_{n-1}B$. Так как площадь первой из этих трапеций равна $\frac{y_0 + y_1}{2} \Delta x$,

площадь второй равна $\frac{y_1 + y_2}{2} \Delta x$ и т.д. (см. рис. 4.55), то

$$\int_a^b f(x) dx \approx \left(\frac{y_0 + y_1}{2} \Delta x + \frac{y_1 + y_2}{2} \Delta x + \dots + \frac{y_{n-1} + y_n}{2} \Delta x \right), \text{ или}$$

$$\int_a^b f(x) dx \approx \frac{b-a}{n} \left(\frac{y_0 + y_n}{2} + y_1 + y_2 + \dots + y_{n-1} \right). \quad (4.107)$$

Формула (4.107) называется *формулой трапеций*. Отметим, что число, стоящее в правой части формулы (4.107), есть среднее арифметическое чисел, стоящих в правых частях формул (4.105) и (4.106).

Число n выбирается произвольно. Чем больше будет это чис-

ло, тем меньше будет шаг $x = \frac{b-a}{n}$, значит, с большей точностью

сумма, написанная в правой части приближенного равенства (4.107), будет соответствовать значению интеграла.

Формула парабол (формула Симпсона). Разделим отрезок $[a, b]$ на четное число равных частей, т.е. $n = 2m$, где $m = 1, 2, \dots$. Площадь криволинейной трапеции, соответствующей первым двум отрезкам $[x_0, x_1]$ и $[x_1, x_2]$ и ограниченной заданной кривой $y = f(x)$, заменим площадью криволинейной трапеции, которая ограниче-

Формула трапеций. Мы получим более точное значение определенного интеграла, если данную кривую $y = f(x)$ заменим не ступенчатой линией, как это было в формуле прямоугольников, а *вписанной ломаной*. Тогда площадь криволинейной трапеции $aABb$ заменится суммой площадей пря-

на *параболой*, проходящей через три точки $M_0(x_0; y_0)$, $M_1(x_1; y_1)$, $M_2(x_2; y_2)$ и имеющей ось, параллельную оси Oy . Такую криволинейную трапецию будем называть *параболической трапецией*.

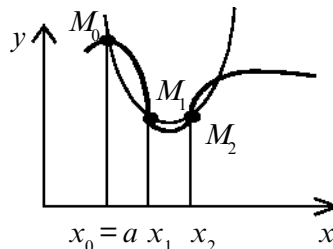


Рис. 4.56

Уравнение параболы с осью, параллельной оси Oy , имеет вид $y = Ax^2 + Bx + C$. Коэффициенты A , B , C однозначно определяются из условия, что парабола проходит через три заданные точки (рис. 4.56).

Аналогичные параболы строим и для других пар отрезков. Сумма площадей параболических трапеций и даст приближенное значение интеграла.

УТВЕРЖДЕНИЕ. Если криволинейная трапеция ограничена параболой $y = Ax^2 + Bx + C$, осью Ox и двумя ординатами, расстояние между которыми равно $2h$, то ее площадь равна:

$$S = \frac{h}{3}(y_0 + 4y_1 + y_2), \quad (4.108)$$

где y_0 и y_2 — крайние ординаты, а y_1 — ордината кривой в середине отрезка.

Теперь, пользуясь формулой (4.108), можем написать следующие приближенные равенства ($h = \Delta x$):

$$\begin{aligned} \int_{a=x_0}^{x_2} f(x)dx &\approx \frac{\Delta x}{3}(y_0 + 4y_1 + y_2), \\ \int_{x_2}^{x_4} f(x)dx &\approx \frac{\Delta x}{3}(y_2 + 4y_3 + y_4), \\ &\dots\dots\dots \\ \int_{x_{2m-2}}^{x_{2m}=b} f(x)dx &\approx \frac{\Delta x}{3}(y_{2m-2} + 4y_{2m-1} + y_{2m}). \end{aligned}$$

Складывая левые и правые части, получим слева искомый интеграл, справа — его приближенное значение:

$$\int_a^b f(x)dx \approx \frac{\Delta x}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + \dots + 2y_{2m-2} + 4y_{2m-1} + y_{2m}), \quad (4.109)$$

или

$$\int_a^b f(x)dx \approx \frac{\Delta x}{3}(y_0 + y_{2m} + 2(y_2 + y_4 + \dots + y_{2m-2}) + 4(y_1 + y_3 + \dots + y_{2m-1})).$$

Это и есть *формула Симпсона*. Здесь число точек деления $2m$ произвольно, но чем это число больше, тем точнее сумма в правой части (4.109) дает значение интеграла.

§ 22. Дифференциальные уравнения

22.1. Понятие о дифференциальном уравнении и его решении

ОПРЕДЕЛЕНИЕ. *Дифференциальным уравнением* называется уравнение, содержащее производные или дифференциалы неизвестной функции.

Дифференциальные уравнения называются *обыкновенными*, если входящие в них функции зависят от одного аргумента.

Порядком дифференциального уравнения называется порядок высшей производной, содержащейся в этом уравнении. Так, уравнение $y'' + 2xy - y^2 = 0$ — второго порядка; $y' - xy = x^2$ — первого порядка.

Функция, обращающая дифференциальное уравнение в тождество, называется *решением* этого уравнения.

Например, функция $y = 5x$ является решением уравнения $y'' - 2xy + 2y = 0$, так как вычислив производные этой функции $y' = 5$, $y'' = 0$ и подставив в данное уравнение их значения и y , получим тождество $-10x + 10x = 0$.

Дифференциальное уравнение первого порядка, его общее решение и начальные условия. Дифференциальное уравнение первого порядка имеет вид:

$$F(x, y, y') = 0, \quad (4.110)$$

или (если его можно разрешить относительно y')

$$y' = f(x, y). \quad (4.111)$$

Решение уравнения (4.110) или (4.111), содержащее произвольную постоянную C , т. е. имеющее вид $y = \varphi(x, C)$, называется *общим решением* этого уравнения. Если это решение получается в неявной форме $\Phi(x, y, C) = 0$, то соотношение $\Phi(x, y, C) = 0$ называется *общим интегралом* уравнения (4.110) или (4.111).

Если придать произвольной постоянной C некоторое фиксированное значение, то из общего решения (общего интеграла) получим частное решение (частный интеграл) этого уравнения.

Для уравнения (4.111) справедлива следующая теорема, называемая теоремой о существовании и единственности решения.

□ ТЕОРЕМА 4.42. (теорема Коши). Если в уравнении (4.111) функция $f(x, y)$ и ее частная производная $f_y(x, y)$ непрерывны в некоторой области D , содержащей точку (x_0, y_0) , то существует единственное решение этого уравнения $y = \varphi(x)$, удовлетворяющее условию: при $x = x_0$ $y = y_0$. ■

Условие, что при $x = x_0$ функция y должна равняться заданному числу y_0 , называется начальным условием. Начальное условие дает возможность выделить из общего решения уравнения частное решение. Для этого из уравнения $y_0 = \varphi(x_0, C)$ определяется конкретное значение $C = C_0$, и тогда искомое частное решение имеет вид: $y = \varphi(x, C_0) = \psi(x)$.

Задача нахождения решения уравнения (4.111), удовлетворяющего начальному условию, называется задачей Коши — из множества интегральных кривых выделяется та, которая проходит через заданную точку (x_0, y_0) области D . В ряде случаев, когда условия теоремы Коши не выполнены, через некоторые точки плоскости Oxy либо не проходит ни одна интегральная кривая, либо проходит более одной интегральной кривой.

▷ ПРИМЕР 4.42. Показать, что функция $y = \sqrt{x^2 + Cx}$ является общим решением уравнения $2xyy' = x^2 + y^2$. Найти частное решение, удовлетворяющее начальному условию $y(-1) = 2$.

Р е ш е н и е. Вычислим производную: $y' = \frac{2x + C}{2\sqrt{x^2 + Cx}}$. Подставив в заданное уравнение значения y и y' , получим тождество $2x^2 + Cx = 2x^2 + Cx$. Следовательно, функция $y = \sqrt{x^2 + Cx}$ является общим решением данного уравнения.

Чтобы найти частное решение, удовлетворяющее заданному начальному условию, надо найти соответствующее значение произвольной постоянной C . Для этого в общее решение под-

ставим начальное условие, т.е. $x = -1, y = 2$. Получим уравнение $2 = \sqrt{1-C}$, откуда $C = -3$. Следовательно, решение $y = \sqrt{x^2 - 3x}$ есть то частное решение, которое удовлетворяет заданному начальному условию. ►

Уравнения с разделяющимися переменными. Уравнение первого порядка $P(x, y)dx + Q(x, y)dy = 0$ называется уравнением с разделяющимися переменными, если функции $P(x, y)$ и $Q(x, y)$ разлагаются на множители, каждый из которых зависит только от одной переменной:

$$f_1(x)f_2(y)dx + \varphi_1(x)\varphi_2(y)dy = 0. \quad (4.112)$$

В уравнении (4.12) путем деления его членов на произведение $f_2(y) \cdot \varphi_1(x)$ переменные разделяются следующим образом:

$$\frac{f_1(x)}{\varphi_1(x)}dx + \frac{\varphi_2(y)}{f_2(y)}dy = 0, \quad (f_2(y) \neq 0; \varphi_1(x) \neq 0).$$

Общий интеграл находится почленным интегрированием:

$$\int \frac{f_1(x)}{\varphi_1(x)}dx + \int \frac{\varphi_2(y)}{f_2(y)}dy = C, \quad (C = \text{const}).$$

В общем случае, разделив на произведение $f_2(y)\varphi_1(x)$, можно потерять те решения исходного уравнения, которые обращают это произведение в нуль. Непосредственной подстановкой легко убедиться, что функция $y = b$, где b — корень уравнения $f_2(y) = 0$, есть решение исходного дифференциального уравнения. Аналогично, функция $x = a$, где a — корень уравнения $\varphi_1(x) = 0$, также является решением исходного уравнения.

► ПРИМЕР 4.43. Найти общий интеграл уравнения:

$$x y y' = 1 - x^2.$$

Решение. Заменим $y' = \frac{dy}{dx}$, умножим обе части уравнения

на dx :

$$xydy = (1 - x^2)dx.$$

Разделим обе части уравнения на x : $udy = \left(\frac{1}{x} - x\right)dx$.

Интегрируя, получим общий интеграл

$$\int ydy = \int \left(\frac{1}{x} - x\right)dx.$$

Тогда решение найдем из уравнения $\frac{y^2}{2} = \ln|x| - \frac{x^2}{2} + C$.

Предлагаем читателю сделать это самостоятельно. ►

Однородные дифференциальные уравнения первого порядка.

ОПРЕДЕЛЕНИЕ. Дифференциальное уравнение $y' = f(x, y)$ называется *однородным*, если функция $f(x, y)$ удовлетворяет условию

$$f(\lambda x, \lambda y) = f(x, y)$$

С помощью подстановки $y = tx$, где t — новая искомая функция от x , однородное уравнение приводится к уравнению с разделяющимися переменными.

► ПРИМЕР 4.44. Решить уравнение $y' = \frac{y}{x} \left(1 + \ln \frac{y}{x}\right)$.

Р е ш е н и е. Правая часть уравнения $f(x, y) = \frac{y}{x} \left(1 + \ln \frac{y}{x}\right)$ обладает свойством

$$f(\lambda x, \lambda y) = \frac{\lambda y}{\lambda x} \left(1 + \ln \frac{\lambda y}{\lambda x}\right) = \frac{y}{x} \left(1 + \ln \frac{y}{x}\right) = f(x, y).$$

Поэтому данное уравнение является однородным дифференциальным уравнением первого порядка. Применим подстановку $y = tx$. Отсюда $y' = t'x + t$. Исходное уравнение принимает вид:

$$t'x + t = t(1 + \ln t); \quad t'x = t \ln t.$$

Заменив $t' = \frac{dt}{dx}$ и разделив переменные, получим $\frac{dt}{t \ln t} = \frac{dx}{x}$.

Интегрируя полученное уравнение, будем иметь:

$$\int \frac{dt}{t \ln t} = \int \frac{dx}{x}; \int \frac{d(\ln t)}{\ln t} = \ln|x| + C; \ln|\ln t| = \ln|x| + \ln C,$$

что после потенцирования даст $\ln t = Cx$, или $t = e^{Cx}$. Так как $t = \frac{y}{x}$,

то $\frac{y}{x} = e^{Cx}$, откуда $y = xe^{Cx}$ — общее решение данного уравнения. ►

ОПРЕДЕЛЕНИЕ. Функция $f(x, y)$ называется **однородной порядка m** , если имеет место тождество:

$$f(\lambda x, \lambda y) = \lambda^m f(x, y). \quad (4.113)$$

ОПРЕДЕЛЕНИЕ. Дифференциальное уравнение первого порядка $M(x, y)dx + N(x, y)dy = 0$ называется **однородным дифференциальным уравнением**, если функции $M(x, y)$ и $N(x, y)$ являются однородными функциями одного и того же порядка.

▷ **ПРИМЕР 4.45.** Решить уравнение $(x - y)dx - (x + y)dy = 0$.

Решение. Коэффициенты при дифференциалах dx и dy , т. е. функции $M(x, y) = -(x + y)$ и $N(x, y) = x - y$, являются однородными функциями первого порядка:

$$M(\lambda x, \lambda y) = -(\lambda x + \lambda y) = -\lambda(x + y) = \lambda \cdot M(x, y),$$

$$N(\lambda x, \lambda y) = \lambda x - \lambda y = \lambda(x - y) = \lambda \cdot N(x, y).$$

Следовательно, данное уравнение — однородное. Положим $y = tx$, где t — некоторая функция переменной x . Так как $dy = t \cdot dx + x \cdot dt$, то данное уравнение примет вид: $(x - tx)(tdx + xdt) - (x + tx)dx = 0$. После упрощений получим: $x(1 - t)dt = (t^2 + 1)dx$. Произведем разделение переменных:

$$\frac{(1-t)dt}{1+t^2} = \frac{dx}{x}.$$

После интегрирования обеих частей уравнения получаем:

$$\int \frac{(1-t)dt}{1+t^2} = \int \frac{dt}{1+t^2} - \int \frac{tdt}{1+t^2} = \operatorname{arctg} t - \frac{1}{2} \ln|1+t^2|;$$

$$\int \frac{dx}{x} = \ln|x| + C.$$

Таким образом,

$$\operatorname{arctg} t - \frac{1}{2} \ln|1+t^2| = \ln|x| + C;$$

$$\operatorname{arctg} t - \ln \sqrt{1+t^2} - \ln|x| = C;$$

$$\operatorname{arctg} t - \ln \sqrt{1+t^2} \cdot x = C.$$

Так как $t = \frac{y}{x}$, то получаем $\operatorname{arctg} \frac{y}{x} - \ln \sqrt{1 + \frac{y^2}{x^2}} \cdot x = C$, или

$\operatorname{arctg} \frac{y}{x} - \ln \sqrt{x^2 + y^2} = C$ — общее решение заданного уравнения. ►

22.2. Линейные уравнения первого порядка

ОПРЕДЕЛЕНИЕ. Уравнение вида $y' + p(x)y = f(x)$, где $p(x)$ и $f(x)$ — заданные непрерывные функции, называется **линейным дифференциальным уравнением первого порядка**.

Для решения линейного уравнения применим подстановку $y = u \cdot v$, причем функцию $u = u(x)$ будем считать новой неизвестной функцией, а функцию $v = v(x)$ выбираем произвольно. Эта подстановка дает

$$u'v + uv' + p(x)uv = f(x) \text{ или } u'v + [v' + p(x)v]u = f(x).$$

Выбираем функцию $v(x)$ так, чтобы $v' + p(x)v = 0$. Тогда предыдущее уравнение сводится к следующим двум уравнениям:

$$1) v' + p(x)v = 0 \text{ и } 2) u'v = f(x),$$

каждое из которых есть дифференциальное уравнение с разделяющимися переменными.

▷ ПРИМЕР 4.46. Решить уравнение $y' - 2xy = 2xe^{x^2}$.

Решение. Убедившись, что данное уравнение линейное, полагаем $y = u \cdot v$, тогда $y' = u'v + uv'$ и данное уравнение преобразуется к виду:

$$u'v + uv' - 2xy = 2xe^{x^2},$$

или

$$u'v + u(v' - 2xv) = 2xe^{x^2}.$$

Так как функция y представлена в виде произведения двух вспомогательных функций u и v , то одну из них можно выбрать произвольно. Выберем в качестве v какой-либо частный интеграл уравнения $v' - 2xv = 0$. Тогда для отыскания функции u получим уравнение $u'v = 2xe^{x^2}$.

Решая первое из этих уравнений, найдем v ; разделяя переменные и интегрируя, найдем его простейший интеграл:

$$\frac{dv}{v} = 2xdx; \ln|v| = x^2 \text{ (положим } C = 0\text{)}.$$

Потенцируя, находим $v = e^{x^2}$.

Подставляя $v = e^{x^2}$ во второе уравнение, получим

$$u'e^{x^2} = 2xe^{x^2}; u' = 2x.$$

Находим общее решение этого уравнения: $u = x^2 + C$. Зная u и v , находим искомую функцию y : $y = u \cdot v = (x^2 + C)e^{x^2}$. ►

22.3. Уравнение Бернулли

ОПРЕДЕЛЕНИЕ. Уравнение вида

$$y' + p(x)y = f(x)y^n, \tag{4.114}$$

где $p(x)$ и $f(x)$ — заданные непрерывные функции, а $n \neq 0$, $n \neq 1$, называется *уравнением Бернулли*.

Отметим, что при $n = 0$ это уравнение становится линейным уравнением, а при $n = 1$ — уравнением с разделяющимися переменными.

При помощи подстановки $z = y^{1-n}$ уравнение Бернулли приводится к линейному уравнению относительно новой функции z .

Уравнение Бернулли можно решить с помощью подстановки $y = u(x) \cdot v(x)$, не сводя его предварительно к линейному.

▷ ПРИМЕР 4.47. Найти частное решение уравнения

$$y' + \frac{y}{x} = -xy^2, \text{ удовлетворяющее начальному условию } y(1) = 1.$$

Р е ш е н и е. Данное уравнение является уравнением Бернулли. Положим $y = u \cdot v$, тогда $y' = u'v + uv'$, и уравнение примет вид:

$$u'v + uv' + \frac{uv}{x} = -xu^2v^2, \text{ или } u'v + u\left(v' + \frac{v}{x}\right) = -xu^2v^2.$$

Отсюда для нахождения u и v имеем два уравнения:

$$1) v' + \frac{v}{x} = 0 \text{ и } 2) u' = -xu^2v^2.$$

Из первого уравнения находим функцию v :

$$\frac{dv}{dx} + \frac{v}{x} = 0; \frac{dv}{v} = -\frac{dx}{x}; \int \frac{dv}{v} = -\int \frac{dx}{x}; \ln|v| = -\ln|x|.$$

Здесь для простоты постоянная $C = 0$. Отсюда $v = \frac{1}{x}$.

Подставляем $v = \frac{1}{x}$ во второе уравнение и решая его, находим функцию u :

$$u' = -xu^2 \cdot \left(\frac{1}{x}\right)^2; u' = -\frac{u^2}{x}; \frac{du}{dx} = -\frac{u^2}{x}; -\frac{du}{u^2} = \frac{dx}{x}; -\int \frac{du}{u^2} = \int \frac{dx}{x} + C;$$

$$\frac{1}{u} = \ln|x| + C; u = \frac{1}{\ln|x| + C}.$$

Тогда $y = u \cdot v = \frac{1}{\ln|x| + C} \cdot \frac{1}{x} = \frac{1}{x(\ln|x| + C)}$ — общее решение данного уравнения.

Подставляя сюда заданные значения переменных $x = 1, y = 1$, находим значение произвольной постоянной $C = 0$.

При $C = 0$ из общего решения получаем $y = \frac{1}{x^2}$. Это и будет частное решение данного уравнения, удовлетворяющее заданному начальному условию $y(1) = 1$. ►

22.4. Дифференциальные уравнения второго порядка

Дифференциальное уравнение второго порядка, его общее решение и начальные условия. Дифференциальное уравнение второго порядка имеет вид

$$F(x, y, y', y'') = 0. \quad (4.115)$$

ОПРЕДЕЛЕНИЕ. Функция $y = \varphi(x, C_1, C_2)$, удовлетворяющая уравнению (4.15) при любых значениях произвольных постоянных C_1 и C_2 , называется его **общим решением**.

ОПРЕДЕЛЕНИЕ. Решение уравнения (4.115), получающееся из общего решения при конкретных значениях произвольных постоянных C_1 и C_2 , называется **частным решением** этого уравнения.

Так как в функцию $y = \varphi(x, C_1, C_2)$ входят две произвольные постоянные C_1 и C_2 , то для выделения из общего решения уравнения (4.115) некоторого частного решения необходимо иметь два начальных условия: если $x = x_0$, то $y = y_0, y' = y'_0, y'' = y''_0$, т. е. $y(x_0) = y_0, y'(x_0) = y'_0$.

$$\text{Тогда } \begin{cases} y_0 = \varphi(x_0, C_1, C_2), \\ y'_0 = \varphi'(x_0, C_1, C_2). \end{cases}$$

Из этой системы можно определить постоянные C_1 и C_2 и тем самым найти частное решение уравнения (4.115).

Уравнения второго порядка, допускающие понижение порядка. В следующих частных случаях уравнения второго порядка сводятся к дифференциальным уравнениям первого порядка.

СЛУЧАЙ 1. Уравнение имеет вид:

$$y'' = f(x), \quad (4.116)$$

где $f(x)$ — непрерывная на интервале (a, b) функция.

Так как $y'' = (y')' = \frac{dy'}{dx}$, то данное уравнение можно записать так:

$$\frac{dy'}{dx} = f(x) \text{ или } dy' = f(x)dx.$$

Интегрируя последнее уравнение, получим:

$$y' = \int f(x)dx + C_1.$$

Интегрируя еще один раз, получим общее решение уравнения (4.116):

$$y = \int (\int f(x) dx) dx + C_1 x + C_2.$$

▷ ПРИМЕР 4.48. Найти общее решение уравнения $y'' = 8e^{2x} + 12x$.

Р е ш е н и е. Так как $y'' = \frac{dy'}{dx}$, то данное уравнение можно записать так:

$$\frac{dy'}{dx} = 8e^{2x} + 12x, \text{ или } dy' = (8e^{2x} + 12x) dx.$$

Интегрируя, получим:

$$y' = 4 \int e^{2x} d(2x) + 12 \int x dx = 4e^{2x} + 6x^2 + C_1.$$

Отсюда $dy = (4e^{2x} + 6x^2 + C_1) dx$ и, значит,

$$y = 2 \int e^{2x} d(2x) + 6 \int x^2 dx + C_1 \int dx = 2e^{2x} + 2x^3 + C_1 x + C_2.$$

Итак, $y = 2e^{2x} + 2x^3 + C_1 x + C_2$ — общее решение заданного уравнения. ►

СЛУЧАЙ 2. Пусть уравнение не содержит y , то есть имеет вид:
 $F(x, y', y'') = 0. \quad (4.117)$

Положим $y' = z$, где z — некоторая функция аргумента x . Тогда $y'' = z'$ и уравнение (4.117) становится уравнением первого порядка:

$$F(x, z, z') = 0.$$

▷ ПРИМЕР 4.49. Найти общее решение уравнения $x y'' = y'$.

Р е ш е н и е. Данное уравнение не содержит явным образом y .

Положим $y' = z$.

Тогда $y'' = z'$. Имеем:

$$xz' = z; \quad x \frac{dz}{dx} = z; \quad \frac{dz}{z} = \frac{dx}{x}.$$

Интегрируя последнее уравнение, получим $\ln|z| = \ln|x| + \ln C_1$

или, потенцируя, $z = C_1 x$. Так как $z = y' = \frac{dy}{dx}$, то $dy = C_1 x dx$. Интегрируя еще раз, получим общее решение заданного уравнения:

$$y = C_1 \cdot \frac{x^2}{2} + C_2. \blacktriangleright$$

СЛУЧАЙ 3. Пусть уравнение не содержит x , то есть имеет вид:

$$F(y, y', y'') = 0. \quad (4.118)$$

Тогда в качестве неизвестной функции опять берется y' , но за аргумент вместо x принимаем y . Пусть $y' = p$. Применяя правило дифференцирования сложной функции, будем иметь

$$y'' = \frac{dy'}{dx} = \frac{dp}{dx} = \frac{dp}{dy} \cdot \frac{dy}{dx} = p \frac{dp}{dy}.$$

С учетом этого уравнение (4.18) примет вид $F\left(y, p, p \frac{dp}{dy}\right) = 0$ — это дифференциальное уравнение первого порядка относительно функции p .

▷ ПРИМЕР 4.50. Найти общее решение уравнения $2yy'' = 1 + (y')^2$.

Р е ш е н и е. Пусть $y' = p$, тогда $y'' = p \frac{dp}{dy}$ и заданное уравнение примет вид $2yp \frac{dp}{dy} = 1 + p^2$. Разделяя в нем переменные, полу-

чим $\frac{2pdp}{1+p^2} = \frac{dy}{y}$.

Интегрируем это уравнение:

$$\int \frac{2pdp}{1+p^2} = \int \frac{dy}{y}; \quad \int \frac{d(1+p^2)}{1+p^2} = \ln|y| + \ln C_1; \quad \ln(1+p^2) = \ln C_1 |y|;$$

$$1+p^2 = C_1 y.$$

Откуда $p^2 = C_1 y - 1$, или $(y')^2 = C_1 y - 1$, и $y' = \pm \sqrt{C_1 y - 1}$.

Далее, так как $y' = \frac{dy}{dx}$, то $\frac{dy}{dx} = \pm \sqrt{C_1 y - 1}$; $\pm \frac{dy}{\sqrt{C_1 y - 1}} = dx$.

Интегрируя последнее уравнение, получим

$$\pm \frac{1}{C_1} \int (C_1 y - 1)^{\pm 1} d(C_1 y - 1) = x + C_2.$$

Возводя обе части этого равенства в квадрат и выражая потом y , получим $y = \frac{C_1}{4}(x + C_2)^2 + \frac{1}{C_1}$ — общее решение заданного уравнения. ►

Линейные однородные уравнения второго порядка с постоянными коэффициентами.

ОПРЕДЕЛЕНИЕ. Уравнение вида

$$y'' + py' + qy = 0, \quad (4.119)$$

где p и q — постоянные действительные числа, называется **линейным однородным** дифференциальным уравнением второго порядка с постоянными коэффициентами; p и q — его коэффициенты.

Общее решение уравнения (4.119) находится с помощью характеристического уравнения

$$k^2 + pk + q = 0, \quad (4.120)$$

которое получается из уравнения (4.119), если, сохраняя в нем коэффициенты p и q , заменить функцию y единицей, а все ее производные k -го порядка соответствующими степенями k .

При этом справедливы правила:

1°. Если корни k_1 и k_2 характеристического уравнения (4.120) действительные и различные, то общее решение уравнения (4.119) выражается формулой:

$$y = C_1 e^{k_1 x} + C_2 e^{k_2 x}. \quad (4.121)$$

2°. Если корни k_1 и k_2 характеристического уравнения (4.120) действительные и равные ($k_1 = k_2$), то общее решение уравнения (4.119) выражается формулой:

$$y = C_1 e^{k_1 x} + C_2 x e^{k_1 x} = e^{k_1 x} (C_1 + C_2 x). \quad (4.122)$$

3°. Если корни k_1 и k_2 характеристического уравнения комплексные ($k_1 = \alpha + \beta i$, $k_2 = \alpha - \beta i$), то общее решение уравнения (4.119) есть

$$y = e^{\alpha x} (C_1 \cos \beta x + C_2 \sin \beta x). \quad (4.123)$$

▷ ПРИМЕР 4.51. Найти общее решение уравнений:

а) $y'' - y' - 6y = 0$; б) $y'' - 22y' + 121y = 0$; в) $y'' + 16y = 0$;

г) $y'' - 4y' + 20y = 0$.

Р е ш е н и е: а) заменяя в данном уравнении функцию y единицей, а ее производные соответствующими степенями k , напомним характеристическое уравнение: $k^2 - k - 6 = 0$.

Корни этого уравнения $k_1 = -2$, $k_2 = 3$ действительные и различные. Поэтому согласно правилу 1° искомое общее решение данного уравнения будет

$$y = C_1 e^{-2x} + C_2 e^{3x};$$

б) по указанному правилу составляем характеристическое уравнение: $k^2 - 22k + 121 = 0$. Решая это уравнение, получим $k_1 = k_2 = 11$. Согласно правилу 2° общим решением данного уравнения будет $y = C_1 e^{11x} + C_2 x e^{11x}$;

в) характеристическое уравнение $k^2 + 16 = 0$ имеет корни $k_1 = 4i$; $k_2 = -4i$. Согласно правилу 3° общее решение данного уравнения имеет вид: $y = e^{\alpha x}(C_1 \cos \beta x + C_2 \sin \beta x)$.

Полагая в этом равенстве $\alpha = 0$, $\beta = 4$, получим общее решение данного уравнения: $y = C_1 \cos 4x + C_2 \sin 4x$;

г) характеристическое уравнение $k^2 - 4k + 20 = 0$ имеет комплексные корни $k_1 = 2 + 4i$, $k_2 = 2 - 4i$. По правилу 3°, полагая в равенстве (4.123) $\alpha = 2$, $\beta = 4$, получим общее решение заданного уравнения

$$y = e^{2x}(C_1 \cos 4x + C_2 \sin 4x). \blacktriangleright$$

Линейные неоднородные дифференциальные уравнения второго порядка с постоянными коэффициентами.

ОПРЕДЕЛЕНИЕ. Уравнение вида

$$y'' + py' + qy = f(x), \quad (4.124)$$

где p и q — действительные числа, $f(x)$ — известная непрерывная функция, называется **линейным неоднородным** дифференциальным уравнением второго порядка с постоянными коэффициентами.

Для общего решения неоднородного уравнения (4.124) справедлива следующая теорема.

□ ТЕОРЕМА 4.43. *Общее решение y неоднородного уравнения (4.124) равно сумме общего решения $y_{\text{одн}}$ соответствующего однородного уравнения*

$$y'' + py' + qy = 0$$

и любого частного решения $\bar{y}$ данного неоднородного уравнения. ■

Согласно этой теореме для решения уравнения (4.124) вначале находится функция $y_{\text{одн}}$ (по правилам 1°–3°), а затем — функция $\bar{y}$. Их сумма и дает общее решение y неоднородного уравнения:

$$y = y_{\text{одн}} + \bar{y}.$$

Для некоторых специальных видов функции $f(x)$ частное решение $\bar{y}$ можно найти методом неопределенных коэффициентов. По виду правой части $f(x)$ можно заранее указать вид частного решения $\bar{y}$, где неизвестны лишь числовые коэффициенты, в следующих простейших случаях.

СЛУЧАЙ 1. $f(x) = P(x)$, где $P(x)$ — многочлен некоторой степени.

В этом случае $\bar{y}$ есть многочлен $Q(x)$ той же самой степени, что и $P(x)$, если число 0 не является корнем характеристического уравнения; если же число 0 является корнем характеристического уравнения кратности r , то $\bar{y} = x^r \cdot Q(x)$.

СЛУЧАЙ 2. $f(x) = e^{mx}$ (a, m — некоторые числа).

В этом случае $\bar{y} = Ae^{mx}$, если число m не является корнем характеристического уравнения, и $\bar{y} = Ax^r e^{mx}$, если число m является корнем характеристического уравнения кратности r .

Здесь A — подлежащий определению коэффициент.

СЛУЧАЙ 3. $f(x) = e^{mx} \cdot P(x)$, где $P(x)$ — многочлен некоторой степени.

В этом случае $\bar{y} = e^{mx} \cdot Q(x)$, если число m не является корнем характеристического уравнения, и $\bar{y} = x^r e^{mx} \cdot Q(x)$, если число m — корень характеристического уравнения кратности r .

Здесь $Q(x)$ — многочлен той же степени, что и $P(x)$, коэффициенты которого подлежат определению.

Заметим, что из случая 3 при $m = 0$ получаем случай 1; а если $P(x) = a$ (многочлен нулевой степени), то $f(x) = ae^{mx}$, и из случая 3 получаем случай 2.

СЛУЧАЙ 4. $f(x) = e^{mx}(a \cos nx + b \sin nx)$.

Тогда $\bar{y} = e^{mx}(A \cos nx + B \sin nx)$, если числа $m \pm ni$ не являются корнями характеристического уравнения, и $\bar{y} = xe^{mx}(A \cos nx + B \sin nx)$, если числа $m \pm ni$ являются корнями характеристического уравнения.

Здесь A и B — подлежащие определению коэффициенты.

СЛУЧАЙ 5. Правая часть уравнения (4.124) — функция $f(x)$ — есть сумма указанных функций.

Тогда частное решение $\bar{y}$ этого уравнения есть сумма частных решений уравнений с той же левой частью, что и уравнение (4.124), а правые части этих уравнений есть каждое слагаемое правой части уравнения (4.124).

▷ **ПРИМЕР 4.52.** Решить уравнения:

а) $y'' + y' - 2y = 6x^2$. б) $y'' - 3y' = 2 - 6x$.

в) $y'' - 6y' + 9y = -8e^x$. г) $y'' + 5y' = e^{-5x}$.

д) $y'' - 4y = xe^{-x}$. е) $y'' + 9y = 12 \cos 3x + 18 \sin 3x$.

ж) $y'' - y' = (2x + 3)e^x$.

Решение. а) Сначала находим общее решение однородного уравнения $y'' + y' - 2y = 0$, соответствующего данному неоднородному уравнению. Его характеристическое уравнение $k^2 + k - 2 = 0$ имеет корни $k_1 = -2$, $k_2 = 1$. Поэтому (согласно правилу 1°) $y_{\text{одн}} = C_1 e^{-2x} + C_2 e^x$.

Теперь находим частное решение $\bar{y}$ данного неоднородного уравнения. Для правой части данного уравнения $f(x) = 6x^2$. Согласно указанному правилу (случай 1, число 0 не является корнем характеристического уравнения) $\bar{y}$ есть многочлен той же степени, что и $f(x) = 6x^2$, то есть многочлен второй степени: $\bar{y} = Ax^2 + Bx + C$. Отсюда, дифференцируя, находим $\bar{y}' = 2Ax + B$, $\bar{y}'' = 2A$

и подставляя $\bar{y}, \bar{y}', \bar{y}''$ в данное уравнение, получим равенство $2A + 2Ax + B - 2Ax^2 - 2Bx - 2C = 6x^2$.

Приравнявая коэффициенты при одинаковых степенях x из обеих его частей, а только при этом условии оно будет тождественным, получим систему

$$\begin{cases} -2A = 6; \\ 2A - 2B = 0; \\ 2A + B - 2C = 0, \end{cases}$$

из которой находим $A = -3, B = -3, C = -4,5$. Следовательно, $y = -3x^2 - 3x - 4,5$, а искомое общее решение данного неоднородного уравнения

$$y = y_{\text{одн}} + \bar{y} = C_1 e^{-2x} + C_2 e^x - 3x^2 - 3x - 4,5.$$

б) Составляем характеристическое уравнение $k^2 - 3k = 0$, определяем его корни $k_1 = 0, k_2 = 3$ и (согласно правилу 2°) находим общее решение $y_{\text{одн}}$ однородного уравнения, соответствующего данному неоднородному уравнению: $y_{\text{одн}} = C_1 + C_2 e^{3x}$.

Частным решением $\bar{y}$ данного неоднородного уравнения в соответствии с его правой частью $f(x) = 2 - 6x$ (случай 1, число 0 — корень характеристического уравнения) будет функция вида $\bar{y} = x(Ax + B) = Ax^2 + Bx$.

Подставляя функцию $\bar{y}$ и ее производные $\bar{y}' = 2Ax + B, \bar{y}'' = 2A$ в данное неоднородное уравнение, получим равенство

$$2A - 6Ax - 3B = 2 - 6x.$$

Приравнявая коэффициенты при одинаковых степенях x обеих его частей, получим систему

$$\begin{cases} -6A = -6, \\ 2A - 3B = 2. \end{cases}$$

Решая ее, находим $A = 1, B = 0$. Следовательно, $\bar{y} = x^2$, общее решение

$$y = y_{\text{одн}} + \bar{y} = C_1 + C_2 e^{3x} + x^2.$$

в) Здесь характеристическое уравнение $k^2 - 6k + 9 = 0$ имеет корни $k_1 = k_2 = 3$. Поэтому общее решение однородного уравне-

ния есть функция $y_{\text{одн}} = C_1 e^{3x} + C_2 x e^{3x}$. Частное решение $\bar{y}$ неоднородного уравнения следует искать в виде $\bar{y} = A e^x$ согласно указанному правилу (случай 2, число $m = 1$ не является корнем характеристического уравнения). Находим $y' = A e^x$, $y'' = A e^x$. Подставим $\bar{y}$, $\bar{y}'$, $\bar{y}''$ в данное уравнение и определим значение коэффициента A :

$$A e^x - 6A e^x + 9A e^x = -8e^x; 4A e^x = -8e^x; A = -2.$$

$$\text{Следовательно, } \bar{y} = -2e^x, y = y_{\text{одн}} + \bar{y} = C_1 e^{3x} + C_2 x e^{3x} - 2e^x.$$

г) Характеристическое уравнение $k^2 + 5k = 0$ имеет корни $k_1 = 0$, $k_2 = -5$. Поэтому общее решение соответствующего однородного уравнения (по правилу 1°) $y_{\text{одн}} = C_1 + C_2 e^{-5x}$.

Согласно указанному правилу и случаю 2 число $m = -5$ является корнем характеристического уравнения $\bar{y} = x A e^{-5x}$. Находим

$$y' = A e^{-5x} - 5A x e^{-5x},$$

$$\bar{y}'' = -5A e^{-5x} - 5A e^{-5x} + 25A x e^{-5x} = -10A e^{-5x} + 25A x e^{-5x}.$$

Подставляя $\bar{y}$, $\bar{y}'$, $\bar{y}''$ в данное уравнение, получим

$$-10A e^{-5x} + 25A x e^{-5x} + 5A e^{-5x} - 25A x e^{-5x} = e^{-5x},$$

отсюда найдем $A = -0,2$.

$$\text{Следовательно, } \bar{y} = -0,2x e^{-5x}, y = y_{\text{одн}} + \bar{y} = C_1 + C_2 e^{-5x} - 0,2x e^{-5x}.$$

д) Характеристическое уравнение $k^2 - 4 = 0$ имеет корни $k_1 = -2$, $k_2 = 2$. Поэтому общее решение соответствующего однородного уравнения:

$$y_{\text{одн}} = C_1 e^{-2x} + C_2 e^{2x}.$$

Частным решением неоднородного уравнения будет функция $\bar{y} = (Ax + B)e^{-x}$ (случай 3, $m = -1$ не является корнем характеристического уравнения). Находим

$$\bar{y}' = A e^{-x} - (Ax + B)e^{-x} = (A - B)e^{-x} - A x e^{-x},$$

$$\bar{y}'' = -(A - B)e^{-x} - A e^{-x} + A x e^{-x} = (B - 2A)e^{-x} + A x e^{-x}.$$

Подставим $\bar{y}$, $\bar{y}'$, $\bar{y}''$ в данное уравнение, получим:

$$(B - 2A)e^{-x} + A x e^{-x} - 4(Ax + B)e^{-x} = x e^{-x};$$

$$(B - 2A + Ax - 4Ax - 4B)e^{-x} = x e^{-x}; -2A - 3B - 3Ax = x.$$

Отсюда, приравнивая коэффициенты при одинаковых степенях x в обеих частях равенства, получим систему уравнений для определения A и B :

$$\begin{cases} -3A = 1, \\ -2A - 3B = 0. \end{cases}$$

Решая ее, находим $A = -\frac{1}{3}$, $B = \frac{2}{9}$. Следовательно,

$$\bar{y} = \left(-\frac{1}{3}x + \frac{2}{9} \right) e^{-x}. \text{ Значит, } y = y_{\text{одн}} + \bar{y} = C_1 e^{-2x} + C_2 e^{2x} + \\ + \left(-\frac{1}{3}x + \frac{2}{9} \right) e^{-x}.$$

е) Характеристическое уравнение $k^2 + 9 = 0$ имеет корни $k_1 = 3i$, $k_2 = -3i$. Поэтому общее решение соответствующего однородного уравнения (см. правило 3°, где $\alpha = 0$, $\beta = 3$) $y_{\text{одн}} = C_1 \cos 3x + C_2 \sin 3x$.

Частное решение $\bar{y}$ данного неоднородного уравнения будет $\bar{y} = x(A \cos 3x + B \sin 3x)$ (случай 4, $m = 0$, $n = 3$ и числа $m \pm ni = \pm 3i$ являются корнями характеристического уравнения).

Дифференцируя дважды это равенство, находим $\bar{y}''$:

$$\begin{aligned} \bar{y}' &= (A \cos 3x + B \sin 3x) + x(-3A \sin 3x + 3B \cos 3x) = \\ &= (A + 3Bx) \cos 3x + (B - 3Ax) \sin 3x; \\ \bar{y}'' &= 3B \cos 3x - (A + 3Bx) \cdot 3 \sin 3x - 3A \sin 3x + (B - 3Ax) \times \\ &\times 3 \cos 3x = (6B - 9Ax) \cos 3x + (-6A - 9Bx) \sin 3x. \end{aligned}$$

Подставив $\bar{y}$ и $\bar{y}''$ в данное уравнение, получим

$$\begin{aligned} 6B \cos 3x - 9Ax \cos 3x - 6A \sin 3x - 9Bx \sin 3x + 9Ax \cos 3x + \\ + 9Bx \sin 3x &= 12 \cos 3x + 18 \sin 3x; \\ 6B \cos 3x - 6A \sin 3x &= 12 \cos 3x + 18 \sin 3x. \end{aligned}$$

Приравнивая коэффициенты у подобных членов в обеих частях равенства, найдем $6B = 12$; $-6A = 18$. Значит, $A = -3$, $B = 2$.

Следовательно, $\bar{y} = x(-3\cos 3x + 2\sin 3x)$,

$$y = y_{\text{одн}} + \bar{y} = C_1 \cos 3x + C_2 \sin 3x + x(2\sin 3x - 3\cos 3x).$$

ж) Характеристическое уравнение $k^2 - k = 0$ имеет корни $k_1 = 0$, $k_2 = 1$. Поэтому общее решение соответствующего однородного уравнения

$$y_{\text{одн}} = C_1 + C_2 e^x.$$

Частным решением неоднородного уравнения будет функция $\bar{y} = x(Ax + B)e^x = (Ax^2 + Bx)e^x$ (согласно указанному правилу, случай 3, $m = 1$ является корнем характеристического уравнения). Находим

$$\begin{aligned} \bar{y}' &= (2Ax + B)e^x + (Ax^2 + Bx)e^x = [Ax^2 + (2A + B)x + B]e^x, \\ \bar{y}'' &= [2Ax + 2A + B]e^x + [Ax^2 + (2A + B)x + B]e^x = \\ &= [Ax^2 + (4A + B)x + 2A + 2B]e^x. \end{aligned}$$

Подставляя $\bar{y}$ и $\bar{y}''$ в данное уравнение и сокращая все слагаемые на множитель $e^x \neq 0$, получаем

$$Ax^2 + (4A + B)x + 2A + 2B - Ax^2 - (2A + B)x - B = 2x + 3,$$

или после упрощения $2Ax + 2A + B = 2x + 3$.

Отсюда, приравнивая коэффициенты при одинаковых степенях x и свободные члены в обеих частях равенства, получим

$$\begin{cases} 2A = 2, \\ 2A + B = 3. \end{cases}$$

Откуда $A = 1$, $B = 1$. Следовательно, $\bar{y} = (x^2 + x)e^x$, а искомое решение данного неоднородного уравнения:

$$y = y_{\text{одн}} + \bar{y} = C_1 + C_2 e^x + x(x + 1)e^x. \blacktriangleright$$

§ 23. Числовые ряды

23.1. Понятие ряда. Признаки сходимости рядов

Числовой ряд. Определение ряда, понятие сходимости.

ОПРЕДЕЛЕНИЕ. *Числовым рядом* называется выражение

$$a_1 + a_2 + a_3 + \dots + a_n + \dots = \sum_{n=1}^{\infty} a_n, \quad (4.125)$$

где числа $a_1, a_2, a_3, \dots, a_n, \dots$, называемые **членами ряда**, образуют известную числовую последовательность.

Число a_n называется *общим членом ряда*. Общий член ряда является функцией от n . Если известно аналитическое выражение этой функции, то, придавая числу n последовательно значения 1, 2, 3, ..., можно найти сколько угодно членов ряда.

ОПРЕДЕЛЕНИЕ. Числовой ряд (4.125) называется *сходящимся*, если сумма первых его членов $S_n = a_1 + a_2 + \dots + a_n$ при $n \rightarrow \infty$ имеет конечный предел. Этот предел называется *суммой* сходящегося ряда. Если же конечный предел суммы S_n не существует, то ряд называется *расходящимся*. В этом случае нет смысла говорить о его сумме.

Если ряд (4.125) сходится и его сумма равна S , то разность $S - S_n = R_n$ называется *n-м остатком ряда*:

$$R_n = a_{n+1} + a_{n+2} + \dots$$

Очевидно, R_n — также числовой ряд.

Для установления сходимости или расходимости рядов применяются признаки сходимости.

Необходимый признак сходимости ряда.

□ **ТЕОРЕМА 4.44** (необходимый признак сходимости). *Если числовой ряд $a_1 + a_2 + a_3 + \dots + a_n + \dots$ сходится, то его общий член a_n стремится к нулю при $n \rightarrow \infty$, то есть $\lim_{n \rightarrow \infty} a_n = 0$.* ■

Это необходимый, но не достаточный признак сходимости для всякого ряда.

Если же $\lim_{n \rightarrow \infty} a_n \neq 0$, то ряд расходится. (Это достаточный признак расходимости для всякого ряда.)

▷ ПРИМЕР 4.53. Написать пять первых членов ряда $\sum_{n=1}^{\infty} \frac{2n+3}{n(n+1)}$.

Р е ш е н и е. Подставляя в формулу общего члена $a_n = \frac{2n+3}{n(n+1)}$ последовательно значения $n = 1, 2, 3, 4, 5$, получим

$$a_1 = \frac{5}{2}; a_2 = \frac{7}{6}; a_3 = \frac{3}{4}; a_4 = \frac{11}{20}; a_5 = \frac{13}{30} . \blacktriangleright$$

▷ ПРИМЕР 4.54. Написать формулу общего члена для каждого ряда:

$$\text{а) } \frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \dots ; \qquad \text{б) } \frac{1}{2} + \frac{3}{2^2} + \frac{5}{2^3} + \dots .$$

Р е ш е н и е. Для ряда а) числители членов ряда — натуральный ряд чисел, знаменатели — числа, которые могут быть получены по формуле $n + 1$, где $n = 1, 2, 3, \dots$.

Следовательно, общий член ряда $a_n = \frac{n}{n+1}$.

Числители членов ряда б) — нечетные числа вида $2n - 1$, а знаменатели могут быть получены по формуле 2^n , то есть $a_n = \frac{2n-1}{2^n}$ — общий член ряда б). $\blacktriangleright$

ОПРЕДЕЛЕНИЕ. Ряд $\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots$ называется *гармоническим*.

Этот ряд расходится, хотя для него выполняется необходимый признак сходимости.

Достаточные признаки сходимости рядов с положительными членами. Применение достаточных признаков при исследовании рядов требуется в тех случаях, когда выполнен необходимый признак сходимости ($\lim_{n \rightarrow \infty} a_n = 0$).

ПЕРВЫЙ ПРИЗНАК СРАВНЕНИЯ. Если *каждый член* ряда

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \dots + a_n + \dots \quad (4.126)$$

с положительными членами (4.126), начиная с некоторого номера, *не превосходит* соответствующего члена сходящегося ряда (4.127)

$$\sum_{n=1}^{\infty} b_n = b_1 + b_2 + b_3 + \dots + b_n + \dots, \quad (4.127)$$

то ряд (4.126) *тоже сходится*.

Если же каждый член ряда (4.126), начиная с некоторого номера, не меньше соответствующего члена расходящегося ряда (4.127), то данный ряд (4.126) *тоже расходится*.

При применении этого признака исследуемый ряд часто сравнивается или с бесконечной геометрической прогрессией:

$$1 + q + q^2 + \dots + q^n + \dots = \sum_{n=0}^{\infty} q^n \quad (q > 0),$$

которая сходится при $0 < q < 1$, а при $q \geq 1$ расходится, или с расходящимся гармоническим рядом:

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots = \sum_{n=1}^{\infty} \frac{1}{n}.$$

ВТОРОЙ ПРИЗНАК СРАВНЕНИЯ. Если существует конечный отличный от нуля предел при $n \rightarrow \infty$ отношения общих членов рядов (4.126) и (4.127) ($\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = l \neq 0$), *то эти ряды одновременно либо сходятся, либо расходятся*.

▷ ПРИМЕР 4.55. Исследовать сходимость ряда $\sum_{n=1}^{\infty} \sin \frac{1}{n}$.

Р е ш е н и е. Сравним данный ряд с расходящимся гармоническим рядом $\sum_{n=1}^{\infty} \frac{1}{n}$ и применим второй признак сравнения. Так как

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \lim_{n \rightarrow \infty} \frac{\sin \frac{1}{n}}{\frac{1}{n}} = 1 \neq 0, \text{ то данный ряд расходится. } \blacktriangleright$$

ПРИЗНАК ДАЛАМБЕРА. Пусть для ряда с *положительными членами*

$$a_1 + a_2 + a_3 + \dots + a_n + \dots$$

существует предел отношения последующего члена к предыдущему при $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = q,$$

тогда *при* $0 < q < 1$ *ряд сходится*, а *при* $q > 1$ *ряд расходится*. При $q = 1$ вопрос о сходимости ряда остается нерешенным.

▷ ПРИМЕР 4.56. Исследовать по признаку Даламбера сходимость рядов:

$$\text{а) } \sum_{n=1}^{\infty} \frac{n^2}{2n}; \quad \text{б) } \sum_{n=1}^{\infty} \frac{n!}{3n}; \quad \text{в) } \sum_{n=1}^{\infty} \frac{3^n}{2^n(2n+1)}.$$

Р е ш е н и е. а) Зная n -й член ряда, находим следующий за ним $(n+1)$ -й член, заменяя в выражении n -го члена n на $n+1$. Затем ищем предел отношения последующего члена a_{n+1} к предыдущему a_n при $n \rightarrow \infty$:

$$a_n = \frac{n^2}{2n}; \quad a_{n+1} = \frac{(n+1)^2}{2^{n+1}}.$$

$$q = \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \lim_{n \rightarrow \infty} \frac{2^n \cdot (n+1)^2}{2^{n+1} \cdot n^2} = \lim_{n \rightarrow \infty} \frac{2^n}{2^n \cdot 2} \cdot \lim_{n \rightarrow \infty} \left(\frac{n+1}{n} \right)^2 =$$

$$= \frac{1}{2} \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^2 = \frac{1}{2}.$$

Так как $q = \frac{1}{2} < 1$, то согласно признаку Даламбера данный ряд сходится.

$$\text{б) } a_n = \frac{n!}{3^n}; \quad a_{n+1} = \frac{(n+1)!}{3^{n+1}}.$$

$$q = \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \lim_{n \rightarrow \infty} \frac{(n+1)! 3^n}{3^{n+1} \cdot n!} = \lim_{n \rightarrow \infty} \frac{n+1}{3} = \infty.$$

Согласно признаку Даламбера данный ряд расходится.

$$\text{в) } a_n = \frac{3^n}{2^n(2n+1)}; \quad a_{n+1} = \frac{3^{n+1}}{2^{n+1}(2n+3)}.$$

$$q = \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \lim_{n \rightarrow \infty} \frac{3^{n+1} \cdot 2^n \cdot (2n+1)}{2^{n+1} \cdot (2n+3) \cdot 3^n} = \frac{3}{2} \cdot \lim_{n \rightarrow \infty} \frac{2n+1}{2n+3} =$$

$$= \frac{3}{2} \lim_{n \rightarrow \infty} \frac{2 + \frac{1}{n}}{2 + \frac{3}{n}} = \frac{3}{2}.$$

Здесь $q = \frac{3}{2} > 1$. По признаку Даламбера данный ряд расходится. ►

ИНТЕГРАЛЬНЫЙ ПРИЗНАК СХОДИМОСТИ (КОШИ).

Если функция $f(x)$ на интервале $[1; \infty]$ непрерывна, положительна, монотонно убывает, то **числовой ряд** $a_1 + a_2 + \dots + a_n + \dots$, где $a_n = f(n)$, **сходится** (расходится), **если сходится** (расходится) **несобственный**

интеграл $\int_1^{\infty} f(x) dx$.

► ПРИМЕР 4.57. Исследовать сходимость ряда $\sum_{n=1}^{\infty} \frac{1}{n^a}$, где a — действительное число.

Р е ш е н и е. При $a \leq 0$ предел при $n \rightarrow \infty$ общего члена ряда отличен от нуля, то есть не выполняется необходимый признак сходимости ряда. Поэтому ряд в этом случае расходится. Для исследования сходимости ряда при $a > 0$ применим признак Коши.

1. Пусть $a = 1$. Тогда $a_n = \frac{1}{n}$ (ряд гармонический); тогда интеграл

$$\int_1^{\infty} \frac{dx}{x} = \lim_{b \rightarrow \infty} \int_1^b \frac{dx}{x} = \lim_{b \rightarrow \infty} [\ln x]_1^b = \lim_{b \rightarrow \infty} (\ln b - \ln 1) = \infty.$$

Несобственный интеграл расходится, тем самым расходится и данный гармонический ряд.

2. Пусть $a > 1$. Тогда

$$\begin{aligned} \int_1^{\infty} \frac{dx}{x^a} &= \lim_{b \rightarrow \infty} \int_1^b x^{-a} dx = \lim_{b \rightarrow \infty} \left[\frac{x^{-a+1}}{1-a} \right]_1^b = \frac{1}{1-a} \lim_{b \rightarrow \infty} \left[\frac{1}{x^{a-1}} \right]_1^b = \\ &= \frac{1}{1-a} \lim_{b \rightarrow \infty} \left(b^{\frac{1}{a-1}} - 1 \right) = \frac{1}{a-1}. \end{aligned}$$

Сходимость этого интеграла обосновывает сходимость ряда.

3. Если $0 < a < 1$, то $\lim_{b \rightarrow \infty} \frac{1}{b^{a-1}} = \infty$ и несобственный интеграл расходится.

Таким образом, данный ряд сходится при $a > 1$ и расходится при $a \leq 1$. ►

23.2. Знакопередающиеся ряды. Их признаки сходимости

ОПРЕДЕЛЕНИЕ. Ряд называется *знакопередающимся* (или *знакопеременным*), если любые два соседних члена его противоположны по своим знакам. Знакопередающийся ряд можно записать так:

$$a_1 - a_2 + a_3 - a_4 + \dots + (-1)^{n+1} a_n + \dots,$$

где все числа a_n ($n = 1, 2, 3, \dots$) положительны.

Сходимость знакопередающегося ряда может быть установлена признаком сходимости знакопередающегося ряда (признаком Лейбница).

ПРИЗНАК ЛЕЙБНИЦА. Если члены знакопередающегося ряда *монотонно убывают по абсолютной величине* и $\lim_{n \rightarrow \infty} a_n = 0$, то такой *ряд сходится*, и его сумма по абсолютной величине не превосходит абсолютной величины первого члена.

▷ ПРИМЕР 4.58. Исследовать сходимость знакочередующихся рядов:

$$\text{а) } \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{\sqrt{n}}; \quad \text{б) } \sum_{n=1}^{\infty} \frac{(-1)^n n}{n+1}.$$

Р е ш е н и е. а) Члены данного ряда монотонно убывают по абсолютной величине:

$$1 > \frac{1}{\sqrt{2}} > \frac{1}{\sqrt{3}} > \dots \quad \text{и} \quad \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} = 0.$$

Поэтому, согласно признаку Лейбница, данный ряд сходится.

б) Члены данного ряда по абсолютной величине монотонно возрастают:

$$\frac{1}{2} < \frac{2}{3} < \frac{3}{4} < \dots \quad \text{и} \quad \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1 \neq 0.$$

Условия признака Лейбница не выполняются, значит, исследуемый ряд расходится. ►

Абсолютная и условная сходимость знакочередующегося ряда.

ОПРЕДЕЛЕНИЕ. Знакопеременный ряд (4.128)

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \dots + a_n + \dots \quad (4.128)$$

называется **абсолютно сходящимся**, если сходится ряд, составленный из абсолютных величин его членов (4.129):

$$\sum_{n=1}^{\infty} |a_n| = |a_1| + |a_2| + |a_3| + \dots + |a_n| + \dots \quad (4.129)$$

ОПРЕДЕЛЕНИЕ. Знакопеременный сходящийся ряд (4.128) называется **условно сходящимся**, если ряд (4.129) расходится.

УТВЕРЖДЕНИЕ. *Всякий абсолютно сходящийся ряд есть ряд сходящийся.*

▷ ПРИМЕР 4.59. Исследовать сходимость знакопеременного ряда (определить: является ли он абсолютно сходящимся, условно сходящимся или расходящимся):

$$а) \frac{1}{\ln 2} - \frac{1}{\ln 3} + \frac{1}{\ln 4} - \frac{1}{\ln 5} + \dots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\ln(n+1)}; б) \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n5^n};$$

$$в) \sum_{n=1}^{\infty} \frac{(-1)^n n^2}{2n^2 + 1}.$$

Р е ш е н и е. а) Члены данного знакочередующегося ряда монотонно убывают по абсолютной величине:

$$\frac{1}{\ln 2} > \frac{1}{\ln 3} > \frac{1}{\ln 4} > \dots \quad \text{и} \quad \lim_{n \rightarrow \infty} \frac{1}{\ln(n+1)} = 0.$$

Поэтому, согласно признаку Лейбница, данный ряд сходится. Чтобы установить, сходится он абсолютно или условно, исследу-

ем ряд с положительными членами $\sum_{n=1}^{\infty} \frac{1}{\ln(n+1)}$, составленный из абсолютных величин членов данного ряда.

Применяя признак сравнения с расходящимся гармоническим рядом, получим $\frac{1}{\ln(n+1)} > \frac{1}{n}$, значит, ряд с положительными членами расходится. Следовательно, данный ряд а) сходится условно.

б) Члены данного знакочередующегося ряда монотонно убывают по абсолютной величине:

$$\frac{1}{5} > \frac{1}{2 \cdot 5^2} > \frac{1}{3 \cdot 5^3} > \dots \quad \text{и} \quad \lim_{n \rightarrow \infty} \frac{1}{n5^n} = 0.$$

Поэтому, согласно признаку Лейбница, он сходится. Ряд, составленный из абсолютных величин членов данного ряда, также сходится согласно признаку Даламбера:

$$\begin{aligned} a_n &= \frac{1}{n5^n}, \quad a_{n+1} = \frac{1}{(n+1)5^{n+1}}, \quad q = \lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \lim_{n \rightarrow \infty} \frac{n5^n}{(n+1)5^{n+1}} = \\ &= \lim_{n \rightarrow \infty} \frac{n}{5(n+1)} = \frac{1}{5} < 1. \end{aligned}$$

Следовательно, данный ряд абсолютно сходящийся.

в) Для данного знакопеременного ряда не выполняется необходимый признак сходимости:

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{n^2}{2n^2 + 1} = \lim_{n \rightarrow \infty} \frac{1}{2 + \frac{1}{n^2}} = \frac{1}{2} \neq 0.$$

Значит, данный ряд расходится. ►

23.3. Степенной ряд и область его сходимости

Функциональные ряды. Их сходимость.

ОПРЕДЕЛЕНИЕ. Ряд $\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + \dots + u_n(x) + \dots$, члены которого являются функциями от переменной x , имеющими общую область определения, называется **функциональным**.

При различных значениях x из функционального ряда получаем числовые ряды, которые могут быть сходящимися или расходящимися.

Совокупность значений x , при которых функциональный ряд сходится, называется его *областью сходимости*.

Из всех функциональных рядов простейшими и наиболее употребительными являются степенные ряды вида:

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots \quad (4.130)$$

Степенной ряд (4.130) сходится при $x = 0$. Определение области сходимости степенного ряда (4.130) базируется на **теореме Абеля**.

□ **ТЕОРЕМА АБЕЛЯ.** Если степенной ряд (4.130) сходится при $x = x_0$ ($x_0 \neq 0$), то он сходится абсолютно при любом значении x , удовлетворяющем неравенству: $|x| < |x_0|$. Если же степенной ряд (4.130) расходится при $x = x_0$, то он расходится при любом значении x , удовлетворяющем неравенству $|x| > |x_0|$. ■

Из теоремы Абеля следует, что если степенной ряд сходится в точке $x_0 \neq 0$, то он сходится в интервале $(-|x_0|, |x_0|)$.

Можно доказать, что для всякого степенного ряда, который имеет точки сходимости (кроме точки $x = 0$) и точки расходимости, существует некоторое число $R > 0$; при этом ряд сходится во всех точках, для которых $|x| < R$. Это число называется *радиусом сходимости ряда* (4.130).

Если ряд (4.130) сходится только при $x = 0$, то полагаем $R = 0$. Если же ряд (4.130) сходится при любом значении x , то полагают $R = \infty$. Интервалом сходимости ряда (4.130) называется интервал $(-R, R)$.

Чтобы найти область сходимости степенного ряда (4.130), надо определить интервал сходимости $(-R, R)$, а затем выяснить вопрос о сходимости ряда на концах интервала, то есть при $x = -R$ и при $x = R$.

Радиус сходимости степенного ряда (4.130) отыскивают с помощью признака Даламбера.

Применим признак Даламбера к ряду, составленному из абсолютных величин членов ряда (4.130):

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \cdot |x| = q \cdot |x|, \quad \text{где } q = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|.$$

Ряд (4.130) сходится абсолютно при всех значениях x , удовлетворяющих неравенству:

$$q|x| < 1, \quad \text{или } |x| < \frac{1}{q}, \quad \text{или } -\frac{1}{q} < x < \frac{1}{q}, \quad -R < x < R,$$

где радиус сходимости $R = \frac{1}{q} = \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right|.$

▷ ПРИМЕР 4.60. Определить область сходимости степенного ряда:

$$\sum_{n=1}^{\infty} \frac{x^n}{n2^n}.$$

Решение. Здесь $a_n = \frac{1}{n2^n}$; $a_{n+1} = \frac{1}{(n+1)2^{n+1}}$. Поскольку

$$R = \lim_{n \rightarrow \infty} \frac{(n+1)2^{n+1}}{n2^n} = 2 \lim_{n \rightarrow \infty} \frac{n+1}{n} = 2,$$

исследуемый ряд сходится при всех значениях x , принадлежащих интервалу $(-2; 2)$.

Граничные точки $x = \pm 2$ этого интервала, для которых $q = 1$ и признак Даламбера не решает вопроса о сходимости ряда, исследуем особо.

При $x = -2$ получим числовой знакочередующийся ряд:

$$\sum_{n=1}^{\infty} \frac{(-2)^n}{2^n n} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n},$$

который сходится согласно признаку Лейбница (члены этого ряда монотонно убывают по абсолютной величине и общий член ряда стремится к нулю).

При $x = 2$ получим гармонический ряд $\sum_{n=1}^{\infty} \frac{1}{n}$, который расходится.

Следовательно, областью сходимости данного ряда является полуоткрытый интервал $[-2; 2]$. ►

Ряды Тейлора и Маклорена. Если функция $f(x)$ имеет в окрестности точки $x = a$ производные $(n+1)$ -го порядка, то по формуле Тейлора (4.131) она может быть представлена в виде суммы многочлена степени n и остаточного члена $R_n(x)$:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x), \quad (4.131)$$

где $R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$, $a < c < x$ или $x < c < a$.

Если в формуле (4.131) положить $n \rightarrow \infty$ и $\lim_{n \rightarrow \infty} R_n(x) = 0$, то получим снова ряд Тейлора:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \dots \quad (4.132)$$

Положив в ряде Тейлора $a = 0$, получим ряд **Маклорена**

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \dots \quad (4.133)$$

Для разложимости функции $f(x)$ в ряд Маклорена необходимо и достаточно, чтобы

$$\lim_{n \rightarrow \infty} \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1} = 0, \quad (4.134)$$

где число c заключено между 0 и x .

Если это условие не выполняется, то степенной ряд (4.133) не представляет собой функцию $f(x)$. Если условие (4.134) выполняется на некотором интервале, то на этом интервале составленный ряд Маклорена сходится к функции $f(x)$.

Для разложения данной функции в ряд Маклорена нужно:

1) вычислить значения этой функции и ее производных при $x = 0$ и подставить их в общее выражение ряда Маклорена (4.133) для произвольной функции;

2) определить совокупность значений x , при которых полученный ряд сходится к данной функции (т.е. при которых выполняется условие (4.134)).

Для многих функций, употребляемых на практике, каждая точка x сходимости ряда Маклорена является и точкой сходимости этого ряда к породившей его функции. Поэтому при разложении многих функций в ряд Маклорена можно вместо проверки выполнения условия (4.134), что во многих случаях весьма затруднительно, исследовать сходимость самого ряда Маклорена как обычного степенного ряда.

▷ ПРИМЕР 4.61. Разложить в ряд Маклорена функцию $f(x) = e^x$.

Решение. Последовательно дифференцируя функцию $f(x) = e^x$, будем иметь:

$$f'(x) = e^x, f''(x) = e^x, \dots, f^{(n)}(x) = e^x.$$

Вычислим значения самой функции и ее производных при $x = 0$:

Подставляя полученные значения $f(0) = 1, f'(0) = 1, f''(0) = 1, \dots, f^{(n)}(0) = 1$ в общее выражение ряда Маклорена для произвольной функции, получим

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

Исследуем сходимость полученного ряда по признаку Даламбера:

$$u_n = \frac{x^n}{n!}, u_{n+1} = \frac{x^{n+1}}{(n+1)!}, R = \lim_{n \rightarrow \infty} \left| \frac{u_n}{u_{n+1}} \right| = \lim_{n \rightarrow \infty} \left| \frac{x^n (n+1)!}{n! x^{n+1}} \right| =$$

$$= \frac{1}{|x|} \lim_{n \rightarrow \infty} (n+1) = \infty,$$

т. е. данный ряд сходится при любом x . Можно доказать, что при любом x он сходится именно к данной функции $f(x) = e^x$. Итак,

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots, (-\infty < x < \infty). \blacktriangleright$$

Аналогично получают разложения в ряд следующих функций:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^{(n+1)} \frac{x^{2n+1}}{(2n+1)!} + \dots, (-\infty < x < \infty).$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + \dots, (-\infty < x < \infty).$$

$$(1+x)^m = 1 + \frac{m}{1!}x + \frac{m(m-1)}{2!}x^2 + \dots + \frac{m(m-1)\dots(m-n+1)}{n!}x^n +$$

$$+ \dots, (-1 < x < 1).$$

$$\ln(1+x) = \frac{x}{1} - \frac{x^2}{2!} + \frac{x^3}{3!} - \dots + (-1)^{n-1} \frac{x^n}{n} + \dots, (-1 < x \leq 1).$$

Операцию разложения элементарных функций в степенные ряды позволяет значительно упростить применение следующих свойств степенных рядов:

1) Два степенных ряда можно почленно складывать и умножать (по правилу умножения многочленов). При этом областью сходимости полученного нового степенного ряда будет совокупность всех точек, в которых одновременно сходятся оба ряда.

2) Степенной ряд в области его сходимости можно почленно интегрировать, а внутри области сходимости можно почленно дифференцировать.

▷ ПРИМЕР 4.62. Разложить в степенной ряд функцию $f(x) = \cos(-3x)$.

Р е ш е н и е. Заменяя x в ряде Маклорена для $\cos x$ на $(-3x)$, получим

$$\cos(-3x) = 1 - \frac{(-3x)^2}{2!} + \frac{(-3x)^4}{4!} - \dots + (-1)^n \frac{(-3x)^{2n}}{(2n)!} + \dots,$$

или

$$\cos(-3x) = 1 - \frac{3^2 x^2}{2!} + \frac{3^4 x^4}{4!} - \frac{3^6 x^6}{6!} + \dots, \quad (-\infty < x < \infty). \blacktriangleright$$

23.4. Современная стоимость денег, дисконтирование

Теория сходимости рядов имеет свое прямое приложение в финансовом корпоративном управлении. Приведем лишь несколько примеров из этой области знаний.

Рассмотрим сначала задачу, обратную той, что рассматривалась в подпараграфе 2.1 главы 2. Пусть требуется накопить через год определенную сумму денег FV . Банк принимает вклады по ставке i . Какую сумму надо иметь сегодня для того, чтобы при помещении ее в банк по ставке i иметь через год заданную сумму FV . Ответ на этот вопрос дает соотношение (2.1), переписанное в виде:

$$PV = \frac{FV}{1+i}. \quad (4.135)$$

Если бы требовалось накопить нужную сумму FV не через один год, а через n лет, то согласно (4.9) (параграф 5.2 главы 4),

$$PV = \frac{FV}{(1+i)^n}. \quad (4.136)$$

Соотношения (4.135), (4.136) решают поставленную задачу, то есть позволяют определить *современную*, или *текущую*, стоимость денег, исходя из будущей стоимости и сложной процентной ставки.

ОПРЕДЕЛЕНИЕ. Процесс приведения будущей суммы денег к современной стоимости называется **дисконтированием**.

Коэффициент, входящий в (4.136):

$$\frac{1}{(1+i)^n}, \quad (4.137)$$

является обратным коэффициенту наращения и называется *коэффициентом дисконтирования*. В задачах о дисконтировании процентную ставку i принято называть *ставкой дисконтирования*. Другие названия ставки дисконтирования — *стоимость привлечения капитала*, *пороговая доходность*, *ставка альтернативного капитала*, *ставка альтернативного вложения*, или *ставка альтернативной доходности*.

Для того чтобы расшифровать последнее название (ставка альтернативного капитала, ставка альтернативного вложения или доходности), рассмотрим простой пример.

Вы собираетесь инвестировать средства в определенный проект, который спустя n лет принесет доход, равный FV . Какую сумму денег следует вложить в данный проект? Для того чтобы ответить на этот вопрос, следует сравнить предлагаемый проект с другими альтернативными вложениями. Пусть i — средняя рыночная ставка доходности (ставка альтернативного вложения). Для того чтобы получить такую же сумму FV через n лет при осуществлении альтернативного проекта, сегодня следует вложить сумму PV , определяемую соотношением (4.136). Следовательно, инвестировать в предлагаемый проект следует сумму, не превышающую

$$\frac{FV}{(1+i)^n}.$$

Данную сумму называют *современной*, или *рыночной*, *стоимостью инвестиционного проекта*.

Приведенные примеры иллюстрируют, почему ставку дисконтирования называют ставкой альтернативного вложения или ставкой альтернативной доходности.

Дисконтирование — важная процедура при проведении финансовых расчетов. Метод дисконтирования широко используется для определения современной рыночной стоимости объекта инвестиций, в частности для определения текущей стоимости ценных бумаг.

Процесс дисконтирования позволяет также сравнивать различные доходы, полученные в разное время, путем приведения стоимости этих будущих потоков к настоящему моменту.

▷ ПРИМЕР 4.63. Какую сумму нужно поместить в банк для того, чтобы через 6 лет накопить сумму 200 000 руб.? Депозитная процентная ставка банка равна 25%.

Р е ш е н и е. В данном случае депозитная ставка банка выбрана в качестве ставки дисконтирования. Согласно (4.34) имеем

$$FV = \frac{200\,000}{1,25^6} = 52\,428,80 \text{ руб.}$$

То есть, для того чтобы через 6 лет накопить 200 000 руб. при ставке 25%, следует поместить на счет 52 428,80 руб. ►

▷ ПРИМЕР 4.64. По векселю через 3 месяца должна быть выплачена сумма 350 000 руб. Найти текущую стоимость векселя, если ставка дисконтирования выбрана 28,5 %.

Р е ш е н и е. Считаем $FV = 350\,000$ руб., $n = 0,25$ лет. Рыночная стоимость векселя определяется с помощью дисконтирования:

$$PV = \frac{350\,000}{(1 + 0,285)^{0,25}} = 328\,732,21 \text{ руб.} \blacktriangleright$$

Текущая стоимость аннуитета. Рассмотрим теперь такое важное понятие финансового анализа, как *аннуитет*.

ОПРЕДЕЛЕНИЕ. *Аннуитетом*, или *рентой*, называется постоянный доход, получаемый через равные промежутки времени.

Примерами аннуитета являются: доход, приносимый облигацией с постоянным купоном без погашения, дивиденды по привилегированным акциям, доход, приносимый сданной в аренду недвижимостью. Доходы, получаемые в разные моменты времени, имеют разную приведенную стоимость. Современная стоимость аннуитета, таким образом, складывается из современных стоимостей всех будущих доходов, представляя собой ряд вида:

$$PV = \sum_{k=1}^n \frac{PMT}{(1+i)^k}. \quad (4.138)$$

Здесь PV — современная стоимость аннуитета, PMT — регулярный ежегодный доход, n — количество лет, в течение которых поступали доходы, i — ставка дисконтирования. Этот ряд сходится (например, по признаку Даламбера). Просуммировав геометрическую прогрессию в правой стороне (4.138), находим:

$$PV = PMT \cdot \frac{1 - (1+i)^{-n}}{i}. \quad (4.139)$$

Коэффициент, входящий в правую часть последнего соотношения, равен

$$\frac{1 - (1+i)^{-n}}{i}. \quad (4.140)$$

Это коэффициент дисконтирования аннуитета.

Соотношение (4.139) определяет стоимость аннуитета в том случае, когда постоянные доходы поступают один раз в конце года. Иначе, можно утверждать, что формула (4.139) определяет рыночную стоимость объекта, приносящего ежегодный постоянный доход.

Если постоянные выплаты PMT происходят несколько раз в году (каждый раз в конце периода), например m раз в году, то можно записать

$$PV = m \cdot PMT \cdot \frac{1 - \left(1 + \frac{j}{m}\right)^{-n \cdot m}}{\frac{j}{m}}, \quad (4.141)$$

где j — номинальная процентная ставка при условии начисления процентов m раз в году, n — количество лет, пока происходят выплаты; всего за n лет будет произведено nm выплат.

▷ ПРИМЕР 4.65. Согласно долговой бумаге на протяжении 5 лет будут производиться ежегодные выплаты в размере 1000 руб. Какова текущая стоимость долговой бумаги, если ставка дисконтирования выбрана 19,25%?

Р е ш е н и е. $PV = 1000 \cdot \frac{1 - (1,1925)^{-5}}{0,1925} = 3040,65$ руб. ►

▷ ПРИМЕР 4.66. В условиях предыдущего примера считать, что выплаты происходят ежеквартально, то есть по 250 руб. каждые три месяца. Доход от ценной бумаги поступает в течение 5 лет. Ставка дисконтирования (номинальная при ежеквартальном начислении процентов $m = 4$) равна $j = 18\%$ ¹. Какова текущая стоимость ценной бумаги?

Р е ш е н и е. Имеем: $PMT = 250$, $j = 0,18$, $n = 5$, $m = 4$.

$$PV = 4 \cdot 250 \cdot \frac{1 - (1 + \frac{0,18}{4})^{-20}}{0,18} = 3251,98 \text{ руб.} \blacktriangleright$$

Видно, что стоимость ценной бумаги выше, чем в предыдущем примере. Это произошло из-за того, что выплаты приблизились к сегодняшнему дню.

Финансовые вычисления по облигациям. Облигации относятся к долговым ценным бумагам. По существу, эмиссия облигаций есть способ получения займа — эмитент выступает в роли заемщика, так как обязуется выплачивать определенный доход по облигациям, а покупатель облигации выступает в роли кредитора. Облигации, как правило, являются ценными бумагами с фиксированным доходом в отличие от акций, когда доход определяется

¹ Номинальная ставка 18% соответствует эффективной годовой ставке 19,25% предыдущего примера.

прибылью предприятия. Доход по облигациям, как правило, бывает меньше, чем по другим ценным бумагам, в то же время облигации считаются более надежным инструментом рынка ценных бумаг. Ведущие рейтинговые агентства, такие как Standard&Poors и Moody's, присваивают облигациям рейтинги. Чем выше рейтинг облигации, тем меньше ее доходность.

По облигациям выплачивается периодически купонный доход, а в конце срока происходит погашение номинала.

Основные параметры облигации — дата покупки облигации, дата погашения, номинальная цена, или номинал облигации, цена погашения, если она отличается от номинала (такая ситуация бывает, как правило, в случае нескольких дат погашения), годовой купонный доход, купонная процентная ставка, количество выплат купонов в году.

Облигация может иметь две или больше даты погашения. В этом случае, как правило, ставка купонного дохода увеличивается к каждой следующей дате погашения. При этом возможны следующие варианты: в первом случае владелец облигации сам выбирает, когда погасить облигацию, во втором случае эмитент оставляет за собой право погасить облигацию в любой из указанных сроков.

Возможна также ситуация, когда эмитент имеет право досрочного выкупа облигаций. Рейтинг таких облигаций ниже, чем у облигаций с запретом досрочного выкупа, так как высока неопределенность для инвестора.

Введем следующие обозначения основных параметров облигаций, которые понадобятся нам в дальнейшем:

N — номинал облигации, выплачивается при погашении;

P — рыночная цена;

$K = \frac{P}{N} \cdot 100$ — курс облигации, он определяет текущую стоимость облигации в процентах от номинала;

g — годовая купонная процентная ставка в процентах или десятичных долях;

$C = g \cdot N$ — *годовой купонный доход* в рублях, он определяет суммарный годовой доход, выплаченный по купонам;

$$i_t = \frac{C}{P} = \frac{g \cdot N}{P} = \frac{g}{K} \cdot 100 \text{ — текущая доходность (для аннуитетов}$$

совпадает с полной доходностью облигации);

i — *полная доходность* облигации за период владения ею; если владелец облигации держит ее вплоть до погашения в конце срока, то величина показывает *доходность к погашению*. В зависимости от задачи буквой i будем еще обозначать ставку дисконтирования.

Если облигация куплена по цене, равной номинальной цене, то говорят, что такая облигация куплена *по номиналу*, если облигация куплена по цене ниже номинала, то говорят, что облигация куплена с *дисконтом*, если по цене выше номинала — с *премией* (последнее не означает, что доход по такой облигации не может быть получен).

Доход, полученный за все время владения облигацией, складывается из купонных выплат и цены погашения облигации, выплачиваемой в конце срока владения (как правило, эта цена совпадает с номинальной ценой).

Обозначим через $C_1, C_2, C_3, \dots, C_n$ купонные доходы, полученные владельцем в течение времени владения облигацией. В конце срока происходит погашение облигации по номиналу N . Сюда относятся выплаты по купонам и цена погашения облигации. Тогда современная (рыночная) стоимость облигации P представляет собой ряд и равна сумме всех дисконтированных доходов:

$$P = \frac{C_1}{1+i} + \frac{C_2}{(1+i)^2} + \frac{C_3}{(1+i)^3} + \dots + \frac{C_n + N}{(1+i)^n}, \quad (4.142)$$

где i — доходность облигации к погашению.

Соотношение (4.142) связывает рыночную цену облигации с доходностью к погашению или со ставкой дисконтирования. Если будущие доходы известны, фиксированы, то соотношение (4.142) позволяет решать две основные задачи: а) определять цену облигации, если известна доходность (ставка дисконтирования); б) определять доходность облигации, если известна цена облигации. Очевидно, что эти две задачи являются обратными друг другу.

Контрольные вопросы

1. Какая последовательность называется монотонной, ограниченной, сходящейся (расходящейся)?
2. В чем суть теорем о пределах функций?
3. Какую функцию называют бесконечно малой, непрерывной, разрывной?
4. Каковы основные правила и формулы дифференцирования функций?
5. Пересекаются ли функции предельных и средних издержек?
6. Каковы основные правила и приемы, используемые при вычислениях неопределенных интегралов?
7. Зачем будущему экономисту нужно умение считать определенные интегралы? Приведите пример их использования.
8. Какие дифференциальные уравнения называют линейными, однородными, первого и второго порядков?
9. Перечислите основные признаки сходимости знакопеременного ряда.
10. Какой ряд называют абсолютно сходящимся?

Упражнения

Доказать, используя определение предела, что:

$$1. \lim_{n \rightarrow \infty} \frac{2n-1}{n+3} = 2. \quad 2. \lim_{x \rightarrow 5} (3x-4) = 11. \quad 3. \lim_{x \rightarrow 2} (x^2-3) = 1.$$

Найти предел:

$$4. \lim_{x \rightarrow 8} \frac{2x-7}{x-8}. \quad 5. \lim_{x \rightarrow 5} (x-5) \sin \frac{1}{x-5}.$$

$$6. \lim_{x \rightarrow \frac{1}{2}} \frac{5x-2x^2-2}{2x-1}. \quad 7. \lim_{x \rightarrow 0} \frac{\sqrt{1+x^2}-1}{\sqrt{x^2+16}-4}.$$

$$8. \lim_{x \rightarrow 1} \frac{x^2 - \sqrt{x}}{\sqrt{x} - 1}. \quad 9. \lim_{x \rightarrow \infty} \frac{1+x-3x^3}{1+x^2+3x^3}.$$

$$10. \lim_{x \rightarrow \infty} \frac{\sqrt{x^2 + 1} + \sqrt{x}}{\sqrt[4]{x^3 + x} - x}.$$

$$11. \lim_{x \rightarrow \infty} \frac{2^x - 3^x}{2^x + 3^x}.$$

$$12. \lim_{x \rightarrow \infty} x(\sqrt{x^2 + 1} - x).$$

$$13. \lim_{x \rightarrow \infty} \left(\sqrt[3]{(x+1)^2} - \sqrt[3]{(x-1)^2} \right)$$

$$14. \lim_{x \rightarrow 0} \frac{1 - \cos 4x}{x^2}.$$

$$15. \lim_{x \rightarrow \infty} (x-5) \sin \left(\frac{1}{x-5} \right)$$

$$16. \lim_{x \rightarrow \pi/6} \frac{\sin(x - \pi/6)}{\sqrt{3}/2 - \cos x}.$$

$$17. \lim_{x \rightarrow 0} \frac{\sin 4x}{\sqrt{x+1} - 1}.$$

$$18. \lim_{x \rightarrow \pi/2} \left(2x \operatorname{tg} x - \frac{\pi}{\cos x} \right)$$

$$19. \lim_{x \rightarrow \infty} \left(1 + \frac{2}{x} \right)^x.$$

$$20. \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x^2} \right)^{5x}.$$

$$21. \lim_{x \rightarrow \infty} \left(\frac{x^2 + 1}{x^2 - 3} \right)^{x^3 - 5}.$$

$$22. \lim_{x \rightarrow 0} \left(\frac{1}{x+1} \right)^{\frac{1}{2x}}.$$

$$23. \lim_{x \rightarrow 0} \left(1 + \operatorname{tg}^2 \sqrt{x} \right)^{\frac{1}{2x}}.$$

24. Первоначальный вклад, положенный в банк под 10% годовых, составил 60 тыс. руб. Найти размер вклада через 5 лет при начислении процентов:

а) ежегодном; б) поквартальном; в) непрерывном.

25. Пусть темп инфляции составляет 3% в день. На сколько изменится первоначальная сумма через 3 месяца?

Найти производную функции:

$$26. y = \left(\frac{x^2 - 1}{x^2 + 1} \right)^4.$$

$$27. y = x^4(8 \ln^2 x - 4 \ln x + 1).$$

$$28. y = \sqrt[3]{x}(e^{3x} - 5).$$

$$29. y = \sqrt[4]{1 + e^{4x}} + \sqrt{5}.$$

$$30. y = \ln \sqrt[3]{\left(\frac{1-3x}{1+3x}\right)^2}. \quad 31. y = \ln \frac{x(1+x^2)}{\sqrt{1-x^2}}.$$

$$32. y = 3x \ln(1-x^2). \quad 33. y = x^3 \ln 2x.$$

$$34. y = \sqrt[3]{\frac{1-e^{4x}}{e^{4x}}}. \quad 35. y = (xe^{2x} + 3)^5.$$

$$36. y = (x^2 - 1) \ln \sqrt{\frac{1-x}{1+x}}. \quad 37. y = \sin(x^2 + 2^x).$$

$$38. y = 4e^{\sqrt{\ln x}}(1 - \sqrt{\ln x}) \quad 39. y = \frac{\ln \cos x}{\cos x}.$$

$$40. y = \cos^2 x + \ln \operatorname{tg} \frac{x}{2}. \quad 41. y = \ln(\sqrt{1+e^x} - 1) - \ln(\sqrt{1+e^x} + 1).$$

Исследовать функцию и построить ее график:

$$42. y = x^2 + x. \quad 43. y = x^2 + \frac{1}{x^2}.$$

$$44. y = \frac{(x-1)^3}{(x+1)^2}. \quad 45. y = x^3 - 12x^2 + 36x.$$

$$46. y = x + \frac{27}{x^3}.$$

47. Найти все частные производные функции:

a) $z = x^2 \sin^2 y$; б) $z = x^{y^2}$; в) $u = e^{x^2 + y^2 + z^2}$.

48. Найти полный дифференциал от функции:

a) $z = x^2 + xy^2 + \sin y$; б) $z = \ln(xy)$.

49. Найти полный дифференциал dz функции:

a) $z = x^2 + y^2$, если $x = \ln u + \frac{1}{v} + w^4$, $y = \frac{1}{u} + vw$, $u = t^2$, $v = \frac{1}{t}$,

$$w = 2t.$$

б) $z = xe^y$, если $x = 3u^2 + 2v - w^3$, $y = au^2 + bv^3 + cw^4$, $u = \ln t$, $v = \frac{1}{t}$,

$$w = t.$$

50. Найти $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, если $z = e^{u-2v}$, $u = \sin x$, $v = x^3 + y^2$.

51. Найти все частные производные функции $z = \arcsin(u + v)$, если $u = \sin x \cos \alpha$, $v = \cos x \sin \alpha$.

Вычислить интеграл:

52. $\int x \operatorname{arctg} x \, dx$.

53. $\int \arcsin x \, dx$.

54. $\int \frac{\arcsin x}{\sqrt{1+x}} \, dx$.

55. $\int \operatorname{arctg} \sqrt{7x-1} \, dx$.

56. $\int \frac{\operatorname{arctg} x}{x^2} \, dx$.

57. $\int \frac{\operatorname{arctg} \sqrt{x}}{\sqrt{x}} \, dx$.

58. $\int \ln x \, dx$.

59. $\int x \ln x \, dx$.

60. $\int x \ln(3x+2) \, dx$.

61. $\int (x^2 + 3x + 2) \ln x \, dx$.

62. $\int (4x^3 + 6x - 7) \ln x \, dx$.

63. $\int \ln(\sqrt{1-x} + \sqrt{1+x}) \, dx$.

64. $\int \frac{x-4}{(x-2)(x-3)} \, dx$.

65. $\int \frac{2x+3}{(x-2)(x+5)} \, dx$.

66. $\int \frac{3x^2 + 2x - 3}{x(x-2)(x+5)} \, dx$.

67. $\int \frac{2x+3}{(x-2)^3} \, dx$.

68. $\int \frac{dx}{(x+1)(x-2)}$.

69. На складе хранится некоторая продукция, пользующаяся равномерным спросом. Ежедневно со склада извлекается 5 тонн продукции, плата за хранение 1 тонны продукции в день — 5 тыс. руб., плата за доставку одной партии — 98 тыс. руб. Найти оптимальный план поставок.

70. Найти частный интеграл дифференциального уравнения при заданных начальных условиях: $y^2 + x^2 y' = 0$, $y(-1) = 1$.

71. Найти общее решение уравнения $y - y' = e^x$.
 72. Найти частное решения уравнения при указанных начальных условиях: $y'' - 5y' + 4y = 0$, $y(0) = 1$, $y'(0) = 1$.
 73. Найти общее решение уравнения: $y'' + y = 6 \sin 2x$.
 74. Написать первые пять членов ряда и проверить, выполняется ли для него необходимый признак сходимости: $\sum_{n=1}^{\infty} \frac{n}{\sqrt{n^2 + 1}}$.

75. Исследовать по признаку Даламбера сходимость ряда:

$$\sum_{n=1}^{\infty} \frac{n(n+1)}{3^n}.$$

76. Найти область сходимости степенного ряда:

$$\sum_{n=1}^{\infty} \frac{x^n}{n^2}.$$

Ответы к упражнениям

4. ∞ . 5. 0. 6. 1,5. 7. 4. 8. 3. 9. -1. 10. -1. 11. -1. 12. 0,5. 13. 0. 14. 8.
 15. 1. 16. 2. 17. 8. 18. -2. 19. e^2 . 20. 1. 21. ∞ . 22. $1/\sqrt{e}$. 23. $\sqrt{e}$.

26. $\frac{16x(x^2 - 1)^3}{(x^2 + 1)^5}$. 27. $32x^3 \ln^2 x$. 28. $\frac{e^{3x}(9x + 1) - 5}{3\sqrt[3]{x^2}}$. 29. $\frac{e^{4x}}{\sqrt[4]{(1 + e^{4x})^3}}$.

30. $\frac{4}{9x^2 - 1}$. 31. $\frac{2x^4 - 3x^2 - 1}{x(x^4 - 1)}$. 32. $3 \left(\ln(1 - x^2) - \frac{2x^2}{1 - x^2} \right)$

33. $x^2 \ln x (3 \ln x + 2)$. 34. $\frac{-4}{3\sqrt[3]{(1 - e^{4x})^2} e^{4x}}$. 35. $5e^{2x}(xe^{2x} + 3)^4(2x + 1)$.

36. $x \ln \frac{1 - x}{1 + x} + 1$. 37. $(2x + 2^x \ln 2) \cos(x^2 + 2^x)$. 38. $-2e^{\sqrt{\ln x}}$.

39. $\frac{\sin x (\ln \cos x - 1)}{\cos^2 x}$. 40. $\frac{1 - \sin x \cdot \sin 2x}{\sin x}$. 41. $\frac{1}{\sqrt{1 + e^x}}$.

42. $y_{\min}(-1/2) = -1/4$, функция убывает на $(-\infty; -1/2)$, возрастает на $(-1/2; +\infty)$, выпукла вниз на $(-\infty; +\infty)$.

43. Асимптота $x = 0$; $y_{\min}(-1) = 2$, функция убывает на $(-\infty; -1)$ и на $(0; 1)$, возрастает на $(-1; 0)$ и на $(1; +\infty)$. Выпукла вниз на $(-\infty; 0)$ и $(0; +\infty)$. Точек перегиба нет.

44. Асимптоты $x = -1$, $y = x - 5$; $y_{\min}(-5) = -27/2$, функция возрастает на $(-\infty; -5)$ и на $(-1; +\infty)$, убывает на $(-5; -1)$. Выпукла вверх на $(-\infty; -1)$ и на $(-1; 1)$, выпукла вниз на $(1; +\infty)$. Точка перегиба $(1; 0)$.

45. $y_{\max}(2) = 32$, $y_{\max}(6) = 0$, функция возрастает на $(-\infty; 2)$ и на $(6; +\infty)$, убывает на $(2; 6)$. Выпукла вверх на $(-\infty; 4)$, выпукла вниз на $(4; +\infty)$. Точка перегиба $(4; 16)$.

46. Асимптоты $x = 0$, $y = x$; $y_{\max}(3) = 4$, $y_{\max}(-3) = -4$, функция возрастает на $(-\infty; -3)$ и на $(3; +\infty)$, убывает на $(-3; 0)$ и $(0; 3)$. Выпукла вверх на $(-\infty; 0)$, выпукла вниз на $(0; +\infty)$. Точек перегиба нет.

$$\mathbf{52.} \frac{x^2+1}{2} \operatorname{arctg} x - \frac{x}{2} + C. \quad \mathbf{53.} x \arcsin x + \sqrt{1-x^2} + C.$$

$$\mathbf{54.} 2\sqrt{1+x} \arcsin x + 4\sqrt{1-x} + C.$$

$$\mathbf{55.} x \operatorname{arctg} \sqrt{7x-1} - \frac{1}{7} \sqrt{7x-1} + C.$$

$$\mathbf{56.} \frac{\operatorname{arctg} x}{x} - \frac{1}{2} \ln \left(1 + \frac{1}{x^2} \right) + C. \quad \mathbf{57.} 2\sqrt{x} \arcsin \sqrt{x} + 2\sqrt{1-x} + C.$$

$$\mathbf{58.} x(\ln x - 1) + C. \quad \mathbf{59.} \frac{x^2}{2} \ln x - \frac{x^2}{4} + C.$$

$$\mathbf{60.} \left(\frac{x^2}{2} - \frac{2}{9} \right) \ln(3x+2) - \frac{x^2}{4} + \frac{x}{3} + C.$$

$$\mathbf{61.} \left(\frac{x^3}{3} + \frac{3}{2}x^2 + 2x \right) \ln x - \frac{x^3}{9} - \frac{3x^2}{4} - 2x + C.$$

$$62. (x^4 + 3x^2 - 7x) \ln x - \left(\frac{x^4}{4} + \frac{3x^2}{2} - 7x \right) + C.$$

$$63. x \ln(\sqrt{1-x} + \sqrt{1+x}) - \frac{1}{2}x + \frac{1}{2} \arcsin x + C.$$

$$64. 2 \ln|x-2| - \ln|x-3| + C. \quad 65. \ln|x-2| + \ln|x+5| + C.$$

$$66. 3 \ln|x| + \ln|x-1| - \ln|x+1| + C. \quad 67. -\frac{7}{2(x-2)^2} - \frac{2}{x-2} + C.$$

$$68. \frac{1}{3} \ln \left| \frac{x-2}{x+1} \right| + C. \quad 70. x + y = 0. \quad 71. y = (x + C)e^x. \quad 72. y = e^x.$$

$$73. y = C_1 \cos x + C_2 \sin x - 2 \sin x.$$

$$74. \frac{1}{\sqrt{2}}, \frac{2}{\sqrt{5}}, \frac{3}{\sqrt{10}}, \frac{4}{\sqrt{17}}, \frac{5}{\sqrt{26}}; \text{ не выполняется.}$$

$$75. \text{Сходится.} \quad 76. [-1; 1].$$

Г л а в а 5

ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКУЮ СТАТИСТИКУ

§1. Основные понятия теории вероятностей

1.1. Опыт и события

Одним из основных понятий теории вероятностей является опыт. Под опытом понимается выполнение комплекса условий, в результате которого происходят или не происходят определенные события (факты).

Простейшие неразложимые результаты опыта называются *элементарными событиями* (исходами) (ω_i), а вся совокупность элементарных событий называется *пространством элементарных событий* $\Omega = \{\omega_i\}$. С каждым опытом связано свое пространство элементарных событий Ω .

▷ПРИМЕР 5.1. Брошена монета. Найти пространство элементарных событий.

Р е ш е н и е. Выделим возможные элементарные исходы: “выпал герб” — ω_1 , “выпала решка” — ω_2 . Пространство элементарных событий $\Omega = \{\omega_1, \omega_2\}$. ►

▷ПРИМЕР 5.2. Найти пространство элементарных событий, если игральная кость подбрасывается один раз.

Р е ш е н и е. Выделим возможные элементарные события: ω_1 — появление 1, ω_2 — появление 2, ω_3 — появление 3, ω_4 — появление 4, ω_5 — появление 5, ω_6 — появление 6. Пространство элементарных событий $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. ►

Любое конечное, или счетное¹, подмножество Ω называется *событием*. Различают три типа событий:

¹ См. подпараграф 1.1 главы 4.

- 1) достоверные (Ω);
- 2) случайные;
- 3) невозможные ($\emptyset$ или $\overline{\Omega}$).

События обычно обозначают первыми прописными буквами латинского алфавита: $A, B, C, \dots$. Например, в примере 5.2 событие $A = \{\omega_2, \omega_4, \omega_6\}$ — появление четного числа очков при подбрасывании игральной кости.

Любое событие назовем *случайным*, если при осуществлении определенной совокупности условий E оно может либо произойти, либо не произойти. В дальнейшем, вместо того чтобы говорить: “совокупность условий E осуществлена”, будем говорить кратко: “произведен опыт”. Таким образом, *событие* будет рассматриваться как *результат опыта* или *испытания*.

Например, в примере 5.2 подбрасывание игральной кости есть испытание. Появление грани с определенным количеством очков — событие.

ОПРЕДЕЛЕНИЕ. События A и B *несовместны*, если в результате одного опыта они не могут происходить одновременно, в противном случае — *совместны*.

Приведем примеры несовместных событий.

1. Из урны с шарами наудачу извлечен шар. Появление синего шара исключает появление красного. События “появился синий шар” и “появился красный шар” — несовместные.

2. В примере 5.1 появление “герба” исключает появление “решки”. События “появился герб” и “появилась решка” — несовместные.

ОПРЕДЕЛЕНИЕ. События $A_1, \dots, A_n$ образуют *полную группу*, если в результате опыта кроме этих событий ничего другого произойти не может.

Например, торговец акциями произвел сделку. Обязательно произойдет одно из следующих трех событий: получение прибыли, неполучение прибыли, убыток. Эти три несовместных события образуют полную группу. Другими словами, появление хотя бы одного из событий полной группы есть достоверное событие.

ОПРЕДЕЛЕНИЕ. Элементы последовательности событий $A_1, A_2, \dots, A_n$ **попарно несовместны**, если любые два из них несовместны.

Например, при подбрасывании игральной кости никакие два элементарных исхода (появление цифр 1, 2, 3, 4, 5, 6) не могут произойти одновременно.

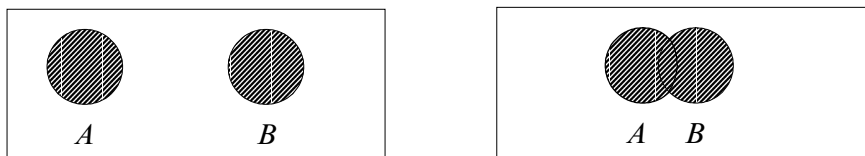
ОПРЕДЕЛЕНИЕ. Несколько событий **равновозможны**, если ни одно из них не имеет объективного преимущества перед другими.

Например, элементарные исходы при подбрасывании монеты, игральной кости равновозможны.

Обычно Ω изображают на плоскости в виде некоторой области, а ω_i в виде точек этой области, устанавливая таким образом соответствие между событиями и точечными множествами.

Над событиями вводятся операции, совпадающие с операциями над множествами: сумма, произведение, отрицание.

ОПРЕДЕЛЕНИЕ. **Суммой** событий A и B называется такое третье событие $A + B$ (или $A \cup B$), которое заключается в наступлении хотя бы одного из событий: или A , или B (рис. 5.1).

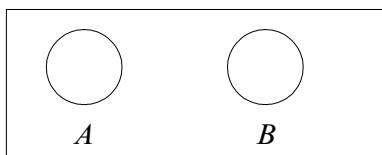


События несовместны

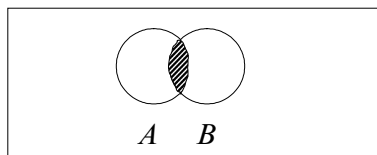
События совместны

Рис. 5.1

ОПРЕДЕЛЕНИЕ. **Произведением** двух событий A и B называется такое третье событие AB (или $A \cap B$), которое заключается в наступлении событий A и B одновременно. Если события A и B несовместны, то $A \cdot B = \emptyset$ (рис. 5.2).



События несовместны



События совместны

Рис. 5.2

ОПРЕДЕЛЕНИЕ. *Отрицанием* события A называется событие $\bar{A}$ (не A), заключающееся в ненаступлении события A . При этом если в результате опыта может произойти событие A , то может произойти и обратное ему событие $\bar{A}$ (затемненная область на рис. 5.3).

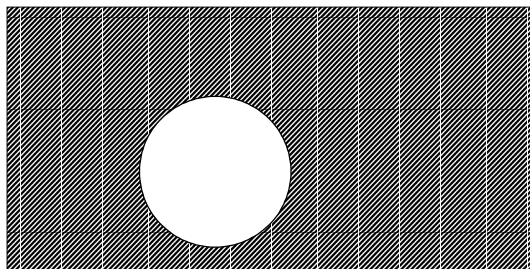


Рис. 5.3

Если наступление события A приводит к наступлению события B и наоборот (наступление B влечет наступление A), то события A и B *равны* ($A = B$).

ОПРЕДЕЛЕНИЕ. Пусть S — множество всех подмножеств Ω , для которого выполняются следующие свойства:

- 1) если $A \in S$ и $B \in S$, то $A + B = A \cup B \in S$;
- 2) если $A \in S$ и $B \in S$, то $A \cdot B = A \cap B \in S$;
- 3) если $A \in S$, то $\bar{A} \in S$.

Тогда множество S называется *алгеброй событий*.

1.2. Определение вероятности события

Вероятность — одно из основных понятий теории вероятностей. Существует несколько подходов к определению этого понятия. Рассмотрим основные из них.

Классическое определение вероятности. Пусть события

$$A_1, A_2, \dots, A_n \in S \quad (5.1)$$

образуют множество элементарных событий, n — число всех возможных элементарных исходов испытания.

Тогда события из (5.1), которые приводят к наступлению события A , называются *благоприятствующими исходами* для события A , $m(A) = m$ — число благоприятствующих исходов.

ОПРЕДЕЛЕНИЕ. *Вероятностью* события A называют отношение числа благоприятствующих этому событию исходов к общему числу всех равновозможных несовместных элементарных исходов, образующих полную группу, т.е.

$$P(A) = m/n.$$

Здесь предполагается, что элементарные исходы несовместны, равновозможны и образуют полную группу.

В примере 5.1 вероятность появления герба

$$P(A) = \frac{1}{2}.$$

В примере 5.2 вероятность появления числа больше четырех

$$P(A) = \frac{2}{6} = \frac{1}{3}.$$

Из классического определения следуют **свойства вероятности**:

1. *Вероятность достоверного события равна единице* $P(\Omega) = 1$.

Действительно, если событие достоверно, то каждый элементарный исход опыта благоприятствует событию. В этом случае $m = n$, следовательно,

$$P(A) = m/n = n/n = 1.$$

2. *Вероятность невозможного события равна нулю.* $P(\bar{\Omega}) = 0$

Действительно, если событие невозможно, то ни один из элемен-

тарных исходов опыта не благоприятствует событию. В этом случае $m = 0$, следовательно,

$$P(A) = m/n = 0/n = 0.$$

3. Вероятность случайного события есть положительное число, заключенное между нулем и единицей.

Действительно, случайному событию благоприятствует лишь часть из общего числа элементарных исходов испытания. В этом случае $0 < m < n$, значит, $0 < m/n < 1$, следовательно,

$$0 < P(A) < 1.$$

Итак, вероятность любого события удовлетворяет двойному неравенству

$$0 \leq P(A) \leq 1, \text{ а } A + \bar{A} = \Omega \text{ — достоверное событие, поэтому } P(A) + P(\bar{A}) = 1, \text{ или } P(\bar{A}) = 1 - P(A).$$

Геометрическое определение вероятности.

ОПРЕДЕЛЕНИЕ. Геометрической вероятностью события A называется отношение меры области, благоприятствующей появлению события A , к мере всей области.

▷ **ПРИМЕР 5.3.** Найти вероятность того, что точка, случайным образом брошенная в квадрат $ABCD$ со стороной 4, попадет в квадрат $A_1B_1C_1D_1$ со стороной 3, находящийся внутри $ABCD$ (рис. 5.4).

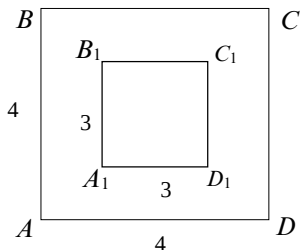


Рис. 5.4

Решение. Вероятность события определяется как отношение меры части области (в данном случае площади), благоприятствующей событию A — $S_{A_1B_1C_1D_1}$, к мере всей области — S_{ABCD} .

$$P(A) = \frac{S_{A_1B_1C_1D_1}}{S_{ABCD}} = \frac{9}{16}. \blacktriangleright$$

1.3. Комбинаторика

Комбинаторика (комбинаторный анализ) — раздел математики, посвященный решению задач выбора и расположения элементов некоторого, обычно конечного, множества в соответствии с заданными правилами. Рождение комбинаторики связано с работами Б. Паскаля и П. Ферма по поводу азартных игр, большой вклад внесли Лейбниц, Бернулли, Эйлер. В настоящее время интерес к комбинаторике связан с развитием математических методов в экономике, теории игр и их применением на практике.

Правило произведения. Пусть из некоторого конечного множества

1-й объект можно выбрать k_1 способами,
2-й объект — k_2 способами,
....., (5.2)

n -й объект — k_n способами.

Тогда произвольный набор перечисленных n объектов из данного множества можно выбрать $k_1 \cdot k_2 \cdot \dots \cdot k_n$ способами.

▷ ПРИМЕР 5.4. Сколько существует трехзначных чисел с разными цифрами?

Решение. В десятичной системе исчисления десять цифр: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. На первом месте может стоять любая из девяти цифр (кроме нуля). На втором месте — любая из оставшихся 9 цифр, кроме выбранной. На последнем месте — любая из оставшихся 8 цифр.

По правилу произведения $9 \cdot 9 \cdot 8 = 648$ трехзначных чисел имеют разные цифры. ►

Правило суммы. При выполнении условий (5.2), любой из объектов можно выбрать $k_1 + k_2 + k_3 + \dots + k_n$ способами.

▷ ПРИМЕР 5.5. Сколько существует способов выбора одного карандаша из коробки, содержащей 5 красных, 7 синих, 3 зеленых карандаша.

Решение. Один карандаш, по правилу суммы, можно выбрать $5 + 7 + 3 = 15$ способами. ►

Обычно в комбинаторике рассматривается идеализированный эксперимент по выбору наудачу m элементов из n . При этом элементы: а) не возвращаются обратно (схема выбора без возвращения); б) возвращаются обратно (схема выбора с возвращением). Рассмотрим обе эти схемы.

Схема выбора без возвращений.

ОПРЕДЕЛЕНИЕ. *Размещениями* называют комбинации, составленные из n различных элементов по m элементов и отличающиеся либо составом элементов, либо их порядком.

Число всех возможных размещений

$$A_n^m = n(n-1)(n-2)\cdots(n-m+1) = \frac{n!}{(n-m)!}, \quad (5.3)$$

где $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$, $1! = 1$, $0! = 1$.

▷ **ПРИМЕР 5.6.** В соревнованиях участвует 10 человек, трое из них займут 1, 2, 3 место. Сколько существует различных вариантов?

Решение. Число различных вариантов равно

$$A_{10}^3 = \frac{10!}{(10-3)!} = \frac{7! \cdot 8 \cdot 9 \cdot 10}{7!} = 8 \cdot 9 \cdot 10 = 720. \blacktriangleright$$

ОПРЕДЕЛЕНИЕ. *Перестановками* называют комбинации, состоящие из одних и тех же n различных элементов и отличающиеся только порядком их расположения.

Число всех возможных перестановок

$$P_n = n!, \quad (5.4)$$

где $n! = 1 \cdot 2 \cdot 3 \cdots n$.

▷ **ПРИМЕР 5.7.** Сколько существует способов расстановки 10 человек в одну линию?

Решение. Общее число способов расстановки определяется как число перестановок из 10 элементов и равно $P_{10} = 3\,628\,800$. ▶

ОПРЕДЕЛЕНИЕ. *Сочетаниями* называют комбинации, составленные из n различных элементов по m элементов, которые отличаются хотя бы одним элементом.

Число сочетаний

$$C_n^m = \frac{n!}{m!(n-m)!}. \quad (5.5)$$

Справедливы тождества:

$$C_n^k = C_n^{n-k}, \quad C_{n+1}^k = C_n^k + C_n^{k-1}, \quad C_n^0 = 1, \quad C_n^1 = n,$$

$$C_n^0 + C_n^1 + \dots + C_n^n = 2^n, \quad \sum_{k=0}^n (C_n^k)^2 = C_{2n}^n.$$

▷ **ПРИМЕР 5.8.** Сколько существует способов выбора трех облигаций из десяти?

Решение. В данном случае при выборе для нас важен только состав наборов по три облигации, порядок выбора роли не играет, поэтому число способов выбора подсчитаем по формуле сочетаний

$$C_{10}^3 = \frac{10!}{(10-3)!3!} = \frac{7! \cdot 8 \cdot 9 \cdot 10}{7! \cdot 1 \cdot 2 \cdot 3} = 120. \blacktriangleright$$

Схема выбора с возвращением. Если при выборе m элементов из n элементы возвращаются обратно и упорядочиваются, то говорят, что это *размещения с повторениями*.

Число размещений с повторениями:

$$\overline{A}_n^m = n^m. \quad (5.6)$$

▷ **ПРИМЕР 5.9.** В гостинице 10 комнат, каждая из которых может разместить четырех человек. Сколько существует вариантов размещения, прибывших четырех гостей?

Решение. Каждый следующий гость из 4 может быть помещен в любую из 10 комнат, поэтому общее число размещений по формуле размещений с повторениями равно

$$\overline{A}_{10}^4 = 10^4 = 10\,000. \blacktriangleright$$

Если при выборе m элементов из n элементы возвращаются обратно без последующего упорядочивания, то говорят, что это *сочетания с повторениями*. Число сочетаний с повторениями из n элементов по m определяется по формуле

$$C_{n+m-1}^m. \quad (5.7)$$

▷ ПРИМЕР 5.10. В магазине продается 10 видов тортов. Очередной покупатель выбил чек на три торта. Считая, что любой набор товаров равновозможен, определить число возможных заказов.

Решение. Число равновозможных заказов по формуле (5.7) равно

$$C_{10+3-1}^3 = C_{12}^3 = \frac{12!}{(12-3)!3!} = \frac{9! \cdot 10 \cdot 11 \cdot 12}{9!3!} = 220. \blacktriangleright$$

Схема упорядоченных разбиений. Пусть $k_1, k_2, \dots, k_r$ — целые числа, такие, что $k_1 + k_2 + \dots + k_r = n, k_i \geq 0 (i = 1, 2, \dots, r)$. Число способов, которыми генеральную совокупность из n элементов можно разделить на r упорядоченных частей (r подмножеств или r групп), из которых первая содержит k_1 элементов, вторая — k_2 элементов и r -я — k_r элементов, обозначается $C_n(k_1, k_2, \dots, k_r)$ и вычисляется по формуле

$$C_n(k_1, k_2, \dots, k_r) = \frac{n!}{k_1! k_2! \cdot \dots \cdot k_r!}. \quad (5.8)$$

Числа, которые определяются по формуле (5.8), называются *полиномиальными коэффициентами*.

▷ ПРИМЕР 5.11. Девять человек размещается в гостинице в четырехместный, трехместный и двухместный номера. Сколько существует способов их размещения

Решение. Число способов размещения по формуле (5.8) равно

$$C_9(4, 3, 2) = \frac{9!}{4!3!2!} = 1260. \blacktriangleright$$

1.4. Примеры вычисления вероятностей

▷ ПРИМЕР 5.12. Набирая номер телефона, абонент забыл последние 3 цифры и набрал их наудачу, помня, что они различны. Найти вероятность того, что набраны нужные цифры.

Р е ш е н и е. Событие A — номер набран верно

$$\tilde{N}(A) = \frac{m(A)}{n} = \frac{1}{720},$$

где $m(A) = 1$ — так как только один набор из 3-х цифр является нужным, а всего таких наборов

$$n = A_{10}^3 = \frac{10!}{(10-3)!} = \frac{7! \cdot 8 \cdot 9 \cdot 10}{7!} = 720. \blacktriangleright$$

▷ ПРИМЕР 5.13. Указать ошибку “решения” задачи: “Брошены две игральные кости. Найти вероятность того, что сумма выпавших очков равна 4 (событие A)”.

“Решение” (ошибочное). Всего возможны два исхода испытания: сумма выпавших очков равна 4, сумма выпавших очков не равна 4. Событию A благоприятствует один исход; общее число исходов равно двум. Следовательно, искомая вероятность

$$P(A) = 1/2.$$

Ошибка этого “решения” состоит в том, что рассматриваемые исходы не являются равновероятными.

Р е ш е н и е (правильное). Общее число равновероятных исходов испытания равно $6 \cdot 6 = 36$ (каждое число выпавших очков на одной кости может сочетаться со всеми числами очков другой кости). Среди этих исходов благоприятствуют событию A только три исхода: (1; 3), (3; 1), (2; 2) (в скобках указаны числа выпавших очков). Следовательно, искомая вероятность

$$P(A) = 3/36 = 1/12. \blacktriangleright$$

Мы постараемся в этой главе рассмотреть некоторые примеры из области корпоративного управления, сводящиеся к задачам теории вероятностей и математической статистики. Для этого нам потребуется ввести одно понятие.

Любое хозяйственное решение, принимаемое в рыночной экономике, представляет собой, по сути, *портфель* (набор, сумму) инвестиций (см. §3 главы 3). Например, семья, решая, как потратить тринадцатую зарплату (премию по итогам года одного из супругов), выбирает несколько инвестиций: часть, предназначен-

ную на покупку квартиры, положить на годовой депозит в банке; другую часть немедленно потратить на ремонт квартиры; третью часть дать в долг давно просящему об этом родственнику и т.д. Также и предприятие, формируя свой план работы на год, делит его на отдельные виды инвестиций: часть денег пойдет на строительство новых цехов, другая — в наем рабочей силы, третья — в банк на поддержание оборотных средств. И, наконец, классический инвестор, решая, где разместить свои средства, составляет свой *портфель* из различных инвестиций: государственных бумаг (ГКО), корпоративных облигаций, акций, деривативов.

▷ ПРИМЕР 5.14. В портфеле инвестора, склонного к риску, 15 различных облигаций, среди которых 10 облигаций с высокой степенью неисполнения обязательств по ним. Инвестор решает наугад исключить из портфеля 3 облигации. Найти вероятность того, что он исключил из портфеля только “бросовые” облигации.

Р е ш е н и е. Пусть событие A — 3 облигации с высокой степенью неисполнения обязательств по ним.

$$m(A) = C_{10}^3 = \frac{10!}{(10-3)! \cdot 3!} = \frac{7! \cdot 8 \cdot 9 \cdot 10}{7! \cdot 1 \cdot 2 \cdot 3} = 120 \text{ — число благоприятствующих исходов.}$$

яствующих исходов.

$$n = C_{15}^3 = \frac{15!}{(15-3)! \cdot 3!} = \frac{12! \cdot 13 \cdot 14 \cdot 15}{12! \cdot 1 \cdot 2 \cdot 3} = 455 \text{ — общее число возможных исходов.}$$

можных исходов.

$$\text{Имеем } \tilde{N}(A) = \frac{m(A)}{n} = \frac{120}{455} = \frac{24}{91}. \blacktriangleright$$

▷ ПРИМЕР 5.15. В коробке 5 красных, 3 зеленых, 2 синих карандаша. Наудачу без возвращения извлекают 3 карандаша. Найти вероятность следующих событий:

A — все извлеченные карандаши разного цвета,

B — все извлеченные карандаши одного цвета,

C — среди извлеченных карандашей 1 синий,

D — среди извлеченных карандашей в точности 2 одного цвета.

Р е ш е н и е. Всего в коробке $5 + 3 + 2 = 10$ карандашей.

1. По правилу произведения $m(A) = 5 \cdot 3 \cdot 2 = 30$ — число исходов, благоприятствующих наступлению события A ; общее число способов выбора из 10 карандашей 3 вычисляется как число сочетаний из 10 по 3:

$$n = C_{10}^3 = \frac{10!}{(10-3)! \cdot 3!} = \frac{7! \cdot 8 \cdot 9 \cdot 10}{7! \cdot 1 \cdot 2 \cdot 3} = 120,$$

отсюда $P(A) = \frac{30}{120} = \frac{1}{4}$.

2. Если все извлеченные карандаши одного цвета, то это либо 3 красных, либо 3 зеленых (3 синих не может быть, так как их в коробке всего 2).

Поэтому по правилу суммы

$$m(B) = C_5^3 + C_3^3 = \frac{5!}{(5-3)! \cdot 3!} + 1 = \frac{3! \cdot 4 \cdot 5}{2! \cdot 3} + 1 = 10 + 1 = 11.$$

Так как $n = 120$, следовательно, $P(B) = \frac{m(B)}{n} = \frac{11}{120}$.

3. Если из 10 карандашей два синего цвета, то один синий можно выбрать двумя способами, а два из восьми карандашей не синего цвета — 28 способами ($C_8^2 = 28$).

Отсюда, по правилу произведения $m(C) = 2 \cdot 28 = 56$,

$$P(C) = \frac{m(C)}{n} = \frac{56}{120} = \frac{7}{15}.$$

4. Событие D — 2 карандаша одного цвета произойдет, если произойдет одно из трех событий: 2 красных карандаша плюс 1 зеленый или синий, 2 зеленых карандаша плюс 1 красный или 1 синий, 2 синих карандаша плюс 1 красный или 1 зеленый.

По правилу произведения $C_5^2 \cdot C_5^1$ — число способов выбора 2 красных карандашей и 1 другого цвета; $C_3^2 \cdot C_7^1$ — число способов выбора 2-х зеленых карандашей и 1 другого цвета; $C_2^2 \cdot C_8^1$ — число способов выбора 2-х синих карандашей и 1 другого цвета.

По правилу суммы

$$m(D) = C_5^2 \cdot C_5^1 + C_3^2 \cdot C_7^1 + C_2^2 \cdot C_8^1 = 79.$$

$$\text{Следовательно, } P(D) = \frac{m(D)}{n} = \frac{79}{120}. \blacktriangleright$$

1.5. Относительная частота

ОПРЕДЕЛЕНИЕ. Отношение числа испытаний, в которых событие появилось, к общему числу фактически произведенных испытаний называется *относительной частотой* события.

Таким образом, относительная частота события A определяется формулой

$$W(A) = m/n,$$

где n — общее число испытаний, m — число появлений события.

Сопоставляя определения вероятности и относительной частоты, делаем вывод: определение вероятности не требует, чтобы испытания производились в действительности; определение же относительной частоты предполагает, что испытания были произведены фактически. Другими словами, *вероятность вычисляют до опыта, а относительную частоту — по опыту*.

▷ **ПРИМЕР 5.16.** Например, отдел технического контроля обнаружил 4 бракованных компакт-диска в партии из 83 случайно отобранных дисков. Найти относительную частоту появления бракованных компакт-дисков.

Решение. Относительная частота появления бракованных компакт-дисков $W(A) = 4/83$. ▶

Длительные наблюдения показали, что если в одинаковых условиях производят опыты, в каждом из которых число испытаний достаточно велико, то относительная частота обнаруживает **свойство устойчивости**. Это свойство состоит в том, что в различ-

ных опытах относительная частота изменяется мало (тем меньше, чем больше произведено испытаний), колеблясь около некоторого постоянного числа. Оказалось, что это постоянное число есть вероятность появления события.

Таким образом, если опытным путем установлена относительная частота, то полученное число можно принять за приближенное значение вероятности.

▷ПРИМЕР 5.17. По данным статистики, относительная частота рождения мальчиков за 1934 г. по месяцам характеризуется следующими числами (числа расположены в порядке следования месяцев начиная с января):

0,514; 0,511; 0,510; 0,529; 0,522; 0,518; 0,538; 0,516; 0,515; 0,509; 0,518; 0,527. Определить относительную частоту рождения мальчиков.

Р е ш е н и е. Относительная частота колеблется около числа 0,518, которое можно принять за приближенное значение вероятности рождения мальчиков. ►

§ 2. Основные теоремы теории вероятностей

2.1. Теорема сложения вероятностей

Теорема сложения вероятностей несовместных событий. Мы уже говорили в § 1 (см. рис. 5.1) о сумме двух событий A и B .

Допустим, что комиссией утверждается бизнес-план и A — утверждение при первом рассмотрении, B — утверждение при втором рассмотрении. Тогда $A + B$ — утверждение при первом рассмотрении, или утверждение при втором, или утверждение в обоих случаях.

В частности, если два события A и B — несовместные, то $A + B$ — событие, состоящее в появлении одного из этих событий (безразлично какого).

ОПРЕДЕЛЕНИЕ. *Суммой нескольких событий* называют событие, которое состоит в появлении хотя бы одного из этих событий.

Например, событие $A + B + C$ состоит в появлении одного из следующих событий: A ; B ; C ; A и B ; A и C ; B и C ; A , B и C .

Пусть события A и B — несовместные, причем вероятности этих событий известны. Как найти вероятность того, что наступит либо событие A , либо событие B ? Ответ на этот вопрос дает теорема сложения вероятностей.

□ **ТЕОРЕМА 5.1.** *Вероятность суммы двух несовместных событий A и B равна сумме вероятностей этих событий:*

$$P(A + B) = P(A) + P(B). \blacksquare$$

СЛЕДСТВИЕ. Пусть $A_1, A_2, \dots, A_n$ — попарно несовместные события. Тогда вероятность появления одного из них (безразлично какого) равна сумме вероятностей этих событий:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Полная группа событий.

□ **ТЕОРЕМА 5.2.** *Вероятность суммы попарно несовместных событий $A_1, A_2, \dots, A_n$, образующих полную группу, равна единице:*

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n) = 1. \blacksquare$$

▷ПРИМЕР 5.18. Вероятность падения курса акции, определенная на основе длительного исследования, равна 0,4, вероятность роста — 0,5. Найти вероятность того, что курс акции не изменится.

Р е ш е н и е. События “падение курса акции”, “рост курса акции”, “курс акции не изменился” образуют полную группу, поэтому сумма вероятностей этих событий равна единице:

$$0,4 + 0,5 + p = 1.$$

Отсюда искомая вероятность

$$p = 1 - 0,9 = 0,1. \blacktriangleright$$

Противоположные события. В §1 (см. рис. 5.3) мы ввели понятие отрицания события A .

ОПРЕДЕЛЕНИЕ. Если два единственно возможных события образуют полную группу, то они называются *противоположными*.

Если одно из двух противоположных событий обозначено через A , то другое мы условились обозначать $\bar{A}$.

Например, успех и неудача при реализации инвестиционного проекта — противоположные события. Если A — успех, то $\bar{A}$ — неудача.

□ **ТЕОРЕМА 5.3.** Если A и $\bar{A}$ — два противоположных события, то сумма вероятностей противоположных событий равна единице:

$$P(A) + P(\bar{A}) = 1.$$

Д о к а з а т е л ь с т в о. Согласно определению противоположные события образуют полную группу, а сумма вероятностей событий, образующих полную группу, равна единице (см. теорему 5.2). ■

З а м е ч а н и е. Если вероятность одного из двух противоположных событий обозначена через p , то вероятность другого события обозначают через q . Таким образом, в силу теоремы 5.3

$$p + q = 1.$$

▷ПРИМЕР 5.19. Вероятность того, что инвестиционный проект будет принят, $p = 0,7$. Найти вероятность того, что проект будет отклонен.

Решение. События “проект принят” и “проект отклонен” — противоположные, поэтому искомая вероятность

$$q = 1 - p = 1 - 0,7 = 0,3. \blacktriangleright$$

Замечание. При решении задач на отыскание вероятности события A часто выгодно сначала вычислить вероятность противоположного события $\bar{A}$, а затем найти искомую вероятность по формуле $P(A) = 1 - P(\bar{A})$.

2.2. Теорема умножения вероятностей

Произведение событий. С понятием произведения двух событий мы познакомились в §1 главы 5 (см. рис. 5.2).

Например, если A — матрица квадратная, B — определитель матрицы равен нулю, то AB — матрица вырожденная (см. §3 главы 1).

ОПРЕДЕЛЕНИЕ. *Произведением нескольких событий* называют событие, состоящее в совместном появлении всех этих событий.

Например, в целях налогообложения ценные бумаги признаются обращающимися на рынке ценных бумаг, если одновременно соблюдены условия A, B, C . Условие A — ценные бумаги допущены к обращению хотя бы одним организатором торговли, имеющим на то право в соответствии с действующим законодательством; условие B — информация об их ценах (котировках) публикуется в средствах массовой информации либо может быть представлена организатором торговли или иным уполномоченным лицом любому заинтересованному лицу в течение трех лет после даты совершения операций с ценными бумагами; условие C — по ним рассчитывается рыночная котировка, когда это предусмотрено соответствующим национальным законодательством. Тогда событие ABC — ценные бумаги признаются обращающимися на рынке ценных бумаг.

Условная вероятность. В начале главы 5 случайное событие определено как событие, которое при осуществлении совокупности условий E может произойти или не произойти. Если при вычислении вероятности события никаких других ограничений, кроме E , не налагается, то такую вероятность называют *безусловной*; если же налагаются и другие дополнительные условия, то вероят-

ность события называют *условной*. Например, часто вычисляют вероятность события B при дополнительном условии, что произошло событие A . Заметим, что и безусловная вероятность, строго говоря, является условной, поскольку предполагается осуществление условий A .

ОПРЕДЕЛЕНИЕ. *Условной вероятностью* события B при условии, что событие A уже произошло, называется число, которое обозначается

$$P(B/A) = P_A(B) = P(AB)/P(A), \quad (P(A) > 0).$$

Аналогично

$$P(A/B) = P_B(A) = P(AB)/P(B), \quad (P(B) > 0) —$$

условная вероятность события A при условии, что событие B уже произошло.

ОПРЕДЕЛЕНИЕ. Два события называют *независимыми*, если появление одного из них не влияет на вероятность появления другого; в противном случае события называют *зависимыми*.

На практике заключение о независимости событий делают исходя из смысла задачи. Например, вероятности сдачи зачета по теории вероятностей каждым из двух студентов не зависят от того, сдал ли зачет другой студент, поэтому события “первый студент получил зачет” и “второй студент получил зачет” независимы.

Произведение вероятностей. Рассмотрим два зависимых события: A и B . Пусть вероятности $P(A)$ и $P_A(B)$ известны. Как найти вероятность совмещения этих событий, т.е. вероятность того, что появится и событие A , и событие B ? Ответ на этот вопрос дает теорема умножения вероятностей.

□ **ТЕОРЕМА 5.4.** *Вероятность произведения двух зависимых событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое событие уже наступило:*

$$P(AB) = P(A)P_A(B). \quad (5.9)$$

Доказательство. По определению условной вероятности $P_A(B) = P(AB)/P(A)$. Отсюда $P(AB) = P(A)P_A(B)$. ■

Применив формулу (5.9) к событию BA , получим $P(BA) = P(B)P_B(A)$, или, поскольку событие BA не отличается от события AB ,

$$P(AB) = P(B)P_B(A). \quad (5.10)$$

Сравнивая формулы (5.9) и (5.10), делаем вывод о справедливости равенства

$$P(A)P_A(B) = P(B)P_B(A). \quad (5.11)$$

СЛЕДСТВИЕ 1. Вероятность произведения n зависимых событий равна произведению вероятности одного из них на условные вероятности всех остальных, причем вероятность каждого последующего события вычисляется в предположении, что все предыдущие события уже произошли:

$$P(A_1 \cdot A_2 \cdot A_3 \cdot \dots \cdot A_n) = P(A_1)P_{A_1}(A_2)P_{A_1 A_2}(A_3) \cdot \dots \cdot P_{A_1 A_2 \dots A_{n-1}}(A_n),$$

где $P_{A_1 A_2 \dots A_{n-1}}(A_n)$ — вероятность события A_n , вычисленная в предположении, что события $A_1, A_2, A_3, \dots, A_{n-1}$ произошли.

В частности, для трех событий

$$P(ABC) = P(A)P_A(B)P_{AB}(C).$$

Заметим, что порядок, в котором расположены события, может быть выбран любым, т.е. безразлично, какое событие считать первым, вторым и т. д.

СЛЕДСТВИЕ 2. Вероятность суммы двух совместных событий A и B равна сумме вероятностей этих событий без вероятности их произведения:

$$P(A + B) = P(A) + P(B) - P(A \cdot B).$$

▷ **ПРИМЕР 5.20.** В портфеле инвестора 10 различных акций, из которых две — с высокой степенью исполнения обязательств по ним, а остальные — с высокой степенью неисполнения обязательств по ним (“бросовые”). Наудачу выбрано 2 акции. Найдем вероятность того, что обе акции “бросовые”.

Решение. Пусть элементарные события таковы: A_1 — первая акция “бросовая”; A_2 — вторая акция “бросовая”. Тогда событие $A = A_1 \cdot A_2$ — обе акции “бросовые”. Вероятность того, что вторая акция “бросовая”, будет зависеть от того, какова первая

акция. Если первая акция “бросовая”, то вероятность того, что вторая акции также “бросовая”, равна условной вероятности события A_2 при условии, что A_1 уже произошло: $P(A_2/A_1) = 7/9$, так как после наступления события A_1 — всего акций останется 9, из них 7 “бросовые”. Отсюда

$$P(A) = P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2 / A_1) = \frac{8}{10} \cdot \frac{7}{9} = \frac{28}{45}. \blacktriangleright$$

▷ ПРИМЕР 5.21. В урне 2 белых и 3 черных шара. Из урны вынимают подряд 2 шара. Найти вероятность того, что оба шара белые.

Решение. Событие A — оба шара белые, элементарное событие A_1 — первым вытащили белый шар, элементарное событие A_2 — вторым вытащили белый шар.

Событие A наступит, если наступят одновременно и A_1 и A_2 , $A = A_1 \cdot A_2$, отсюда

$$P(A) = P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2 / A_1) = \frac{2}{5} \times \frac{1}{4} = \frac{1}{10}. \blacktriangleright$$

Независимые события. Теорема умножения вероятностей для независимых событий. Пусть вероятность события B не зависит от появления события A .

Событие B является *независимым* от события A , если условная вероятность события B равна его безусловной вероятности:

$$P_A(B) = P(B). \quad (5.12)$$

Подставив (5.12) в соотношение (5.11), получим

$$P(A)P(B) = P(B)P_B(A).$$

Отсюда

$$P_B(A) = P(A),$$

т.е. условная вероятность события A в предположении, что наступило событие B , равна его безусловной вероятности. Другими словами, событие A не зависит от события B .

Итак, если событие B не зависит от события A , то и событие A не зависит от события B , это означает, что *свойство независимости событий взаимно*.

Для *независимых* событий теорема умножения вероятностей формулируется следующим образом.

□ ТЕОРЕМА 5.5. *Вероятность произведения двух независимых событий равна произведению вероятностей этих событий:*

$$P(A \cdot B) = P(A)P(B). \blacksquare \quad (5.13)$$

Равенство (5.13) также принимают в качестве определения независимых событий.

▷ ПРИМЕР 5.22. Найти вероятность совместной сдачи зачета двумя студентами, если вероятность сдачи зачета первым студентом (событие A) равна 0,8, а вторым (событие B) — 0,7.

Р е ш е н и е. События A и B независимы, поэтому по теореме умножения вероятностей искомая вероятность

$$P(AB) = P(A)P(B) = 0,7 \cdot 0,8 = 0,56. \blacktriangleright$$

З а м е ч а н и е. Если события A и B независимы, то независимы также события A и $\bar{B}$, $\bar{A}$ и B , $\bar{A}$ и $\bar{B}$. В примере 5.20 событие $\bar{A}$ — первый студент зачет не сдаст, $\bar{B}$ — второй студент зачет не сдаст.

Действительно,

$$A = A\bar{B} + AB.$$

Следовательно,

$$P(A) = P(A\bar{B}) + P(AB), \text{ или } P(A) = P(A\bar{B}) + P(A)P(B).$$

Отсюда

$$P(A\bar{B}) = P(A)[1 - P(B)], \text{ или } P(A\bar{B}) = P(A)P(\bar{B}),$$

т.е. события A и $\bar{B}$ независимы.

Независимость событий $\bar{A}$ и B , $\bar{A}$ и $\bar{B}$ — следствие теоремы 5.5.

ОПРЕДЕЛЕНИЕ. Несколько событий называют *попарно независимыми*, если каждые два из них независимы.

Например, события A , B , C попарно независимы, если независимы события A и B , A и C , B и C .

Для того чтобы обобщить теорему умножения вероятностей на несколько событий, введем понятие независимости событий в совокупности.

ОПРЕДЕЛЕНИЕ. Несколько событий называют *независимыми в совокупности* (или просто независимыми), если независимы каждые два из них и независимы каждое событие и все возможные произведения остальных.

Например, если события A_1, A_2, A_3 независимы в совокупности, то независимы события A_1 и A_2 ; A_2 и A_3 ; A_3 и A_1 ; A_1 и $A_2 A_3$; A_2 и $A_1 A_3$; A_3 и $A_1 A_2$. Из сказанного следует, что если события независимы в совокупности, то условная вероятность появления любого из них, вычисленная в предположении, что наступили какие-либо другие события из числа остальных, равна его безусловной вероятности.

Подчеркнем, что если несколько событий независимы попарно, то отсюда еще не следует их независимость в совокупности. В этом смысле требование независимости событий в совокупности сильнее требования их попарной независимости.

Поясним сказанное на примере. Пусть в урне имеется 4 шара, окрашенные: один — в красный цвет (A), один — в синий цвет (B), один — в черный цвет (C) и один — во все эти три цвета (ABC). Вычислим вероятность того, что извлеченный из урны шар имеет красный цвет.

Так как из четырех шаров два имеют красный цвет, то $P(A) = 2/4 = 1/2$. Рассуждая аналогично, найдем $P(B) = 1/2$, $P(C) = 1/2$. Допустим теперь, что взятый шар имеет синий цвет, т.е. событие B уже произошло. Изменится ли вероятность того, что извлеченный шар имеет красный цвет, т.е. изменится ли вероятность события A ? Из двух шаров, имеющих синий цвет, один шар имеет и красный цвет, поэтому вероятность события A по-прежнему равна $1/2$. Другими словами, условная вероятность события A , вычисленная в предположении, что наступило событие B , равна его безусловной вероятности. Следовательно, события A и B независимы. Аналогично придем к выводу, что события A и C , B и C независимы. Итак, события A , B и C попарно независимы.

Независимы ли эти события в совокупности? Оказывается, нет. Действительно, пусть извлеченный шар имеет два цвета, например, синий и черный. Чему равна вероятность того, что этот шар имеет красный цвет? Лишь один шар окрашен во все три цвета,

поэтому взятый шар имеет и красный цвет. Таким образом, допустив, что события B и C произошли, приходим к выводу, что событие A обязательно наступит. Следовательно, это событие достоверное и вероятность его равна единице. Другими словами, условная вероятность $P_{BC}(A) = 1$ события A не равна его безусловной вероятности $P(A) = 1/2$. Итак, попарно независимые события A, B, C не являются независимыми в совокупности.

Приведем теперь следствие из теоремы умножения вероятностей.

СЛЕДСТВИЕ. *Вероятность совместного появления нескольких событий, независимых в совокупности, равна произведению вероятностей этих событий:*

$$P(A_1, A_2, A_3, \dots, A_n) = P(A_1)P(A_2)P(A_3) \dots P(A_n).$$

З а м е ч а н и е. Если события $A_1, A_2, \dots, A_n$ независимы в совокупности, то и противоположные им события $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$ также независимы в совокупности.

▷ **ПРИМЕР 5.23.** Найти вероятность совместного появления герба при одном бросании двух монет.

Р е ш е н и е. Вероятность появления “герба” первой монеты (событие A)

$$P(A) = 1/2.$$

Вероятность появления “герба” второй монеты (событие B)

$$P(B) = 1/2.$$

События A и B независимы, поэтому искомая вероятность по теореме умножения вероятностей равна

$$P(AB) = P(A)P(B) = 1/2 \cdot 1/2 = 1/4. \blacktriangleright$$

Вероятность появления хотя бы одного события. Пусть в результате испытания могут появиться n событий, независимых в совокупности, либо некоторые из них (в частности, только одно или ни одного), причем вероятности появления каждого из событий известны. Как найти вероятность того, что наступит хотя бы одно из этих событий? Например, если в результате испытания могут появиться три события, появление хотя бы одного из этих них означает наступление либо одного, либо двух, либо трех событий. Ответ на поставленный вопрос дает следующая теорема.

□ ТЕОРЕМА 5.6. Вероятность наступления хотя бы одного из событий $A_1, A_2, \dots, A_n$, независимых в совокупности, равна разности между единицей и произведением вероятностей противоположных событий $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n$:

$$P(A) = 1 - q_1 q_2 \cdots q_n. \blacksquare$$

Частный случай: если события $A_1, A_2, \dots, A_n$ имеют одинаковую вероятность, равную p , то вероятность появления хотя бы одного из этих событий

$$P(A) = 1 - q^n. \quad (5.14)$$

▷ ПРИМЕР 5.24. Вероятности получения дохода частным инвестором при вложении средств в паевой инвестиционный фонд A (ПИФ A), ПИФ B и ПИФ C таковы: $p_1 = 0,9$; $p_2 = 0,8$; $p_3 = 0,7$. Найти вероятность получения инвестором прибыли хотя бы с одной инвестиции (событие A) в конкретный биржевой день.

Р е ш е н и е. Вероятность получения дохода при вложении средств в каждый из трех фондов не зависит от результатов вложения средств в другие фонды. Поэтому рассматриваемые события A_1 (получение дохода от инвестиции в ПИФ A), A_2 (получение дохода от инвестиции в ПИФ B) и A_3 (получение дохода от инвестиции в ПИФ C) независимы в совокупности.

Вероятности событий, противоположных событиям A_1, A_2 и A_3 (т.е. вероятности неполучения дохода), соответственно равны:

$$q_1 = 1 - p_1 = 1 - 0,9 = 0,1; \quad q_2 = 1 - p_2 = 1 - 0,8 = 0,2;$$

$$q_3 = 1 - p_3 = 1 - 0,7 = 0,3.$$

Искомая вероятность

$$P(A) = 1 - q_1 q_2 q_3 = 1 - 0,1 \cdot 0,2 \cdot 0,3 = 0,994. \blacktriangleright$$

▷ ПРИМЕР 5.25. В компьютерном классе имеется 4 принтера. Для каждого принтера вероятность того, что он работает в данный момент, равна 0,9. Найти вероятность того, что в данный момент работает хотя бы один принтер (событие A).

Р е ш е н и е. События “принтер работает” и “принтер не работает” (в данный момент) — противоположны, поэтому сумма их вероятностей равна единице:

$$p + q = 1.$$

Отсюда вероятность того, что принтер в данный момент не работает, равна

$$q = 1 - p = 1 - 0,9 = 0,1.$$

Искомая вероятность

$$P(A) = 1 - q^4 = 1 - 0,1^4 = 0,9999.$$

Так как полученная вероятность весьма близка к единице, мы вправе заключить, что в данный момент работает хотя бы один из принтеров. ►

▷ ПРИМЕР 5.26. Вероятность того, что торговец акциями, торгующий во внутридневном режиме (day-Trader), при однократной покупке акций компании D закроет свою позицию (продаст акции) с прибылью, равна 0,4. Сколько сделок должен провести day-Trader, чтобы с вероятностью не менее 0,9 он получил прибыль хотя бы от одной сделки?

Р е ш е н и е. Обозначим через A событие “при n сделках day-Trader получает прибыль хотя бы от одной сделки”. События, состоящие в получении прибыли при первой, второй сделках и т.д., независимы в совокупности, поэтому применима формула (5.14)

$$P(A) = 1 - q.$$

Приняв во внимание, что по условию $P(A) \geq 0,9$, $p = 0,4$ (следовательно, $q = 1 - 0,4 = 0,6$), получим $1 - 0,6^n \geq 0,9$; отсюда $0,6^n \leq 0,1$. Прологарифмируем это неравенство по основанию 10: $n \lg 0,6 \leq \lg 0,1$. Отсюда, учитывая, что $\lg 0,6 < 0$, имеем

$$n \geq \lg 0,1 / \lg 0,6 = -1 / (-0,2218) = 4,5.$$

Итак, $n \geq 5$, т.е. day-Trader должен провести не менее 5 сделок. ►

2.3. Вероятности гипотез и формулы полной вероятности

Формула Байеса. Пусть событие A может наступать только одновременно с одним из попарно несовместных событий $H_1, H_2, \dots, H_n$, образующих полную группу. Тогда вероятность события A определяется по формуле полной вероятности:

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + \dots + P(H_n) \cdot P(A/H_n),$$

или

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i), \quad (5.15)$$

где события $H_1, H_2, \dots, H_n$ — гипотезы, а $P(A/H_i)$ — условная вероятность наступления события A при наступлении i -й гипотезы ($i = 1, 2, \dots, n$).

Условная вероятность гипотезы H_i при условии того, что событие A произошло, определяется по формуле вероятности гипотез, или *формуле Байеса* (она позволяет пересмотреть вероятности гипотез после наступления события A):

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{P(A)}. \quad (5.16)$$

▷ ПРИМЕР 5.27. Команда стрелков состоит из 5 человек. Трое из них попадают с вероятностью 0,8, а двое — с вероятностью 0,6. Наудачу из команды берется стрелок и проводит выстрел.

а) Какова вероятность того, что стрелок попадет?

б) Если стрелок попал в цель, то какова вероятность, что это один из 3-х (один из 2-х)?

Р е ш е н и е. а) Событие A может произойти, если произойдет одно из несовместных событий: H_1 — наудачу взятый стрелок — один из трех, H_2 — наудачу взятый стрелок — один из двух. Для определения вероятности события A воспользуемся формулой (5.15):

$$\begin{aligned} P(A) &= P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) = \frac{3}{5} \cdot \frac{8}{10} + \frac{2}{5} \cdot \frac{6}{10} = \\ &= \frac{36}{50} = \frac{18}{25}, \end{aligned}$$

$$\text{так как } P(H_1) = \frac{3}{5}, P(A/H_1) = 0,8; P(H_2) = \frac{2}{5}, P(A/H_2) = 0,6.$$

б) по формуле (5.16)

$$P(H_1/A) = \frac{P(H_1) \cdot P(A/H_1)}{P(A)} = \frac{\frac{24}{50}}{\frac{36}{50}} = \frac{2}{3},$$

$$P(H_2 / A) = \frac{P(H_2) \cdot P(A / H_2)}{P(A)} = \frac{\frac{12}{50}}{\frac{36}{50}} = \frac{1}{3},$$

$$\text{или } P(H_2 / A) = 1 - P(H_1 / A) = 1 - \frac{2}{3} = \frac{1}{3}. \blacktriangleright$$

Формула Бернулли. Пусть некоторый опыт повторяется в неизменных условиях n раз, причем каждый раз может либо наступить (успех), либо не наступить (неудача) некоторое событие A , где $P(A) = p$ — вероятность успеха, $P(\bar{A}) = 1 - p = q$ — вероятность неудачи. Тогда вероятность того, что в k случаях из n произойдет событие A , вычисляется по формуле Бернулли:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}.$$

Условия, приводящие к формуле Бернулли, называются *схемой независимых повторных испытаний*, или *схемой Бернулли*. Так как вероятности $P_n(k)$ для различных значений k представляют собой слагаемые в разложении бинома Ньютона:

$$(p + q)^n = C_n^0 \cdot p^0 \cdot q^n + C_n^1 \cdot p^1 \cdot q^{n-1} + \dots + C_n^k \cdot p^k \cdot q^{n-k} + \dots + C_n^n \cdot p^n \cdot q^0,$$

то распределение вероятностей $P_n(k)$, где $0 \leq k \leq n$, называется *биномиальным*.

Если в каждом из независимых испытаний вероятности наступления события A разные, то вероятность наступления события A k раз в n опытах определяется как коэффициент при k -й степени полинома:

$$\varphi_n(Z) = \prod_{i=1}^n (q_i + p_i Z) = a_n Z^n + a_{n-1} Z^{n-1} + \dots + a_1 Z^1 + a_0,$$

где $\varphi_n(Z)$ — производящая функция.

▷ ПРИМЕР 5.28. Частный инвестор вкладывает средства в три паевых инвестиционных фонда: ПИФ 1, ПИФ 2, ПИФ 3. Вероятности получения инвестором прибыли при вложении своих средств в каждый из трех фондов различны и равны: $p_1 = 0,7$; $p_2 = 0,8$; $p_3 = 0,9$.

Найти а) вероятность неполучения дохода; б) вероятность получения дохода с одной из инвестиций; в) вероятность неполучения с любых двух инвестиций; г) вероятность получения с трех инвестиций.

Р е ш е н и е. Производящая функция

$$\varphi_3(Z) = (0,3 + 0,7 Z) \cdot (0,2 + 0,8 Z) \cdot (0,1 + 0,9 Z) = 0,504 Z^3 + 0,398 Z^2 + 0,092 Z + 0,006.$$

Отсюда вероятность:

- а) неполучения дохода $P_3(0) = 0,006$;
- б) получения дохода с любой одной из инвестиций $P_3(1) = 0,092$;
- в) получения дохода с любых двух из инвестиций $P_3(2) = 0,398$;
- г) получения дохода со всех трех инвестиций $P_3(3) = 0,504$. ►

§ 3. Случайные величины

3.1. Определение случайной величины

Ранее мы уже приводили примеры событий, состоящих в появлении того или иного числа. Например, при бросании игральной кости могли появиться числа 1, 2, 3, 4, 5 и 6. Наперед определить число выпавших очков невозможно, поскольку оно зависит от многих случайных причин, которые полностью не могут быть учтены. Числа 1, 2, 3, 4, 5 и 6 есть *возможные значения* этой величины.

ОПРЕДЕЛЕНИЕ. *Случайной величиной* называют такую величину, которая в результате опыта может принимать те или иные значения, причем до опыта мы не можем сказать, какое именно значение она примет.

Более точно, случайная величина — это действительная функция, определенная на пространстве элементарных событий Ω .

Приведем примеры случайных величин:

1. Число родившихся девочек среди 100 новорожденных есть случайная величина, которая имеет следующие возможные значения: 0, 1, 2, ..., 100.

2. Расстояние, которое пролетит снаряд при выстреле из орудия, есть случайная величина. Действительно, расстояние зависит не только от установки прицела, но и от многих других причин, которые не могут быть полностью учтены. Возможные значения этой величины принадлежат некоторому промежутку (a, b) .

Будем далее обозначать случайные величины прописными буквами X, Y, Z , а их возможные значения соответствующими строчными буквами x, y, z . Например, если случайная величина X имеет три возможных значения, то они будут обозначены так: x_1, x_2, x_3 .

3.2. Определение функции распределения

Пусть x — действительное число. Вероятность события, состоящего в том, что X примет значение, меньшее x , т.е. вероятность события $X < x$, обозначим через $F(x)$. Разумеется, если

x изменяется, то, вообще говоря, изменяется и $F(x)$, т.е. $F(x)$ — функция от x .

ОПРЕДЕЛЕНИЕ. *Функцией распределения* называется числовая функция числового аргумента $F(x)$, определяющая вероятность того, что случайная величина X в результате опыта примет значение, меньшее x , т.е.

$$F(x) = P(X < x), x \in \mathbf{R}$$

где $\mathbf{R}$ — множество действительных чисел.

Геометрически это равенство можно истолковать так: $F(x)$ есть вероятность того, что случайная величина примет значение, которое изображается на числовой оси точкой, лежащей левее точки x .

Иногда вместо термина “функция распределения” используют термин “интегральная функция”.

3.3. Дискретные и непрерывные случайные величины

Случайные величины могут быть трех типов: *дискретные, непрерывные, смешанные* (дискретно-непрерывные).

Дискретная случайная величина может принимать конечное или бесконечное счетное число значений. Например, подбрасываем монету 5 раз. Случайная величина X — число появлений герба может принимать значения 0, 1, 2, 3, 4, 5.

ОПРЕДЕЛЕНИЕ. *Дискретной (разрывной)* называется случайная величина, которая принимает отдельные, изолированные возможные значения с определенными вероятностями.

ОПРЕДЕЛЕНИЕ. Случайную величину называют *непрерывной*, если ее функция распределения есть непрерывная, кусочно-дифференцируемая функция с непрерывной производной.

Непрерывная случайная величина в отличие от дискретной случайной величины принимает бесконечное несчетное число значений. Например, мишень имеет форму круга радиуса R . По этой мишени произвели выстрел с обязательным попаданием. Обозначим через Y расстояние от центра мишени до точки попадания, $Y \in [0; R]$. Y — непрерывная случайная величина, так как она принимает бесконечное несчетное число значений.

3.4. Свойства функции распределения вероятностей случайной величины

Свойства функции распределения.

1. Значения функции распределения принадлежат отрезку $[0; 1]$:

$$0 \leq F(x) \leq 1 \text{ при } \forall x \in \mathbf{R}.$$

2. $F(x)$ — неубывающая функция, т.е.

$$F(x_2) \geq F(x_1), \text{ если } x_2 > x_1.$$

СЛЕДСТВИЕ 1. Вероятность того, что случайная величина примет значение, заключенное в интервале (a, b) , равна приращению функции распределения на этом интервале:

$$P(a \leq X \leq b) = F(b) - F(a).$$

СЛЕДСТВИЕ 2. Вероятность того, что **непрерывная** случайная величина X примет одно определенное значение, равна нулю.

3. Если возможные значения случайной величины принадлежат интервалу (a, b) , то: 1) $F(x) = 0$ при $x \leq a$; 2) $F(x) = 1$ при $x \geq b$.

СЛЕДСТВИЕ. Если возможные значения **непрерывной** случайной величины расположены на всей оси x , то справедливы следующие предельные отношения:

$$\lim_{x \rightarrow -\infty} F(x) = 0; \quad \lim_{x \rightarrow +\infty} F(x) = 1.$$

Любая функция, удовлетворяющая этим свойствам, является функцией распределения некоторой случайной величины. Нестрого говоря, функция распределения содержит всю вероятностную информацию о случайной величине X .

График функции распределения. Описанные выше свойства позволяют представить, как выглядит график функции распределения непрерывной случайной величины.

График расположен в полосе, ограниченной прямыми $y = 0$, $y = 1$ (согласно свойству 1).

При возрастании x в интервале (a, b) , в котором заключены все возможные значения случайной величины, график “поднимается вверх” (согласно свойству 2).

При $x \leq a$ ординаты графика равны нулю; при $x \geq b$ ординаты графика равны единице (согласно свойству 3).

График функции распределения непрерывной случайной величины изображен на рис. 5.5.

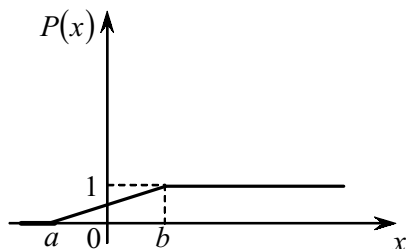


Рис. 5.5

3.5. Закон распределения дискретной случайной величины

Несмотря на то, что функция распределения универсальным образом описывает вероятностный характер случайной величины, *дискретную* случайную величину можно задать таблично, графически и аналитически. Соответствие между возможными значениями и их вероятностями называют *законом распределения дискретной случайной величины*.

Пусть X — дискретная случайная величина, которая принимает значения $x_1, x_2, \dots, x_n, \dots$ с некоторой вероятностью p_i , где $i = 1, 2, \dots, n, \dots$. Тогда можно говорить о вероятности того, что случайная величина X приняла значение x_i : $p_i = P(X = x_i)$. Значения x_i и соответствующие вероятности p_i представляют в виде таблицы.

При *табличном* задании закона распределения дискретной случайной величины первая строка таблицы содержит возможные значения, а вторая — их вероятности:

X	x_1	x_2	$\dots$	x_n
p	p_1	p_2	$\dots$	p_n

Приняв во внимание, что в одном испытании случайная величина принимает одно и только одно возможное значение, заклю-

чаем, что события $X = x_1, X = x_2, \dots, X = x_n$ образуют полную группу. Следовательно, сумма вероятностей этих событий, т.е. сумма вероятностей второй строки таблицы, равна единице:

$$p_1 + p_2 + \dots + p_n = 1.$$

Если множество возможных значений X бесконечно (счетно), то ряд $p_1 + p_2 + \dots$ сходится и его сумма равна единице.

▷ ПРИМЕР 5.29. В денежной лотерее выпущено 100 билетов. Разыгрывается один выигрыш 50 руб. и 10 выигрышей по одному рублю. Найти закон распределения случайной величины X — стоимости возможного выигрыша для владельца одного лотерейного билета.

Р е ш е н и е. Напишем возможные значения X : $x_1 = 50, x_2 = 1, x_3 = 0$. Вероятности этих возможных значений таковы: $p_1 = 0,01, p_2 = 0,1, p_3 = 1 - (p_1 + p_2) = 0,89$.

Напишем искомый закон распределения:

X	50	1	0
p	0,01	0,1	0,89

Для наглядности закон распределения дискретной случайной величины можно изобразить и *г р а ф и ч е с к и*, для чего в прямоугольной системе координат строят точки (x_i, p_i) , а затем соединяют их отрезками прямых. Полученную фигуру называют *многоугольником распределения*.

▷ ПРИМЕР 5.30. Дискретная случайная величина X задана законом распределения:

X	1	3	6	8
p	0,2	0,1	0,4	0,3

Построить многоугольник распределения.

Р е ш е н и е. Построим прямоугольную систему координат, причем по оси абсцисс будем откладывать возможные значения, а по оси ординат — соответствующие вероятности. Построим точки $M_1(1; 0,2), M_2(3; 0,1), M_3(6; 0,4)$ и $M_4(8; 0,3)$ (рис. 5.6). Соединив

эти точки отрезками прямых, получим искомый многоугольник распределения. ►

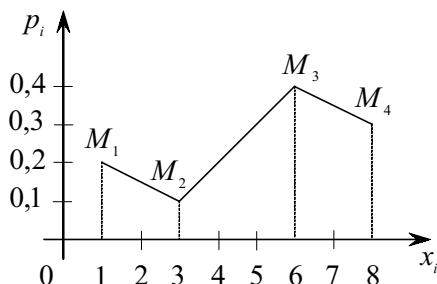


Рис. 5.6

Над случайными величинами устанавливаются операции сложения и умножения.

ОПРЕДЕЛЕНИЕ. *Суммой* двух случайных величин X и Y называется случайная величина, которая получается в результате сложения всех значений случайной величины X и всех значений случайной величины Y ; соответствующие вероятности перемножаются.

ОПРЕДЕЛЕНИЕ. *Произведением* двух случайных величин X и Y называется случайная величина, которая получается в результате перемножения всех значений случайной величины X и всех значений случайной величины Y ; соответствующие вероятности не перемножаются.

► **ПРИМЕР 5.31.** Дискретные случайные величины X и Y заданы в виде таблиц:

x_i	0	1	2	3
p_i	0,1	0,4	0,3	0,2

y_i	-1	0	1
p_i	0,2	0,3	0,5

Найти: 1) $X + C$, где $C = 2$; 2) $X + Y$.

Решение.

1) $Z = X + C$, $C = 2$.

z_i	0+2	1+2	2+2	3+2
p_i	0,1	0,4	0,3	0,2

z_i	2	3	4	5
p_i	0,1	0,4	0,3	0,2

$$2) Z = X + Y.$$

z_i	0-1	0+0	0+1	1-1	1+0	1+1	2-1	2+0	2+1	3-1	3+0	3+1
p_i	0,02	0,03	0,05	0,08	0,12	0,2	0,06	0,09	0,15	0,04	0,06	0,1

Одинаковые значения случайных величин можно записать один раз, предварительно сложив соответствующие вероятности:

z_i	-1	0	1	2	3	4
p_i	0,02	0,11	0,23	0,33	0,21	0,1

З а м е ч а н и е. График функции распределения *дискретной* случайной величины имеет ступенчатый вид. Убедимся в этом на примере.

▷ **ПРИМЕР 5.32.** Дискретная случайная величина задана таблицей распределения:

X	1	4	8
p	0,3	0,1	0,6

Найти функцию распределения и построить ее график.

Р е ш е н и е. Если $x \leq 1$, то $F(x) = 0$ (согласно свойству 3). Если $1 < x \leq 4$, то $F(x) = 0,3$. Действительно, X может принять значение 1 с вероятностью 0,3.

Если $4 < x \leq 8$, то $F(x) = 0,4$. Действительно, если x удовлетворяет неравенству $4 < x \leq 8$, то $F(x)$ равно вероятности события $X < x$, которое может быть осуществлено, когда X примет значение 1 (вероятность этого события равна 0,3) или значение 4 (вероятность этого события равна 0,1). Поскольку эти два события несовместны, то по теореме сложения вероятностей (см. теорему 5.1) вероятность события $X < x$ равна сумме вероятностей $0,3 + 0,1 = 0,4$.

Если $x > 8$, то $F(x) = 1$. Действительно, событие $X \leq 8$ достоверно, следовательно, его вероятность равна единице.

Итак, функция распределения аналитически может быть записана так:

$$F(x) = \begin{cases} 0 & \text{при } x \leq 1, \\ 0,3 & \text{при } 1 < x \leq 4, \\ 0,4 & \text{при } 4 < x \leq 8, \\ 1 & \text{при } x > 8. \end{cases}$$

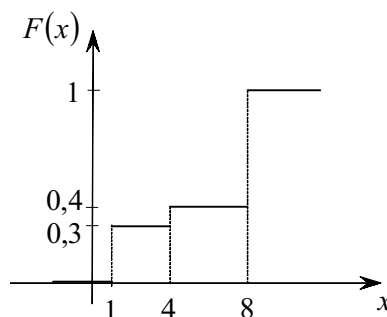


График этой функции приведен на рис. 5.7. ►

Рис. 5.7

3.6. Плотность распределения вероятностей непрерывной случайной величины

Определение плотности распределения. Непрерывная случайная величина задавалась нами ранее с помощью функции распределения. Этот способ задания не является единственным. Непрерывную случайную величину можно также задать, используя другую функцию, которую называют плотностью распределения, или плотностью вероятности (иногда ее называют дифференциальной функцией).

ОПРЕДЕЛЕНИЕ. *Плотностью распределения вероятностей* непрерывной случайной величины X называется функция $f(x)$ — первая производная от функции распределения $F(x)$:
 $f(x) = F'(x)$.

Из определения следует, что функция распределения является первообразной для плотности распределения.

Заметим, что для описания распределения вероятностей *дискретной* случайной величины плотность распределения неприемима.

Вероятность попадания непрерывной случайной величины в заданный интервал. Зная плотность распределения, можно вычислить вероятность того, что непрерывная случайная величина примет значение, принадлежащее заданному интервалу. Вычисление основано на следующей теореме.

□ **ТЕОРЕМА 5.7.** Вероятность того, что непрерывная случайная величина X примет значение, принадлежащее интервалу (a, b) , равна определенному интегралу от плотности распределения, взятому в пределах от a до b :

$$P(a < X < b) = \int_a^b f(x) dx. \blacksquare$$

В частности, если $f(x)$ — четная функция и концы интервала симметричны относительно начала координат, то

$$P(-a < X < a) = 2 \int_0^a f(x) dx.$$

▷ **ПРИМЕР 5.33.** Задана плотность вероятности случайной величины X

$$f(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ 2x & \text{при } 0 < x \leq 1, \\ 0 & \text{при } x > 1. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, принадлежащее интервалу $(0,5; 1)$.

Р е ш е н и е. Искомая вероятность

$$\begin{aligned} P(0,5 < X < 1) &= \int_{0,5}^1 2x dx = 2 \int_{0,5}^1 x dx = 2 \cdot \frac{x^2}{2} \Big|_{0,5}^1 x^2 \Big|_{0,5}^1 = \\ &= 1 - 0,25 = 0,75. \blacktriangleright \end{aligned}$$

Нахождение функции распределения по известной плотности распределения. Зная плотность распределения $f(x)$, можно найти функцию распределения $F(x)$ по формуле

$$F(x) = \int_{-\infty}^x f(x) dx. \quad (5.17)$$

▷ П Р И М Е Р 5.34. Найти функцию распределения по данной плотности распределения:

$$f(x) = \begin{cases} 0 & \text{при } x \leq a, \\ 1/(b-a) & \text{при } a < x \leq b, \\ 0 & \text{при } x > b \end{cases}$$

и построить ее график.

Р е ш е н и е. Воспользуемся формулой (5.17).

Если $x \leq a$, то $f(x) = 0$, следовательно, $F(x) = 0$. Если $a < x \leq b$, то $f(x) = 1/(b-a)$, следовательно,

$$F(x) = \int_{-\infty}^x f(x) dx = \int_{-\infty}^a 0 dx + \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a}.$$

Если $x > b$, то

$$F(x) = \int_{-\infty}^x f(x) dx = \int_{-\infty}^a 0 dx + \int_a^b \frac{1}{b-a} dx + \int_b^x 0 dx = \frac{b-a}{b-a} = 1.$$

Итак, искомая функция распределения

$$F(x) = \begin{cases} 0 & \text{при } x \leq a, \\ (x-a)/(b-a) & \text{при } a < x \leq b, \\ 1 & \text{при } x > b. \end{cases}$$

График функции изображен на рис. 5.8. ►

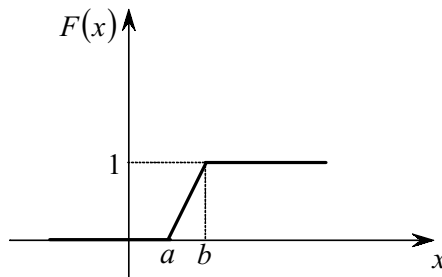


Рис. 5.8

Свойства плотности распределения.

1. *Плотность распределения — неотрицательная функция:*

$$f(x) \geq 0.$$

График плотности распределения называют *кривой распределения*.

2. *Несобственный интеграл от плотности распределения в пределах от $-\infty$ до ∞ равен единице:*

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

§ 4. Числовые характеристики случайных величин

4.1. Числовые характеристики дискретных случайных величин

На практике нет необходимости характеризовать случайную величину полностью. Обычно достаточно указать только отдельные числовые параметры, которые описывают случайную величину суммарно; такие числовые параметры называются *числовыми характеристиками случайной величины*. К числу важных числовых характеристик относятся *математическое ожидание*, *мода*; характеристики рассеяния: *дисперсия*, *среднее квадратическое отклонение*.

ОПРЕДЕЛЕНИЕ. *Математическим ожиданием* дискретной случайной величины называется сумма произведений всех ее возможных значений на их вероятности.

Пусть случайная величина X может принимать только значения $x_1, x_2, \dots, x_n$, вероятности которых соответственно равны $p_1, p_2, \dots, p_n$. Тогда математическое ожидание $M(X)$ случайной величины X определяется равенством

$$M(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n.$$

Если дискретная случайная величина X принимает счетное множество возможных значений, то

$$M(X) = \sum_i x_i p_i,$$

причем математическое ожидание существует, если ряд в правой части равенства сходится абсолютно.

З а м е ч а н и е. Из определения следует, что математическое ожидание дискретной случайной величины есть неслучайная (постоянная) величина. Рекомендуем запомнить это утверждение, так как оно будет использоваться многократно. В дальнейшем будет показано, что математическое ожидание непрерывной случайной величины также есть постоянная величина.

▷ П Р И М Е Р 5.35. Найти математическое ожидание случайной величины X , зная ее закон распределения:

X	2	1	3
p	0,1	0,6	0,3

Р е ш е н и е. Искомое математическое ожидание равно сумме произведений всех возможных значений случайной величины на их вероятности:

$$M(X) = 2 \cdot 0,1 + 1 \cdot 0,6 + 3 \cdot 0,3 = 1,7. \blacktriangleright$$

▷ П Р И М Е Р 5.36. Найти математическое ожидание числа появлений события A в одном испытании, если вероятность события A равна p .

Р е ш е н и е. Случайная величина X — число появлений события A в одном испытании — может принимать только два значения: $x_1 = 1$ (событие A наступило) с вероятностью p и $x_2 = 0$ (событие A не наступило) с вероятностью $q = 1 - p$. Искомое математическое ожидание

$$M(X) = 1 \cdot p + 0 \cdot q = p. \blacktriangleright$$

Итак, математическое ожидание числа появлений события в одном испытании равно вероятности этого события.

Свойства математического ожидания.

1. Математическое ожидание постоянной величины равно самой постоянной:

$$M(C) = C.$$

2. Постоянный множитель можно выносить за знак математического ожидания:

$$M(CX) = CM(X).$$

3. Математическое ожидание произведения двух независимых случайных величин равно произведению их математических ожиданий:

$$M(XY) = M(X)M(Y).$$

С Л Е Д С Т В И Е. Математическое ожидание произведения нескольких взаимно независимых случайных величин равно произведению их математических ожиданий.

Например, для трех случайных величин

$$M(XYZ) = M(XY \cdot Z) = M(XY)M(Z) = M(X)M(Y)M(Z).$$

Для произвольного числа случайных величин доказательство проводится методом математической индукции.

▷ ПРИМЕР 5.37. Независимые случайные величины X и Y заданы следующими законами распределения:

X	1	2	3	Y	5	10
p	0,6	0,1	0,3	p	0,8	0,2

Найти математическое ожидание случайной величины XY .

Р е ш е н и е. Найдем математическое ожидание каждой из величин:

$$M(X) = 1 \cdot 0,6 + 2 \cdot 0,1 + 3 \cdot 0,3 = 1,7;$$

$$M(Y) = 5 \cdot 0,8 + 10 \cdot 0,2 = 6.$$

Случайные величины X и Y независимы, поэтому искомое математическое ожидание

$$M(XY) = M(X)M(Y) = 1,7 \cdot 6 = 10,2. \blacktriangleright$$

4. Математическое ожидание суммы двух случайных величин равно сумме математических ожиданий слагаемых:

$$M(X + Y) = M(X) + M(Y).$$

СЛЕДСТВИЕ. Математическое ожидание суммы нескольких случайных величин равно сумме математических ожиданий слагаемых.

Например для трех случайных величин

$$M(X + Y + Z) = M[(X + Y) + Z] = M(X + Y) + M(Z) = M(X) + M(Y) + M(Z).$$

ОПРЕДЕЛЕНИЕ. *Модой* $M_0(X)$ называется значение случайной величины, имеющее наиболее вероятное значение.

Математическое ожидание числа появлений события в независимых испытаниях. Пусть производится n независимых испытаний, в каждом из которых вероятность появления события A постоянна и равна p . Чему равно среднее число появлений события A в этих испытаниях? Ответ на этот вопрос дает следующая теорема.

□ ТЕОРЕМА 5.8. Математическое ожидание $M(X)$ числа появления события A в n независимых испытаниях равно произведению числа испытаний на вероятность появления события в каждом испытании:

$$M(X) = np. \blacksquare$$

▷ ПРИМЕР 5.38. Вероятность попадания в цель при стрельбе из орудия $p = 0,7$. Найти математическое ожидание общего числа попаданий, если будет произведено 10 выстрелов.

Р е ш е н и е. Попадание при каждом выстреле не зависит от исходов других выстрелов, поэтому рассматриваемые события независимы и, следовательно, искомое математическое ожидание

$$M(X) = np = 10 \cdot 0,7 = 7 \text{ (попаданий)}. \blacktriangleright$$

Одного математического ожидания недостаточно для характеристики случайных величин, так как можно указать такие случайные величины, которые имеют одинаковые математические ожидания, но различные возможные значения. Рассмотрим, например, дискретные случайные величины X и Y , заданные следующими законами распределения:

X	-1	1	Y	-100	100
p	$0,5$	$0,5$	p	$0,5$	$0,5$

Найдем математические ожидания этих величин:

$$M(X) = -1 \cdot 0,5 + 1 \cdot 0,5 = 0,$$

$$M(Y) = -100 \cdot 0,5 + 100 \cdot 0,5 = 0.$$

Хотя математические ожидания обеих величин одинаковы, их возможные значения различны. Таким образом, зная лишь математическое ожидание случайной величины, нельзя судить ни о том, какие возможные значения она может принимать, ни о том, как они рассеяны вокруг математического ожидания. Другими словами, математическое ожидание полностью математическую величину не характеризует и необходимо ввести другие характеристики.

Отклонение случайной величины от ее математического ожидания. Пусть X — случайная величина и $M(X)$ — ее математическое

ожидание. Рассмотрим в качестве новой случайной величины разность $X - M(X)$.

ОПРЕДЕЛЕНИЕ. *Отклонением* называют разность между случайной величиной и ее математическим ожиданием.

Пусть закон распределения X известен:

X	x_1	x_2	$\dots$	x_n
p	p_1	p_2	$\dots$	p_n

Напишем закон распределения отклонения. Для того чтобы отклонение приняло значение $x_1 - M(X)$ достаточно, чтобы случайная величина приняла значение x_1 . Вероятность же этого события равна p_1 , следовательно, и вероятность того, что отклонение примет значение $x_1 - M(X)$, также равна p_1 . Аналогично обстоит дело и для остальных возможных значений отклонения.

Таким образом, отклонение имеет следующий закон распределения:

$X - M(X)$	$x_1 - M(X)$	$x_2 - M(X)$	$\dots$	$x_n - M(X)$
p	p_1	p_2	$\dots$	p_n

Приведем важное свойство отклонения.

□ **ТЕОРЕМА 5.9.** *Математическое ожидание отклонения равно нулю:*

$$M[X - M(X)] = 0. \blacksquare$$

▷ **ПРИМЕР 5.39.** Задан закон распределения дискретной случайной величины X :

X	1	2
p	0,2	0,8

Убедиться, что математическое ожидание отклонения равно нулю.

Р е ш е н и е. Найдем математическое ожидание X :

$$M(X) = 1 \cdot 0,2 + 2 \cdot 0,8 = 1,8.$$

Найдем возможные значения отклонения, для чего из возможных значений X вычтем математическое ожидание $M(X)$: $1 - 1,8 = -0,8$; $2 - 1,8 = 0,2$. Напишем закон распределения отклонения:

$X - M(X)$	$-0,8$	$0,2$
p	$0,2$	$0,8$

Найдем математическое ожидание отклонения:

$$M[X - M(X)] = (-0,8) \cdot 0,2 + 0,2 \cdot 0,8 = 0.$$

Итак, математическое ожидание отклонения равно нулю, как и должно быть. ►

Определение дисперсии дискретной случайной величины.

ОПРЕДЕЛЕНИЕ. *Дисперсией* (рассеянием) дискретной случайной величины называют математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

$$D(X) = M[X - M(X)]^2.$$

Пусть случайная величина задана законом распределения

X	x_1	x_2	$\dots$	x_n
p	p_1	p_2	$\dots$	p_n

Тогда квадрат отклонения имеет следующий закон распределения:

$[X - M(X)]^2$	$[x_1 - M(X)]^2$	$[x_2 - M(X)]^2$	$\dots$	$[x_n - M(X)]^2$
p	p_1	p_2	$\dots$	p_n

По определению дисперсии

$$D(X) = M[X - M(X)]^2 = [x_1 - M(X)]^2 p_1 + [x_2 - M(X)]^2 p_2 + \dots + [x_n - M(X)]^2 p_n.$$

Таким образом, для того чтобы найти дисперсию, достаточно вычислить сумму произведений возможных значений квадрата отклонения на их вероятности.

З а м е ч а н и е. Из определения следует, что дисперсия дискретной случайной величины есть неслучайная (постоянная) величина.

В дальнейшем читатель узнает, что дисперсия непрерывной случайной величины также есть постоянная величина.

▷ ПРИМЕР 5.40. Найти дисперсию случайной величины X , которая задана следующим законом распределения:

X	1	2	5
p	0,3	0,5	0,2

Р е ш е н и е. Найдем математическое ожидание:

$$M(X) = 1 \cdot 0,3 + 2 \cdot 0,5 + 5 \cdot 0,2 = 2,3.$$

Найдем все возможные значения квадрата отклонения:

$$[x_1 - M(X)]^2 = (1 - 2,3)^2 = 1,69;$$

$$[x_2 - M(X)]^2 = (2 - 2,3)^2 = 0,09;$$

$$[x_3 - M(X)]^2 = (5 - 2,3)^2 = 7,29.$$

Напишем закон распределения квадрата отклонения:

$[X - M(X)]^2$	1,69	0,09	7,29
p	0,3	0,5	0,2

По определению

$$D(X) = 1,69 \cdot 0,3 + 0,09 \cdot 0,5 + 7,29 \cdot 0,2 = 2,01. \blacktriangleright$$

Вычисление, основанное на определении дисперсии, оказалось относительно громоздким. Рассмотрим формулу, которая приводит к цели значительно быстрее.

Формула для вычисления дисперсии. Для вычисления дисперсии часто бывает удобно пользоваться следующей теоремой.

□ ТЕОРЕМА 5.10. *Дисперсия равна разности между математическим ожиданием квадрата случайной величины X и квадратом ее математического ожидания:*

$$D(X) = M(X^2) - [M(X)]^2. \blacksquare$$

▷ ПРИМЕР 5.41. Найти дисперсию случайной величины X , которая задана следующим законом распределения:

X	2	3	5
p	0,1	0,6	0,3

Решение. Найдем математическое ожидание $M(X)$:

$$M(X) = 2 \cdot 0,1 + 3 \cdot 0,6 + 5 \cdot 0,3 = 3,5.$$

Напишем закон распределения случайной величины X^2 :

X^2	4	9	25
p	0,1	0,6	0,3

Найдем математические ожидания $M(X^2)$:

$$M(X^2) = 4 \cdot 0,1 + 9 \cdot 0,6 + 25 \cdot 0,3 = 13,3.$$

Искомая дисперсия

$$D(X) = M(X^2) - [M(X)]^2 = 13,3 - (3,5)^2 = 1,05. \blacktriangleright$$

Свойства дисперсии.

1. Дисперсия постоянной величины C равна нулю:

$$D(C) = 0.$$

2. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат:

$$D(CX) = C^2 D(X).$$

3. Дисперсия суммы (разности) двух независимых случайных величин равна сумме дисперсий этих величин:

$$D(X \pm Y) = D(X) + D(Y).$$

СЛЕДСТВИЕ 1. Дисперсия суммы нескольких взаимно независимых случайных величин равна сумме дисперсий этих величин.

СЛЕДСТВИЕ 2. Дисперсия суммы постоянной величины и случайной равна дисперсии случайной величины:

$$D(C + X) = D(X).$$

▷ **ПРИМЕР 5.42.** Случайные величины X и Y независимы. Найти дисперсию случайной величины $Z = 2X + 5Y$, если известно, что $D(X) = 5$, $D(Y) = 4$.

Решение. Так как величины X и Y независимы, то независимы также и величины $2X$ и $5Y$. Используя свойства дисперсии, получим

$$D(Z) = D(2X + 5Y) = D(2X) + D(5Y) = 4D(X) + 25D(Y) = 4 \cdot 5 + 25 \cdot 4 = 120. \blacktriangleright$$

4. Для любых случайных величин X и Y

$$D(X \pm Y) = D(X) + D(Y) \pm 2\text{cov}(X, Y),$$

где $C_{xy} = \text{cov}(X, Y) = M((X - m_x)(Y - m_y))$ — ковариация случайных величин X и Y , а $m_x = M(X)$, $m_y = M(Y)$.

Для двух случайных величин X и Y мерой их зависимости служит коэффициент корреляции

$$\rho_{xy} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)D(Y)}}.$$

Коэффициент корреляции обладает следующими свойствами:

1°. $-1 \leq \rho_{xy} \leq 1$ для любых случайных величин X и Y .

2°. Если $\rho_{xy} = 1$, то $Y = aX + b$, $a > 0$; если $\rho_{xy} = -1$, то $Y = aX + b$, $a < 0$.

3°. Если случайные величины X и Y независимы, то $r(X, Y) = 0$.

Случайные величины X и Y , для которых $\rho_{xy} = 0$, называются *некоррелированными*. Таким образом, в силу свойства 3° независимые случайные величины некоррелированы, обратное, вообще говоря, неверно. Отсюда следует, что свойство дисперсии 2 следует из свойства 4, если случайные величины *некоррелированы*, а тем более, — если они независимы.

Как используются ковариация C_{xy} и коэффициент корреляции ρ_{xy} в теории вероятностей и математической статистике мы покажем на примере в подпараграфе 6.1 главы 5.

Дисперсия характеризует средний квадрат отклонения дискретной случайной величины, поэтому на практике часто используют в качестве характеристики разброса *среднее квадратическое*

отклонение $\sigma(X) = \sqrt{D(X)}$, которое имеет ту же размерность, что и случайная величина X . Иногда требуется найти *коэффициент*

вариации $V = \frac{\sigma(X)}{M(X)}$, определяемый данным отношением.

4.2. Числовые характеристики непрерывных случайных величин

ОПРЕДЕЛЕНИЕ. *Математическим ожиданием непрерывной случайной величины X , возможные значения которой принадлежат отрезку $[a, b]$, называют определенный интеграл*

$$M(X) = \int_a^b x f(x) dx. \quad (5.18)$$

Если возможные значения принадлежат всей оси Ox , то

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Предполагается, что несобственный интеграл сходится абсолютно, т.е. существует интеграл

$$\int_{-\infty}^{\infty} |x| f(x) dx.$$

Если бы это требование не выполнялось, то значение интеграла зависело бы от скорости стремления (в отдельности) нижнего предела к $-\infty$, а верхнего — к $+\infty$.

Мода непрерывной случайной величины X будет определяться как максимум ее дифференциальной функции:

$$M_0(X) = \max_{(-\infty; +\infty)} f(x).$$

По аналогии с дисперсией дискретной величины определяется и дисперсия непрерывной величины.

ОПРЕДЕЛЕНИЕ. *Дисперсией непрерывной случайной величины называют математическое ожидание квадрата ее отклонения.*

Если возможные значения X принадлежат отрезку $[a, b]$, то

$$D(X) = \int_a^b [x - M(X)]^2 f(x) dx;$$

если возможные значения принадлежат всей оси Ox , то

$$D(X) = \int_{-\infty}^{\infty} [x - M(X)]^2 f(x) dx.$$

З а м е ч а н и е. Для вычисления дисперсии существуют более удобные формулы:

$$D(X) = \int_a^b x^2 f(x) dx - [M(X)]^2,$$

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x) dx - [M(X)]^2. \quad (5.19)$$

Среднее квадратическое отклонение непрерывной случайной величины определяется, как и для дискретной величины, равенством

$$\sigma(X) = \sqrt{D(X)}.$$

З а м е ч а н и е. Свойства математического ожидания и дисперсии дискретных величин сохраняются и для непрерывных величин.

▷ П Р И М Е Р 5.43. Найти математическое ожидание и дисперсию случайной величины X , заданной функцией распределения

$$F(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ x & \text{при } 0 < x \leq 1, \\ 1 & \text{при } x > 1. \end{cases}$$

Р е ш е н и е. Найдем плотность распределения:

$$f(x) = F'(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ 1 & \text{при } 0 < x \leq 1, \\ 0 & \text{при } x > 1. \end{cases}$$

Найдем математическое ожидание по формуле (5.18):

$$M(X) = \int_0^1 x \cdot 1 \cdot dx = \left. \frac{x^2}{2} \right|_0^1 = 1/2.$$

Найдем дисперсию по формуле (5.19):

$$D(X) = \int_0^1 x^2 \cdot 1 \cdot dx - \left(\frac{1}{2} \right)^2 = \left. \frac{x^3}{3} \right|_0^1 - 1/4 = 1/12. \blacktriangleright$$

§ 5. Некоторые специальные распределения случайных величин

Рассмотрим некоторые случайные величины, распределенные по конкретному закону, часто используемые в теории вероятностей, математической статистике и экономике.

1. Пуассоновское распределение. Пусть *дискретная* случайная величина X принимает бесконечное счетное число значений: 0, 1, 2, 3, 4, 5, ..., k , ... с вероятностью, определяющейся по *формуле Пуассона*:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

описывающей пуассоновское распределение, где $\lambda > 0$ — *параметр распределения Пуассона*:

x_i	0	1	2	...	k	...
p_i	$e^{-\lambda}$	$\lambda \cdot e^{-\lambda}$	$\frac{\lambda^2}{2!} \cdot e^{-\lambda}$	...	$\frac{\lambda^k}{k!} \cdot e^{-\lambda}$	...

Тогда эта случайная величина называется *пуассоновской* случайной величиной с параметром λ .

Основные характеристики этой случайной величины:

$$M(X) = D(X) = \lambda.$$

Заметим, что равенство дисперсии математическому ожиданию справедливо только для распределения Пуассона.

2. Геометрический закон распределения. Пусть $P(A) = p$ — вероятность наступления события A в каждом опыте, соответственно, $q = 1 - p$ — вероятность ненаступления события A .

Вероятность наступления события A в k -м опыте определяется по формуле

$$P(X = k) = p \cdot q^{k-1}. \quad (5.20)$$

Дискретная случайная величина X , распределенная по геометрическому закону, принимает значения 1, 2, ..., k , ... с вероятностью, определяемой по формуле (5.20):

x_i	1	2	3	...	k	...
P_i	$p \cdot q^0$	$p \cdot q^1$	$p \cdot q^2$	...	$p \cdot q^{k-1}$	...

▷ ПРИМЕР 5.44. Вероятность принятия бизнес-плана равна 0,6. Найти вероятность того, что бизнес-план будет принят при второй, третьей, k -й попытке.

Р е ш е н и е.

$$P(X=1) = p = 0,6;$$

$$P(X=2) = p \cdot q = 0,6 \cdot 0,4 = 0,24;$$

$$P(X=3) = p \cdot q^2 = 0,6 \cdot 0,4 \cdot 0,4 = 0,096;$$

.....

$$P(X=k) = p \cdot q^{k-1}.$$

x_i	1	2	3	...	k	...
P_i	0,6	0,24	0,096	...	$p \cdot q^{k-1}$	...

Сумма вероятностей, как и для других законов, равна единице:

$$S = \frac{p}{1-q} = \frac{0,6}{1-0,4} = \frac{0,6}{0,6} = 1 \text{ — согласно формуле суммы беско-}$$

нечной геометрической прогрессии со знаменателем $q < 1$. ►

3. Биномиальное распределение. Предположим, что проводится серия n независимых испытаний, причем каждое испытание имеет два исхода — “успех” (событие A) или “неудача” (событие $\bar{A}$), и пусть p — вероятность успеха в отдельном испытании, тогда $q = 1 - p$ — вероятность неудачи. Рассмотрим в качестве *дискретной* случайной величины X число появлений события A в этих испытаниях.

Поставим перед собой задачу найти закон распределения величины X . Для ее решения требуется определить возможные значения X и их вероятности. Очевидно, событие A в n испытаниях может либо не появиться, либо появиться один раз, либо два раза, ..., либо n раз. Таким образом, возможные значения X таковы:

$x_1 = 0, x_2 = 1, x_3 = 2, \dots, x_n = n - 1$. Остается найти вероятности этих возможных значений, для чего достаточно воспользоваться **формулой Бернулли**:

$$P_n(k) = C_n^k p^k q^{n-k}, \quad (5.21)$$

где $k = 0, 1, 2, \dots, n$.

Формула (5.21) является *аналитическим* выражением искомого закона распределения.

Биномиальным называют распределение вероятностей, определяемое формулой Бернулли. Закон называют биномиальным потому, что правую часть равенства (5.21) можно рассматривать как общий член разложения **бинома Ньютона**:

$$(p + q)^n = C_n^n p^n + C_n^{n-1} p^{n-1} q + \dots + C_n^k p^k q^{n-k} + \dots + C_n^0 q^n.$$

Таким образом, первый член разложения p^n определяет вероятность наступления рассматриваемого события n раз в n независимых испытаниях; второй член разложения $np^{n-1}q$ определяет вероятность наступления события $(n - 1)$ раз; ...; последний член q^n определяет вероятность того, что событие не появится ни разу.

Напишем биномиальный закон в виде таблицы:

X	n	$n-1$	$\dots$	k	$\dots$	0
p	p^n	$np^{n-1}q$	$\dots$	$C_n^k p^k q^{n-k}$	$\dots$	q^n

Основные характеристики этой случайной величины X :

$$M(X) = np, \quad D(X) = npq, \quad \sigma(X) = \sqrt{npq}.$$

При $n \rightarrow \infty$ и $p \rightarrow 0$ биномиальный закон приближается к закону распределения Пуассона, где $\lambda = np$.

4. Равномерное распределение. Будем считать, что случайная величина X распределена по *равномерному* (прямоугольному) закону, если все значения случайной величины лежат внутри некоторого интервала и все они равновероятны (точнее, обладают одной плотностью вероятности).

Непрерывная случайная величина X , плотность распределения которой задается формулой

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$

называется *равномерной* на отрезке $[a, b]$. Нетрудно проверить, что

$$M(X) = \frac{b+a}{2}, \quad D(X) = \frac{(b-a)^2}{12}.$$

5. Показательное (экспоненциальное) распределение. *Непрерывная* случайная величина X , плотность распределения которой задается формулой

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0, \end{cases}$$

называется *показательной* или *экспоненциальной* с параметром λ .

Распределение показательной случайной величины является *показательным распределением*.

В большом числе случаев показательное распределение описывает время безотказной работы прибора, при этом число λ интерпретируется как интенсивность отказа. Это распределение находит также широкое применение в демографии и в экономике.

Основные характеристики показательной случайной величины X :

$$M(X) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda^2}.$$

6. Нормальное (гауссовское) распределение. *Непрерывная* случайная величина X , плотность распределения которой задается формулой

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}, \quad (5.22)$$

называется *нормальной*, или *гауссовской*, с параметрами a и σ^2 . Часто используется обозначение $X \sim N(a, \sigma^2)$. Нормальная случайная величина с $a = 0$ и $\sigma^2 = 1$ называется *стандартной нормальной*

величиной. Если $X \sim N(a, \sigma^2)$, то случайная величина $Z = \frac{X - a}{\sigma}$

является стандартной.

Распределение вероятностей нормальной (гауссовской) случайной величины является **нормальным** (гауссовским) распределением.

Гауссовское распределение играет исключительно важную роль в теории вероятностей и математической статистике. Поэтому мы рассмотрим его более подробно, чем все остальные.

Мы видим, что нормальное распределение определяется двумя параметрами: a и σ . Достаточно знать эти параметры, чтобы задать нормальное распределение. Покажем, что вероятностный смысл этих параметров таков: a есть математическое ожидание, σ — среднее квадратическое отклонение нормального распределения.

1. По определению математического ожидания непрерывной случайной величины

$$M(X) = \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-(x-a)^2/(2\sigma^2)} dx.$$

Введем новую переменную $z = (x - a)/\sigma$. Отсюда $x = \sigma z + a$, $dx = \sigma dz$. Приняв во внимание, что новые пределы интегрирования равны старым, получим

$$M(X) = \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma z + a) e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma z e^{-z^2/2} dz + \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-z^2/2} dz.$$

Первое слагаемое равно нулю (под знаком интеграла нечетная функция; пределы интегрирования симметричны относительно начала координат). Второе слагаемое равно a , так как интеграл Пуассона

$$\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}.$$

Итак, $M(X) = a$, т.е. *математическое ожидание нормального распределения равно параметру a .*

2. По определению дисперсии непрерывной случайной величины, учитывая, что $M(X) = a$, имеем

$$D(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-a)^2 e^{-(x-a)^2/(2\sigma^2)} dx.$$

Введем новую переменную $z = (x - a)/\sigma$. Отсюда $x - a = \sigma z$, $dx = \sigma dz$. Приняв во внимание, что новые пределы интегрирования равны старым, получим

$$D(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz.$$

Обозначим $u = z$, $dv = ze^{-z^2/2} dz$. Интегрируя по частям (см. подпараграф 19.3 главы 4), найдем

$$D(X) = \sigma^2.$$

Следовательно,

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\sigma^2} = \sigma.$$

Итак, *среднее квадратическое отклонение нормального распределения равно параметру σ* .

ОПРЕДЕЛЕНИЕ. *Общим* называется нормальное распределение с произвольными параметрами a и σ ($\sigma > 0$).

ОПРЕДЕЛЕНИЕ. *Нормированным* (стандартным) называется нормальное распределение с параметрами $a = 0$ и $\sigma = 1$.

Например, если X — нормальная величина с параметрами a и σ , то $U = (X - a)/\sigma$ — нормированная нормальная величина, причем $M(U) = 0$, $\sigma(U) = 1$.

Плотность нормированного распределения

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Эта функция табулирована.

Функция $F(x)$ общего нормального распределения

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(z-a)^2/(2\sigma^2)} dz,$$

а функция нормированного распределения

$$F_0(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz.$$

Функция $F_0(x)$ табулирована.

Легко проверить, что

$$F(x) = F_0((x-a)/\sigma).$$

Учитывая, что $\int_{-\infty}^{\infty} f(x)dx = 1$ и, в силу симметрии $f(x)$ относи-

тельно нуля

$$\int_{-\infty}^0 f(x)dx = 0,5 \text{ (т.е. } P(-\infty < X < 0) = 0,5), \text{ легко получить, что}$$

$$F_0(x) = 0,5 + \Phi(x),$$

$$\text{где } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz \text{ — функция Лапласа.}$$

Действительно,

$$F_0(x) = P(-\infty < X < x) = P(-\infty < X < 0) + P(0 < X < x) = 0,5 + \Phi(x).$$

Нормальная кривая. График плотности нормального распределения называют *нормальной кривой* (*кривой Гаусса*) (рис. 5.9).

Исследуем функцию

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}$$

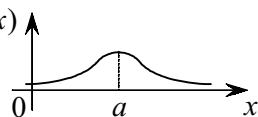


Рис. 5.9

методами дифференциального исчисления.

1. Очевидно, что функция определена на всей оси Ox .
2. При всех значениях x функция принимает положительные значения, т.е. нормальная кривая расположена над осью Ox .
3. Предел функции при неограниченном возрастании x (по абсолютной величине) равен нулю: $\lim_{|x| \rightarrow \infty} y(x) = 0$, т.е. ось Ox служит горизонтальной асимптотой графика.
4. Исследуем функцию на экстремум. Найдем первую производную:

$$y' = -\frac{x-a}{\sigma^3\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)}.$$

Легко видеть, что $y' = 0$ при $x = a$, $y' > 0$ при $x < a$, $y' < 0$ при $x > a$. Следовательно, при $x = a$ функция имеет максимум, равный $\frac{1}{\sigma\sqrt{2\pi}}$.

5. Разность $x - a$ содержится в аналитическом выражении функции в квадрате, т.е. график функции симметричен относительно прямой $x = a$.

6. Исследуем функцию на точки перегиба. Найдем вторую производную:

$$y'' = \frac{1}{\sigma^3\sqrt{2\pi}} e^{-(x-a)^2/(2\sigma^2)} \left[1 - \frac{(x-a)^2}{\sigma^2} \right].$$

Легко видеть, что при $x = a + \sigma$ и $x = a - \sigma$ вторая производная равна нулю, а при переходе через эти точки она меняет знак (в обеих этих точках значение функции равно $1/(\sigma\sqrt{2\pi e})$). Таким образом, точки графика $(a - \sigma, 1/(\sigma\sqrt{2\pi e}))$ и $(a + \sigma, 1/(\sigma\sqrt{2\pi e}))$ являются точками перегиба.

Вероятность попадания нормальной случайной величины в заданный интервал. Как известно, если случайная величина X задана плотностью распределения $f(x)$, вероятность того, что X примет значение, принадлежащее интервалу (α, β) , такова:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx.$$

Пусть случайная величина X распределена по нормальному закону. Тогда вероятность того, что X примет значение, принадлежащее интервалу (α, β) , равна

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-(x-a)^2/(2\sigma^2)} dx.$$

Преобразуем эту формулу так, чтобы можно было пользоваться готовыми таблицами. Так, например, табулирована функция Лапласа. Введем новую переменную $z = (x - a)/\sigma$. Отсюда $x = \sigma z + a$, $dx = \sigma dz$. Найдем новые пределы интегрирования. Если $x = \alpha$, то $z = (\alpha - a)/\sigma$; если $x = \beta$, то $z = (\beta - a)/\sigma$.

Таким образом, имеем

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^{(\beta-a)/\sigma} e^{-z^2/2} \sigma dz = \frac{1}{\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^0 e^{-z^2/2} dz + \\ + \frac{1}{\sqrt{2\pi}} \int_0^{(\beta-a)/\sigma} e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_0^{(\beta-a)/\sigma} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_0^{(\alpha-a)/\sigma} e^{-z^2/2} dz.$$

Пользуясь функцией Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz,$$

окончательно получим

$$P(\alpha < X < \beta) = \Phi((\beta - a)/\sigma) - \Phi((\alpha - a)/\sigma). \quad (5.23)$$

7. Логарифмически нормальное распределение. Пусть $X \sim N(a, \sigma^2)$. Случайная величина $Y = ex$ называется *логарифмически нормальной*. Можно показать, что плотность распределения этой величины определяется формулой

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln x - a)^2}{2\sigma^2}\right), \quad x > 0; \quad p(x) = 0, \quad x \leq 0. \quad (5.24)$$

Логарифмически нормальное распределение возникает в ситуации, когда исследуемая случайная величина формируется под воздействием большого числа мультипликативных случайных факторов, например при исследовании фондовых рынков. Можно показать, что

$$M(Y) = \exp\left(a + \frac{\sigma^2}{2}\right), \quad D(Y) = \exp(2a + \sigma^2)[\exp(\sigma^2) - 1].$$

8. χ^2 -распределение. Пусть $X_1, \dots, X_n$ — независимые стандартные нормальные случайные величины (т.е. $X_i \sim N(0,1)$, $i = 1, \dots, n$).

Говорят, что случайная величина $\chi^2(n) = \sum_{i=1}^n X_i^2$ имеет χ^2 -распределение с n степенями свободы. Плотность распределения величины $\chi^2(n)$ задается формулой

¹ Запись вида $\exp(t)$ равносильна записи e^t .

$$f(x) = \frac{1}{2^{n/2} \Gamma(n/2)} \left(\frac{x}{2}\right)^{\frac{n}{2}-1} \cdot \exp\left(-\frac{x}{2}\right), \quad x > 0; \quad f(x) = 0, \quad x \leq 0,$$

где Γ — гамма-функция, определяемая равенством

$$\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx, \quad t > 0, \text{ в частности, } \Gamma(k) = (k-1)! \text{ для целых } k.$$

Нетрудно проверить, что

$$M(\chi^2) = n, \quad D(\chi^2) = 2n.$$

9. Распределение Стьюдента (t -распределение). Пусть $X_0, X_1, \dots, X_n$ — независимые стандартные нормальные случайные величины. Распределение случайной величины

$$t(n) = \frac{X_0}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}} = \frac{X_0}{\sqrt{\frac{1}{n} \chi^2(n)}}$$

называется *распределением Стьюдента*, или t -распределением с n степенями свободы. Плотность распределения величины $t(n)$ задается формулой

$$f(x) = \frac{\Gamma((n+1)/2)}{(n\pi)^{1/2} \Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-\frac{(n+1)}{2}}.$$

При $n = 1$ соответствующее распределение носит название *распределение Коши*.

Можно показать, что при $n > 1$

$$M(t(n)) = 0, \quad D(t(n)) = \frac{n}{n-2}.$$

10. Распределение Фишера (F -распределение). Пусть $X_1, \dots, X_m, Y_1, \dots, Y_n$ — совокупность независимых стандартных нормальных случайных величин. Распределение случайной величины

$$F(m, n) = \frac{\frac{1}{m} \sum_{i=1}^m X_i^2}{\frac{1}{n} \sum_{i=1}^n Y_i^2} = \frac{\frac{1}{m} \chi^2(m)}{\frac{1}{n} \chi^2(n)}$$

называется *распределением Фишера* или *F-распределением* с (m, n) *степенями свободы*. Плотность распределения величины $F(m, n)$ задается формулой

$$f(x) = \frac{\Gamma[(m+n)/2] m^{\frac{m}{2}} n^{\frac{n}{2}}}{\Gamma(m/2)\Gamma(n/2)} \cdot \frac{x^{\frac{m}{2}-1}}{(mx+n)^{\frac{(m+n)}{2}}}, \quad x > 0; \quad f(x) = 0,$$

$$x \leq 0.$$

Распределения χ^2 , Стьюдента и Фишера активно применяются в статистике при построении доверительных интервалов для оцениваемых параметров и при проверке статистических гипотез.

§ 6. Некоторые аспекты из области математической статистики

6.1. Основные задачи математической статистики. Измерение риска, присущего инвестиционному портфелю

Мы не ставим своей целью дать в рамках данного учебника полный содержательный курс теории вероятностей и математической статистики (ему посвящены замечательные работы, например, [4], [5], [11]), но здесь хотим лишь проинформировать читателей, показать на конкретных примерах важность изучения математической статистики для будущих экономистов.

В подпараграфе 1.4 главы 5 мы говорили о *портфеле* инвестиций и перечислили несколько видов инвестиций. Главное, что отличает все эти виды инвестиций — *степень риска*. Наименее рискованными инвестициями, которые вы можете осуществить, являются денежные вложения в государственные ценные бумаги (ГКО). Риск непогашения этих бумаг минимален, а короткий срок означает относительную стабильность цен на них.

По сути, инвестор, который хочет дать займы, скажем на 3 месяца, может обеспечить себе гарантированный доход, купив ГКО со сроком 3 месяца. Однако при этом инвестор не может быть уверен в том, какова их *реальная* норма доходности (см. §3 главы 3, подпараграф 23.4 главы 4), поскольку остается неопределенность, порождаемая инфляцией, действиями правительства, общей экономической ситуацией и т.д.

Формируя портфель из долгосрочных правительственных облигаций, инвестор приобретает активы, цены на которые изменяются с изменением процентных ставок. (Цены на облигации падают, когда процентные ставки растут, и наоборот.)

Переходя от правительственных облигаций к облигациям компаний, инвестор берет на себя дополнительный риск, связанный с *невыполнением обязательств по ним (неплатежеспособностью)* конкретных юридических лиц. Приобретая обыкновенные акции, инвестор разделяет риски предприятия, акции которого он приобрел.

Математическая статистика позволяет обосновать ответ на вопросы: о случайности и закономерности изучаемого явления (влияет ли инфляция на доходность¹ акций); как зависит результативный признак от факторного (как зависит доходность акций от уровня инфляции при прочих равных условиях); сколько необходимо провести наблюдений для объективного суждения об изучаемом явлении; какой фактор сильнее влияет на результат (влияние результатов парламентских выборов или цен на нефть на доходность акций).

Методы математической статистики можно разделить на *описательные* и *аналитические*.

Первая задача математической статистики — указать способы сбора и группировки статистических сведений, полученных в результате наблюдений или в результате специально поставленных экспериментов; *описать* реальные наблюдения с помощью таблиц, графиков, характеристик положения (среднее арифметическое, мода), характеристик рассеяния (среднее квадратическое отклонение, дисперсия) и т.д.

Вторая задача математической статистики — разработать методы *анализа* статистических данных в зависимости от целей исследования. Сюда относятся:

а) оценка неизвестной вероятности события; оценка неизвестной функции распределения; оценка параметров распределения, вид которого известен; оценка зависимости случайной величины от одной или нескольких случайных величин и др.;

б) проверка статистических гипотез о виде неизвестного распределения или о величине параметров распределения, вид которого известен.

Аналитические методы позволяют на основании выборочных наблюдений сделать статистически значимые выводы о наличии закономерностей для всей совокупности.

Современная математическая статистика разрабатывает способы определения числа необходимых испытаний до начала ис-

¹ Доходность — отношение величины получаемого дохода за период времени к цене акции.

следования (планирование эксперимента), в ходе исследования (последовательный анализ) и решает многие другие задачи. Современную математическую статистику определяют как науку о принятии решений в условиях неопределенности.

Итак, задача математической статистики состоит в создании методов сбора и обработки статистических данных для получения научных и практических выводов.

ОПРЕДЕЛЕНИЕ. *Генеральной совокупностью* называют совокупность объектов, из которых производится выборка.

ОПРЕДЕЛЕНИЕ. Последовательность наблюдений $X_1, \dots, X_n$ называется *случайной выборкой объема n* (или *выборочной совокупностью*), если $X_1, \dots, X_n$ получены как независимые реализации некоторой случайной величины X с распределением $F(x)$. При этом также говорят, что $X_1, \dots, X_n$ есть *выборка из генеральной совокупности X* (или $F(x)$). С теоретико-вероятностной точки зрения случайная выборка $X_1, \dots, X_n$ может рассматриваться как последовательность независимых случайных величин, имеющих одно и то же распределение $F(x)$.

ОПРЕДЕЛЕНИЕ. *Объемом совокупности* (выборочной или генеральной) называют число объектов этой совокупности.

Например, если из 50 акций 50 различных компаний отобрано для финансового анализа 10 акций, то объем генеральной совокупности $N = 50$, а объем выборки $n = 10$.

В реальных социально-экономических системах данные обычно представляют собой наблюдения за происходящим процессом, например курсом валюты на бирже в течение месяца. Результаты наблюдений в общем случае — ряд чисел, расположенных в беспорядке, который для изучения необходимо упорядочить (проранжировать).

ОПРЕДЕЛЕНИЕ. Операция, заключенная в расположении значений признака по неубыванию, называется *ранжированием* опытных данных.

После операции ранжирования опытные данные можно сгруппировать так, чтобы в каждой группе признак воспринимал одно и то же значение, которое называется *вариантом* (x_i). Число элементов в каждой группе называется *частотой* варианта (n_i). Пусть из генеральной совокупности извлечена выборка, причем x_1 наблюдалось n_1 раз, x_2 — n_2 раз, x_k — n_k раз и $\sum_{i=1}^k n_i = n$ — объем выборки. Отношение частот к объему выборки $\frac{n_i}{n} = W_i$ называют *относительными частотами*.

ОПРЕДЕЛЕНИЕ. Последовательность вариантов, расположенных в возрастающем порядке, называется *вариационным рядом* (вариация — изменение).

ОПРЕДЕЛЕНИЕ. Перечень вариантов и соответствующих им частот, или относительных частот, называют *статистическим распределением выборки*.

Заметим, что в теории вероятностей под *распределением* понимают соответствие между возможными значениями случайной величины и их вероятностями, а в математической статистике — соответствие между наблюдаемыми вариантами и их частотами, или относительными частотами.

Вариационные ряды изображают графически с помощью полигона и гистограммы.

ОПРЕДЕЛЕНИЕ. *Полигон частот* — это ломаная, отрезки которой соединяют точки $(x_1; n_1)$, $(x_2; n_2)$, ..., $(x_k; n_k)$.

ОПРЕДЕЛЕНИЕ. *Полигон относительных частот* — это ломаная, отрезки которой соединяют точки $(x_1; \frac{n_1}{n})$, $(x_2; \frac{n_2}{n})$, ..., $(x_k; \frac{n_k}{n})$.

ОПРЕДЕЛЕНИЕ. *Гистограммой частот* называется фигура, состоящая из прямоугольников с основанием h и высотами n_i . Для *гистограммы относительных частот* в качестве высоты принимают $\frac{n_i}{n}$.

Гистограмма относительных частот является аналогом дифференциальной функции случайной величины (плотности распределения).

Вернемся к вопросу об измерении риска, присущего инвестиционному портфелю. Пусть у вас есть два значения, от которых вы можете оттолкнуться. Вы знаете ставку дисконта для безопасных проектов и ставку для проектов со “средним риском”. Но вы *не знаете* пока, как вычислить ставки дисконта для активов, не вписывающихся в эти простые случаи. Для того чтобы их определить, вы должны знать: 1) как измерить риск и 2) какова связь между возникновением риска и требуемыми премиями за риск.

На рис. 5.10 показаны среднегодовые нормы доходности акций компании А, рассчитанные для фондового рынка в некоторой гипотетической экономике. Колебания доходности от года к году очень значительны. Самое большое значение годовой доходности 54,0% наблюдалось в 1934 г. Однако, за 4 года произошло снижение более чем на 25%, в 1932 г. отмечалось самое низкое значение доходности, которое составило –43,3%.

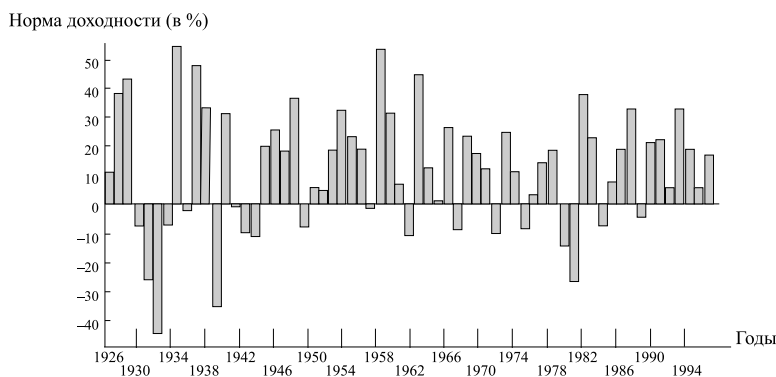


Рис. 5.10

Другой способ представления данных — гистограмма, или частотное распределение. Это показано на рис. 5.11, где изменчивость доходности от года к году представлена широким “разбросом” результатов.

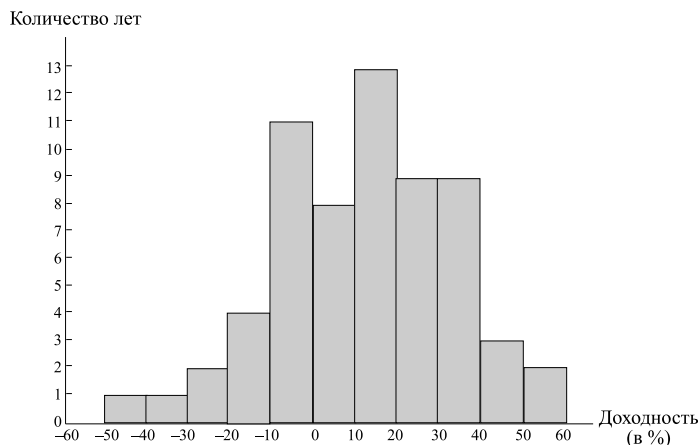


Рис. 5.11

Дисперсия и стандартное отклонение. Стандартными статистическими показателями разброса результатов служат *дисперсия* и *стандартное отклонение*. Дисперсия рыночной доходности $D(X)$ представляет собой ожидаемое отклонение от ожидаемой доходности в квадрате. Это можно выразить так

$$\text{Дисперсия } (X) = \text{ожидаемое значение } (X_i - \bar{X})^2,$$

где X_i — фактическая доходность, $\bar{X}$ — ожидаемая доходность, или иначе

$$D(X) = \sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2, \quad (5.25)$$

где $\bar{X}$ есть среднее арифметическое значений фактической доходности:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Стандартное отклонение равно квадратному корню из дисперсии:

$$\sigma = \sqrt{D(X)}.$$

Формула (5.25) верна, если все значения $X_1, \dots, X_n$ признака выборки объема n различны. Если же значения признака $X_1, \dots, X_k$ имеют соответственно частоты $n_1, n_2, \dots, n_k$, причем $n_1 + n_2 + n_3 + \dots + n_n = n$, то

$$D(X) = \sigma^2 = \frac{1}{n} \sum_{i=1}^n n_i (X_i - \bar{X})^2 = \sum_{i=1}^n W_i (X_i - \bar{X})^2.$$

Приведем очень простой пример, показывающий, как вычисляются дисперсия и стандартное отклонение. Допустим, что вам представилась возможность сыграть в следующую игру. Сначала вы инвестируете 100 усл. ед. Затем подбрасываете две монеты. Если выпадет “орел” — прибавляете к первоначальной сумме 20%, если “решка” — отнимаете 10%. Очевидно, существует четыре вероятных результата:

“орел” + “орел”: +40%;
 “орел” + “решка”: +10%;
 “решка” + “орел”: +10%;
 “решка” + “решка”: –20%.

Составим таблицу распределения частот:

X_i	+40	+10	–20
W_i	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Относительная частота равна 1 к 4 (или 0,25), что вы получите 40%, равна 2 к 4 (или 0,5), что вы получите 10%, и 1 к 4 (или 0,25), что вы потеряете 20%. Ожидаемая доходность игры, следовательно, представляет собой средневзвешенную значений фактической доходности:

$$\bar{X} = (0,25 \cdot 40) + (0,5 \cdot 10) + (0,25 \cdot -20) = 10\%.$$

Таблица 5.1

(1) Процентная норма доходности X_i	(2) Отклонение ожидаемой доходности $X_i - \bar{X}$	(3) Квадрат отклонения $(X_i - \bar{X})^2$	(4) Относительная частота W_i	(5) W_i × $(X_i - \bar{X})^2$
+40	+30	900	0,25	225
+10	0	0	0,5	0
–20	–30	900	0,25	225
$D(X) = \sigma^2 = \frac{1}{n} \sum_{i=1}^n n_i (X_i - \bar{X})^2 = \sum_{i=1}^n W_i (X_i - \bar{X})^2 = 450$ $\sigma = \sqrt{D(X)} = \sqrt{450} \approx 21,21$				

Из таблицы видно, что дисперсия процентных доходов составляет 450. Стандартное отклонение — корень квадратный из 450 — приблизительно равно 21,21. Норма доходности имеет те же единицы измерения, поэтому мы можем сказать, что изменчивость результатов игры составляет 21,21%.

Один из способов охарактеризовать неопределенность — сказать, что событий происходит меньше, чем можно ожидать. Риск, присущий активам, можно точно выразить описанием всех возможных результатов и вероятности их возникновения, как мы делали в игре с подбрасыванием монет. Однако для реальных активов это сделать трудно, а часто и невозможно. Поэтому мы используем дисперсию и стандартное отклонение, чтобы описать разброс возможных результатов¹.

Эти показатели и являются естественными измерителями риска. Если бы исход с подбрасыванием монет был предопределен, стандартное отклонение равнялось бы нулю. Фактическое стандартное отклонение будет положительно, поскольку мы не знаем, что произойдет на самом деле. Рассмотрим вторую игру, похожую на первую, за исключением того, что теперь при выпадении “орла” прибавляется 35%, при выпадении “решки” отнимается 25%. И опять возможны следующие четыре исхода:

“орел” + “орел”:	+70%;
“орел” + “решка”:	+10%;
“решка” + “орел”:	+10%;
“решка” + “решка”:	–50%.

В данной игре ожидаемая доходность, как и в первом случае, равна 10%, но стандартное отклонение вдвое больше — 42,43% против 21,21% в первой игре. Это говорит о том, что вторая игра в два раза рискованнее первой.

Вычисление портфельного риска. Если есть две случайные выборки $X_1, \dots, X_n$ и $Y_1, \dots, Y_n$, то можно определить:

¹ Какой из этих показателей использовать — исключительно вопрос удобства. Так как стандартное отклонение выражается в тех же единицах, что и норма доходности, в целом, этот показатель использовать удобнее. Однако, когда мы говорим о доле риска, объясняемого каким-то фактором, обычно менее опасно пользоваться показателем дисперсии.

ковариацию:

$$\text{cov}(X, Y) = C_{XY} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}); \quad (5.26)$$

коэффициент корреляции:

$$\rho_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y}. \quad (5.27)$$

Необходимость введения этих понятий (см. подпараграф 4.2. главы 5) и их пояснение представим примером.

Предположим, что портфель состоит на 60% из акций компании *A* и на 40% из акций компании *B*. Вы ожидаете, что в наступающем году доходность акций компании *A* составит 15%, а акций компании *B* — 21%. Ожидаемая доходность вашего инвестиционного портфеля представляет собой средневзвешенную ожидаемых значений доходности отдельных акций¹:

Ожидаемая доходность портфеля = $(0,60 \cdot 15) + (0,40 \cdot 21) = 17,4\%$.

Расчет ожидаемой доходности портфеля — достаточно легкая процедура. Самая тяжелая часть работы — это определить риск портфеля. Допустим, в прошлом стандартное отклонение доходности составляло для акций компании *A* примерно 28%, для акций компании *B* — примерно 42%. Вы полагаете, что эти цифры по-прежнему служат верным показателем отклонения возможных *будущих* доходов. Скорее всего, сначала вы будете склонны допустить, что стандартное отклонение доходности вашего портфеля представляет собой средневзвешенную стандартных отклонений доходности отдельных акций, т.е.

$(0,60 \cdot 28) + (0,40 \cdot 42) = 33,6\%$.

Это было бы верно, *только* если цены двух видов акций изменялись бы совершенно одинаково.

¹ Давайте это проверим. Предположим, вы инвестируете 60 усл. ед. в акции компании *A* и 40 долл. в акции компании *B*. Ожидаемый долларовый доход по вашим акциям *A* равен $0,15(60) = 9,00$ усл. ед., и по акциям *B* — $0,21(40) = 8,40$ усл. ед. Ожидаемый долларовый доход от вашего портфеля составляет $9,00 + 8,40 = 17,40$ усл. ед. *Норма* доходности портфеля равна $17,40/100 = 0,174$, или 17,4%.

	Акция 1	Акция 2
Акция 1	$x_1^2 \sigma_1^2$	$x_1 x_2 C_{12} = x_1 x_2 \rho_{12} \sigma_1 \sigma_2$
Акция 2	$x_1 x_2 C_{12} = x_1 x_2 \rho_{12} \sigma_1 \sigma_2$	$x_2^2 \sigma_2^2$

Рис. 5.12

На рис. 5.12 представлена процедура точного вычисления риска портфеля, состоящего из двух акций. Вам нужно заполнить таблицу из 4 прямоугольников. В верхнем левом прямоугольнике вы взвешиваете дисперсию доходности акции 1 (σ_1^2) по *квадрату* доли инвестиций в акции 1 (x_1^2). Заполняя нижний правый прямоугольник, вы взвешиваете дисперсию доходности акции 2 (σ_2^2) по *квадрату доли* инвестиций в акции 2 (x_2^2). Дисперсия для портфеля из двух акций равна сумме значений в этих четырех прямоугольниках. Значение x_i равно доле инвестиций в акции i ; σ_i^2 — дисперсия доходности акций i ; C_{ij} — *ковариация* доходности акций i и j ($\rho_{ij} \sigma_i \sigma_j$); ρ_{ij} — *корреляция* доходности акций i, j .

Элементы прямоугольников, расположенных по этой диагонали, зависят от дисперсий акций 1 и 2, элементы двух других прямоугольников зависят от их *ковариаций*. Как вы можете предположить, ковариация служит для измерения степени совместной изменчивости двух акций. Ковариация может быть выражена умножением коэффициента *корреляции* ρ_{12} , который характеризует степень линейной функциональной зависимости между случайными величинами, на два стандартных отклонения¹:

$$C_{12} = \rho_{12} \sigma_1 \sigma_2.$$

Для двух случайных величин коэффициент корреляции имеет следующие свойства:

- 1) $-1 \leq \rho \leq 1$;
- 2) если $\rho = \pm 1$, то между случайными величинами X и Y существует функциональная линейная зависимость;
- 3) если $\rho = 0$, то случайные величины X и Y некоррелированы, что не означает независимости вообще;

¹ Отметим, что ковариация любой ценной бумаги с ней самой равна просто ее дисперсии.

4) если X и Y образуют систему нормально распределенных случайных величин, то из их некоррелированности следует их независимость.

Раздел математической статистики, изучающий методы оценки выборочного коэффициента корреляции, проверку его значимости, называется *корреляционным анализом*.

Большинство же акций имеет тенденцию к одновременному изменению. В этом случае коэффициент корреляции ρ_{12} положителен, следовательно, положителен и коэффициент ковариации σ_{12} . Если различные акции движутся совершенно не связанно, тогда коэффициент корреляции и ковариация равны нулю; если акции изменяются в противоположных направлениях, коэффициент корреляции и ковариация отрицательны. Вы взвешиваете дисперсии умножением на квадрат доли инвестиции, таким же образом вы должны взвесить ковариацию умножением на *произведение* двух соответствующих вложений в x_1 и x_2 .

Когда вы заполните все четыре прямоугольника, вы просто складываете полученные в них величины и находите дисперсию портфеля:

$$\text{Дисперсия портфеля} = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2(x_1 x_2 \rho_{12} \sigma_1 \sigma_2).$$

Стандартное отклонение портфеля равно корню квадратному из дисперсии.

Теперь мы можем попробовать привести некоторые цифровые примеры для акций компаний A и B . Заметим, что при совершенной корреляции двух акций значение стандартного отклонения состоящего из них портфеля будет находиться на 40%-ной отметке промежутка между стандартными отклонениями этих акций (или, иными словами, стандартное отклонение портфеля будет превышать меньшее из стандартных отклонений двух акций на 40% от разности между этими стандартными отклонениями).

Давайте проверим это, заполнив четыре прямоугольника при условии $\rho_{12} = +1$.

	A	B
A	$x_1^2 \sigma_1^2 = (0,60)^2 \cdot (28)^2$	$x_1 x_2 \rho_{12} \sigma_1 \sigma_2 = 0,6 \cdot 0,40 \cdot 1 \cdot 28 \cdot 42$
B	$x_1 x_2 \rho_{12} \sigma_1 \sigma_2 = 0,6 \cdot 0,40 \cdot 1 \cdot 28 \cdot 42$	$x_2^2 \sigma_2^2 = (0,40)^2 \cdot (42)^2$

Дисперсия вашего портфеля акций будет равна сумме полученных значений:

$$\text{Дисперсия портфеля} = [(0,60)^2 \cdot (28)^2] + [(0,40)^2 \cdot (42)^2] + 2(0,60 \times 0,40 \cdot 1 \cdot 28 \cdot 42) = 1129.$$

Стандартное отклонение равно $\sqrt{1129} = 33,6\%$, или превышает 28 на 40% от разности между 42 и 28.

Акции компаний *A* и *B* изменяются не совершенно одинаково. Если исходить из прошлого опыта, то коэффициент корреляции между двумя акциями составляет примерно 0,4. Если мы проделаем те же вычисления при условии, что $\rho_{12} = +0,4$, то обнаружим:

$$\text{Дисперсия портфеля} = [(0,60)^2 \cdot (28)^2] + [(0,40)^2 \cdot (42)^2] + 2(0,60 \times 0,40 \cdot 0,4 \cdot 28 \cdot 42) = 790,27.$$

Стандартное отклонение равно $\sqrt{790} = 28,1\%$. Теперь величина риска отличается от 28 значительно меньше чем на 40% от разности между 42 и 28, на самом деле она *практически* равна риску при условии инвестирования только в акции компании *A*.

В качестве иллюстрации давайте предположим, что взаимосвязь между акциями компаний *A* и *B* такова, что коэффициент корреляции отрицателен. Поскольку мы представляем нереальную ситуацию, пойдем еще дальше и допустим, что между акциями существует совершенно отрицательная корреляция ($\rho_{12} = -1$). В этом случае:

$$\text{Дисперсия портфеля} = [(0,60)^2 \cdot (28)^2] + [(0,40)^2 \cdot (42)^2] + 2(0,60 \times 0,40 \cdot (-1) \cdot 28 \cdot 42) = 0.$$

При отрицательной корреляции всегда существует стратегия формирования портфеля (представленная особым набором акций в нем), позволяющая полностью исключить риск¹. Очень жаль, что в реальности такой совершенно отрицательной корреляции между обыкновенными акциями не бывает.

Чтобы определить дисперсию портфеля, состоящую из *N* акций, необходимо заполнить матричную таблицу, подобную той,

¹ Поскольку стандартное отклонение по акциям компании *B* в 1,5 раза превышает отклонение по акциям компании *A*, чтобы исключить риск для портфеля из двух данных акций, необходимо инвестировать в 1,5 раза больше средств в акции компании *A*.

что изображена на рис. 5.13. Квадраты, расположенные по диагонали, указывают на значения дисперсии ($x_i^2 \sigma_i^2$), а недиагональные квадраты — на значения ковариации ($x_i x_j \sigma_{ij}$).

Каждый прямоугольник на диагонали — затемненные квадраты на рис. 5.13 — содержит значение дисперсии, взвешенной по квадрату доли инвестиций в соответствующие ценные бумаги. Остальные квадраты содержат информацию о ковариации между двумя ценными бумагами, взвешенной по произведению соответствующих долей инвестиций¹.

Обратили ли вы внимание, глядя на рис. 5.13, как вырастает значение ковариации при увеличении количества ценных бумаг в портфеле? Если мы берем портфель, состоящий из двух видов ценных бумаг, количество квадратов в таблице со значениями дисперсий равно количеству квадратов с ковариацией. Если количество разных ценных бумаг больше двух, тогда квадратов с ковариацией гораздо больше, чем квадратов с дисперсией. Следовательно, судить об изменчивости портфеля следует главным образом по ковариации.

Рассмотрим портфель с равными долями инвестиций в N акций. Следовательно, в каждую акцию инвестируется $1/N$ совокупных инвестиций. Следовательно, дисперсия в каждом квадрате равна $(1/N)^2$ общей дисперсии, а ковариация — $(1/N)^2$ общей ковариации. Имеется N квадратов с дисперсиями и $N^2 - N$ квадратов с ковариацией. Следовательно:

Количество акций

	1	2	3	4	5	6	7	N
1								
2								
3								
4								
5								
6								
7								
N								

Рис. 5.13

¹ Формула эквивалентна “сложению всех квадратов”:

$$\text{Дисперсия портфеля} = \sum_{i=1}^N \sum_{j=1}^N x_i x_j C_{ij}.$$

Отметим, что когда $i = j$, значение C_{ij} просто равно дисперсии акции i .

$$\begin{aligned} \text{Дисперсия портфеля} &= N \left(\frac{1}{N} \right)^2 \times \text{средняя дисперсия} + \\ &+ (N^2 - N) \left(\frac{1}{N} \right)^2 \times \text{средняя ковариация} = \left(\frac{1}{N} \right) \times \text{средняя дисперсия} + \\ &+ \left(1 - \frac{1}{N} \right) \times \text{средняя ковариация}. \end{aligned}$$

Заметим, что если число N возрастает, то значение дисперсии портфеля почти приближается к среднему значению ковариации. Если бы средняя ковариация равнялась нулю, то можно было бы *полностью* избежать риска, располагая достаточным количеством ценных бумаг. К сожалению, обычные акции изменяются независимо друг от друга. Большинство акций, которые может приобрести инвестор, связаны друг с другом, т.е. имеют положительную ковариацию, которая ограничивает эффект увеличения количества акций портфеля.

▷ П Р И М Е Р 5.45. Стандартное отклонение доходности акций компании A за период t составляет 30%, акций компании B за тот же период — 20%. Коэффициент корреляции $\rho = 0,5$. Определить риск портфеля (стандартное отклонение доходности за период t), если инвестор купил акций компании A на 5000 руб., B — на 20 000 руб.

Р е ш е н и е. Инвестор всего купил акций на 25 000 руб. Тогда доля акций компании A от общей суммы составляет 20%. Доля акций компании B — 80%. По условию $\sigma_1 = 0,3$, $\sigma_2 = 0,2$. Дисперсию портфеля вычислим по формуле:

$$\sigma^2 = x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2(x_1 x_2 \rho_{12} \sigma_1 \sigma_2).$$

Тогда риск портфеля

$$\begin{aligned} \sigma &= \sqrt{x_1^2 \sigma_1^2 + x_2^2 \sigma_2^2 + 2(x_1 x_2 \rho_{12} \sigma_1 \sigma_2)} = \\ &= \sqrt{20^2 \cdot 0,3^2 + 80^2 \cdot 0,2^2 + 2 \cdot 20 \cdot 80 \cdot 0,3 \cdot 0,2 \cdot 0,5} \approx 19,7\%. \blacktriangleright \end{aligned}$$

Оценивание параметров. Предположим, что распределение случайной величины X (генеральной совокупности) зависит от

некоторого (возможно, многомерного) неизвестного параметра θ : $F(x) = F(x; \theta)$, $\theta \in \Theta \subseteq \mathbf{R}^n$. Общая задача оценивания заключается в получении каких-либо выводов о параметре θ на основании наблюдений $X_1, \dots, X_n$. Различают точечное и интервальное оценивание.

Любая функция $\varphi_n: \mathbf{R}^n \rightarrow \Theta$ называется *точечной оценкой* (или просто оценкой) параметра θ . Часто используется обозначение $\hat{\theta} = \varphi_n(X_1, \dots, X_n)$. В русскоязычной литературе по статистике, как правило, одним и тем же термином “оценка” называют как функцию φ_n , так и ее значение $\hat{\theta}$ для конкретных наблюдений $X_1, \dots, X_n$. В английском языке эти объекты различают, называя φ_n — *estimator*, а величину $\hat{\theta}$ — *estimation*. Поэтому правильнее было бы называть функцию φ_n методом оценивания, сохранив название “оценка” за величиной $\hat{\theta}$, однако такая терминология не является общепринятой. Заметим, что оценка, являясь функцией случайных наблюдений, также есть случайная величина.

Множество $D_n(X_1, \dots, X_n) \subseteq \Theta$ называется *доверительным множеством с уровнем доверия α* (или $100\alpha\%$ — доверительным множеством), если $P(\theta \in D_n(X_1, \dots, X_n)) = \alpha$, где $0 < \alpha < 1$. Часто это множество называют *интервальной оценкой* параметра θ с уровнем доверия α . Термин “интервальная” связан с тем, что в случае одномерного параметра в качестве доверительных множеств рассматриваются, как правило, интервалы.

Проверка гипотез. Одной из традиционных задач статистики является проверка статистических гипотез. Простейшая схема выглядит следующим образом. Выдвигается основная, или *нулевая*, гипотеза о том, что неизвестный параметр принадлежит некоторому заданному подмножеству $Z_0 \subset \Theta$, и *альтернативная* гипотеза о том, что параметр принадлежит другому подмножеству $Z_1 \subset \Theta$. Обычно используются обозначения: $H_0: \theta \in Z_0$ и $H_1: \theta \in Z_1$. Требуется на основании наблюдений $X_1, \dots, X_n$ принять (проверить) нулевую гипотезу H_0 или отвергнуть ее в пользу альтернативной гипотезы H_1 . Ниже мы более подробно рассмотрим задачи оценивания и проверки гипотез.

6.2. Оценивание параметров

Пусть $\hat{\theta} = \varphi_n(X_1, \dots, X_n)$ — некоторая оценка параметра θ .

ОПРЕДЕЛЕНИЕ. Оценка $\hat{\theta} = \varphi_n(X_1, \dots, X_n)$ называется *несмещенной*, если

$$M[\varphi_n(X_1, \dots, X_n)] = \theta \text{ для } \forall \theta \in \Theta. \quad (5.28)$$

В формуле (5.28) математическое ожидание берется по распределению, порожденному функцией $F(x; \theta)$. Чтобы подчеркнуть это, условие несмещенности часто записывают в виде $M_\theta(\hat{\theta}) = \theta$. Выборочное среднее $\bar{X}$ является примером несмещенной оценки математического ожидания случайной величины X , определяющей генеральную совокупность. Действительно, поскольку $M(X_i) = m$ для каждого i , то из свойств математического ожидания следует, что $M(\bar{X}) = m$.

ОПРЕДЕЛЕНИЕ. Оценка $\hat{\theta} = \varphi_n(X_1, \dots, X_n)$ называется *состоятельной*, если $p \lim_{n \rightarrow \infty} \varphi_n(X_1, \dots, X_n) = \theta$.

ОПРЕДЕЛЕНИЕ. Оценка $\hat{\theta} = \varphi_n(X_1, \dots, X_n)$ называется *эффективной*, если $M(\varphi_n(X_1, \dots, X_n) - \theta)^2 = \min_g M(g(X_1, \dots, X_n) - \theta)^2$, где минимум берется по всем возможным оценкам g параметра θ .

Ясно, что в классе несмещенных оценок эффективность означает минимальность дисперсии.

Содержательно несмещенность оценки означает, что при ее использовании мы не получаем систематической ошибки; состоятельность оценки гарантирует приближение оценки к истинному значению параметра при увеличении объема выборки, а эффективная оценка является наилучшей в смысле минимума среднеквадратического отклонения. Отметим, что несмещенность и эффективность — это свойства, не зависящие от объема выборки n , в то время как состоятельность является асимптотическим свойством при стремлении n к бесконечности. Смещенность или несмещенность конкретной оценки проверяется, как правило, непосредственными вычислениями.

Неравенство Рао — Крамера. Пусть $p(x; \theta)$ — плотность распределения случайного вектора $X = (X_1, \dots, X_n)$, $x \in \mathbf{R}^n$, зависящая от одномерного параметра θ . Предположим, что выполнены следующие условия:

1) множество $G = \{x \in \mathbf{R}^n: p(x; \theta) > 0\}$ не зависит от θ (условие регулярности),

2) плотность $p(x; \theta)$ дифференцируема по θ при каждом $x \in G$ и

$$I_n = \int_G \left[\frac{\partial \ln p(x; \theta)}{\partial \theta} \right]^2 p(x; \theta) dx = M \left[\frac{\partial \ln p(x; \theta)}{\partial \theta} \right]^2 < \infty. \quad (5.29)$$

Тогда для любой несмещенной оценки $\hat{\theta} = \varphi_n(X_1, \dots, X_n)$ параметра θ выполнено неравенство (Рао — Крамера):

$$M(\hat{\theta} - \theta)^2 = D(\hat{\theta}) \geq \frac{1}{I_n}. \quad (5.30)$$

Число I_n в (5.29) называется *информационным количеством* θ в X . Если компоненты $X_1, \dots, X_n$ вектора X независимы и одинаково распределены с плотностью $p(x; \theta)$, $x \in \mathbf{R}^1$, то можно показать, что $I_n = nI_1$, где I_1 — информационное количество θ в одной компоненте X_k :

$$I_1 = M \left[\frac{\partial \ln p(X_k; \theta)}{\partial \theta} \right]^2.$$

Неравенство Рао — Крамера устанавливает нижнюю границу для дисперсии оценки, поэтому если для какой-то несмещенной оценки $\hat{\theta}$ в (5.30) достигается равенство, можно утверждать, что оценка $\hat{\theta}$ эффективна (в классе несмещенных оценок). Именно таким образом можно доказать, что выборочное среднее $\bar{X}$ есть эффективная оценка среднего значения для нормальной генеральной совокупности. Неравенство Рао — Крамера обобщается на случай смещенных оценок, а также на случай многомерного параметра θ (число I_n при этом заменяется на соответствующую матрицу). Отметим, что условие регулярности является существенным — можно привести примеры, когда его отсутствие приводит к нарушению неравенства (5.30).

Здесь мы кратко опишем лишь два общих метода оценивания неизвестных параметров.

Метод максимального правдоподобия. Пусть $X = (X_1, \dots, X_n)$, $x \in \mathbf{R}^n$ — случайный вектор, имеющий плотность распределения $p(x; \theta)$, которая зависит от неизвестного параметра θ . *Функцией правдоподобия (likelihood function)* называется случайная величина $L = L(X; \theta) = p(X; \theta)$. Наряду с L рассматривают также *логарифмическую функцию правдоподобия* $l = \ln L$. *Оценкой максимального правдоподобия* называется величина $\hat{\theta} = \hat{\theta}_{ML}$, максимизирующая (при каждом фиксированном значении X) функцию правдоподобия L (или, что эквивалентно, логарифмическую функцию правдоподобия l), т.е. такая функция $\hat{\theta} = \hat{\theta}(X)$, что $L(X; \hat{\theta}(X)) = \max_{\theta} L(X; \theta)$. Индекс ML соответствует английскому термину *Maximum Likelihood estimation*. В регулярном случае необходимым

условием максимума является уравнение $\frac{\partial L(X; \theta)}{\partial \theta} = 0$, которое называется *уравнением правдоподобия*. Можно заметить, что если компоненты $X_1, \dots, X_n$ вектора X независимы и одинаково распределены с плотностью $p(x; \theta)$, $x \in \mathbf{R}^1$, то функция правдоподобия есть произведение функций правдоподобия каждой компоненты (соответственно логарифмическая функция правдоподобия распадается в сумму индивидуальных логарифмических функций).

Для широкого класса задач оценки максимального правдоподобия являются состоятельными и асимптотически эффективными. В то же время они могут быть смещенными. Например, с помощью непосредственных вычислений можно показать, что для нормальной генеральной совокупности оценки максимального правдоподобия среднего значения, дисперсии и математического ожидания дисперсии есть соответственно

$$\hat{m}_{ML} = \bar{X} \text{ и } \hat{\sigma}_{ML}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{n} \sigma^2, \quad M(\hat{\sigma}_{ML}^2) = \frac{n-1}{n} \sigma^2.$$

Недостатком метода является необходимость знать распределение вектора X .

Метод моментов. Пусть $X_1, \dots, X_n$ — случайная выборка из генеральной совокупности X , распределение которой зависит от не-

которого неизвестного параметра θ : $F(x) = F(x; \theta)$, $\theta \in \Theta \subseteq \mathbf{R}^m$. Предположим, что заданы функции $g_1(x), \dots, g_m(x)$, такие, что

а) в явном виде вычисляются математические ожидания

$$f_k(\theta) = f_k(\theta_1, \dots, \theta_m) = M g_k(X), \quad k = 1, \dots, m; \quad (5.31)$$

б) система (5.31), рассматриваемая как система уравнений относительно θ , имеет единственное решение.

Для получения оценки параметра θ в правой части системы (5.31) математические ожидания заменяют их выборочными ана-

логами $\overline{g_k(X)} = \frac{1}{n} \sum_{i=1}^n g_k(X_i)$, $k = 1, \dots, m$, т.е. строят систему уравнений (относительно θ)

$$f_k(\theta) = f_k(\theta_1, \dots, \theta_m) = \overline{g_k(X)}, \quad k = 1, \dots, m, \quad (5.32)$$

решение которой и дает оценку параметра θ методом моментов. Функции $g_1(x), \dots, g_m(x)$, участвующие в оценивании, носят название *моментных функций*. Термин “момент” связан с тем обстоятельством, что часто используют функции $g_k(x) = x^k$, иными словами, правыми частями системы (5.31) являются моменты случайной величины X . Оценки, полученные методом моментов, в широком числе случаев состоятельны, но, как правило, менее эффективны, чем оценки максимального правдоподобия.

Доверительные интервалы. Помимо получения точечной оценки неизвестного параметра часто ставится задача построения такой области, в которую параметр попадает с заданной вероятностью или с заданным уровнем доверия. Иными словами, требуется для случайной выборки $X = (X_1, \dots, X_n)$ и числа $0 < \alpha < 1$ найти такое множество $D_n = D_n(X_1, \dots, X_n) \subseteq \Theta$, что $P(\theta \in D_n(X_1, \dots, X_n)) = \alpha$. Это множество D_n называется *доверительным множеством с уровнем доверия α* (или $100\alpha\%$ — *доверительным множеством*). Для одномерного параметра θ доверительное множество обычно называют *доверительным интервалом*. Часто этот термин распространяют и на многомерный случай. Для построения доверительных множеств используется следующий общий подход. Предположим, что существует такое преобразование $h_n(x_1, \dots, x_n; \theta)$, что

распределение случайной величины $h_n(X_1, \dots, X_n; \theta)$ уже не зависит от θ и может быть эффективно найдено (например, затабулировано). Тогда для заданного $0 < \alpha < 1$ можно найти интервал I_n , такой, что $P(h_n(X_1, \dots, X_n; \theta) \in I_n) = \alpha$. Если при этом включение $h_n(X_1, \dots, X_n; \theta) \in I_n$ можно “разрешить” относительно θ , т.е. найти такое множество $D_n = D_n(X_1, \dots, X_n) \subseteq \Theta$, что $h_n(X_1, \dots, X_n; \theta) \in I_n \Leftrightarrow \theta \in D_n(X_1, \dots, X_n)$, то из самой конструкции следует, что D_n будет доверительным множеством с уровнем доверия α . Подчеркнем, что доверительное множество не единственно. Рассмотрим два примера построения доверительных интервалов.

Доверительный интервал для среднего значения нормальной генеральной совокупности. Пусть $X_1, \dots, X_n$ — случайная выборка из нормальной генеральной совокупности с параметрами (m, σ^2) .

Величина $\frac{(\bar{X} - m)\sqrt{n}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}} = \frac{(\bar{X} - m)\sqrt{n}}{s_x}$ имеет распределение

Стьюдента с $n - 1$ степенями свободы (преобразование h_n). Поэтому для заданного $0 < \alpha < 1$ в силу симметричности распределения Стьюдента можно найти (например, по таблицам) число t_α (двусторонний α -квантиль) таким образом, чтобы

$$P\left(\left|\frac{\sqrt{n}(\bar{X} - m)}{s_x}\right| \leq t_\alpha\right) = \alpha. \text{ Разрешая относительно } m, \text{ получаем для}$$

этого параметра доверительный интервал $\left(\bar{X} - \frac{s_x t_\alpha}{\sqrt{n}}, \bar{X} + \frac{s_x t_\alpha}{\sqrt{n}}\right)$.

Доверительный интервал для дисперсии нормальной генеральной совокупности. Случайная величина $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{\sigma^2} s_x^2$

имеет распределение $\chi^2(n-1)$ (преобразование h_n). Поэтому для заданного $0 < \alpha < 1$ можно найти интервал $[l_\alpha, u_\alpha]$, такой что

$P\left(l_\alpha \leq \frac{n-1}{\sigma^2} s_X^2 \leq u_\alpha\right) = \alpha$. Разрешая относительно σ^2 , получаем для

этого параметра доверительный интервал $\left(\frac{(n-1)s_X^2}{u_\alpha}, \frac{(n-1)s_X^2}{l_\alpha}\right)$.

6.3. Проверка гипотез

Проверка гипотез используется всякий раз, когда необходим обоснованный вывод о преимуществах того или иного способа инвестиций, стрельбы, различных измерений, процесса управления, об эффективности введения инновационных технологий в обучении, о значимости математической модели и т. д. В рамках параметрического подхода общая схема проверки гипотезы может быть описана так. Пусть $X_1, \dots, X_n$ — случайная выборка из некоторой генеральной совокупности с функцией распределения $F(x) = F(x; \theta)$, $\theta \in \Theta \subseteq \mathbf{R}^m$. Относительно параметра θ выдвигаются две гипотезы, а именно, $H_0: \theta \in Z_0$ и $H_1: \theta \in Z_1$, где $Z_0 \subset \Theta$, $Z_1 \subset \Theta$ — некоторые заданные множества. Гипотезу H_0 называют *основной* или *нулевой*, а гипотезу H_1 — *альтернативной*. Если множество Z состоит из одной точки ($Z = \{\theta_n\}$), то соответствующая гипотеза называется *простой*, в противном случае она называется сложной.

Простая гипотеза, в отличие от сложной, полностью определяет теоретическую функцию распределения случайной величины. Например, гипотеза “вероятность появления события в схеме Бернулли равна 1/2”, “закон распределения случайной величины нормальный с параметрами $a = 0$, $\sigma^2 = 1$ ” являются простыми, а гипотезы “вероятность появления события в схеме Бернулли заключена между 0,3 и 0,6”, “закон распределения не является нормальным” — сложными.

ОПРЕДЕЛЕНИЕ. *Статистическим тестом*, или просто *тестом*, называется любая процедура, основанная на наблюдениях $X_1, \dots, X_n$, результатом которой является одно из двух возможных решений:

- 1) не отвергать (принять) нулевую гипотезу H_0 ;

2) отвергнуть нулевую гипотезу H_0 в пользу альтернативной гипотезы H_1 .

Поскольку тест использует случайную выборку $X_1, \dots, X_n$, то, естественно, могут возникать ошибочные решения. В связи с этим возникают две ошибки теста:

ошибка первого рода: нулевая гипотеза отвергается, когда она верна;

ошибка второго рода: нулевая гипотеза принимается, когда верна альтернативная гипотеза.

Вероятности ошибок первого и второго рода можно обозначить $\alpha = P(H_1|H_0)$ и $\beta = P(H_0|H_1)$ соответственно. Величину $1 - \alpha$ называют значимостью теста, а величину $1 - \beta$ — его мощностью.

Естественно при построении теста стремиться уменьшить эти ошибки, однако нетрудно понять, что невозможно минимизировать их одновременно. Поэтому обычно поступают следующим образом: фиксируют значимость теста и стараются найти такой тест, у которого мощность максимальна (именно здесь в явном виде проявляется несимметричность гипотез, деление их на основную и альтернативную).

На практике для построения тестов часто используют следующий подход. Предположим, что можно найти такую статистику $t_n = t_n(X_1, \dots, X_n)$, что если гипотеза H_0 верна, то распределение случайной величины t_n известно (например, затабулировано). Тогда для заданного значения α ошибки первого рода можно найти такую область K_α , что $P(t_n \in K_\alpha) = 1 - \alpha$ (подчеркнем, что вероятность вычисляется в предположении, что верна нулевая гипотеза). Тогда тест определяется следующим образом:

1) на основании наблюдений $X_1, \dots, X_n$ вычисляется значение статистики t_n ;

2) для заданного уровня значимости $1 - \alpha$ находится область K_α ,

3) если $t_n \in K_\alpha$, то нулевая гипотеза не отвергается (принимается); если $t_n \notin K_\alpha$, то нулевая гипотеза отвергается в пользу альтернативной.

Статистику t_n называют *проверочной статистикой*, а область K_α — критической областью. На практике часто проверочные ста-

тистики имеют распределения: стандартное, нормальное, χ^2 , Стьюдента и Фишера. В этих случаях при использовании подобного рода тестов для каждого значения проверочной статистики, полученной в эксперименте, находится еще так называемое *P-значение*. Если статистика t_n , распределение которой при нулевой гипотезе принадлежит к одному из указанных четырех типов, приняла значение c , то соответствующим *P-значением* называется число $P(|t_n| > |c|)$ — для нормального распределения и распределения Стьюдента и число $P(t_n > c)$ для распределений χ^2 и Фишера. Таким образом, *P-значение* сразу позволяет судить о значимости нулевой гипотезы. Все компьютерные пакеты, как правило, приводят *P-значения* вычисленных статистик.

Можно установить связь между описанной выше процедурой и построением доверительных интервалов. Действительно, предположим, что для неизвестного параметра θ построен доверительный интервал D_n с уровнем доверия $1 - \alpha$ и предположим, что нулевая гипотеза является простой, т. е. $H_0: \theta = \theta_0$. Тогда нулевая гипотеза не отвергается, если $\theta_0 \in D_n$, и отвергается в противном случае.

По своему прикладному содержанию статистические гипотезы можно подразделить на несколько основных типов:

- о ранстве числовых характеристик генеральных совокупностей;
- о числовых значениях параметров;
- о законе распределения;
- об однородности выборок (т. е. принадлежности их одной и той же генеральной совокупности).

Более подробный теоретический материал на эту тему и примеры решения задач вы можете найти в специальной литературе. В пределах данного учебника многие понятия вводятся в ознакомительном порядке с целью упрощения изложения курса.

§7. Математическая модель прогнозирования экономических процессов

7.1. Экономическое прогнозирование

Экономико-математическое моделирование в настоящее время — один из основных инструментов экономического анализа. Под инструментом здесь понимается не только конкретное использование экономико-математических методов и моделей, а также соответствующих технических средств для их реализации, но и методологический подход, взгляд на экономические процессы и явления, на их внутреннюю структуру, свойства, развитие с точки зрения аппарата математического моделирования. Использование экономико-математических методов и моделей позволяет получить новые качественные выводы об экономических процессах и явлениях.

Процесс системного анализа экономической системы можно разбить на следующие этапы:

- 1) постановка задачи, определение целей и критериев оценки;
- 2) анализ исследуемой системы;
- 3) разработка концепции развития системы и подготовка возможных вариантов решений и их последствий.

Реализация второго и третьего этапов в настоящее время в связи с возросшей сложностью экономических систем объективно невозможна без использования экономико-математических методов и моделей. Их применение позволяет в значительной степени пересмотреть существующие методы учета и экономического анализа, использовать намного больше информации, производить альтернативные, многовариантные расчеты, получать более устойчивые оценки.

Рассмотрим прогнозирование экономических процессов, а именно:

- методологию экономического прогнозирования;
- основные этапы построения моделей экономического прогнозирования;

классификацию методов экономического прогнозирования;
экстраполяцию экономических процессов с использованием кривых роста;
методы выбора наилучшей кривой роста;
методы вычисления параметров кривых роста;
критерии точности и адекватности моделей экономического прогнозирования;
точные и интервальные прогнозы.

Экономическое прогнозирование (ЭП) — это процесс разработки экономических прогнозов, основанных на научных методах познания экономических явлений, и использования всей совокупности методов, средств и способов экономической прогностики. В то же время ЭП является частью **прогностики** — прикладной научной дисциплины, изучающей закономерности и способы разработки прогнозов развития объектов любой природы. В качестве *объекта* ЭП рассматривает процесс конкретного расширения воспроизводства, а в качестве *предмета* — познание возможных состояний функционирующих экономических объектов в будущем, исследование закономерностей их поведения и способов разработки экономических прогнозов.

Основное содержание ЭП — качественный и количественный анализ реальных экономических процессов, выявление объективных условий, факторов и тенденций развития. Содержание ЭП предопределяет основные принципы разработки прогнозов, к числу которых относятся:

системность;
адекватность;
альтернативность.

Системность ЭП определяет анализ явления как единого целого и как совокупности относительно самостоятельных направлений прогнозирования. Практическая реализация этого принципа предполагает создание моделей, которые соответствовали бы содержанию каждого отдельного блока и одновременно позволяли бы построить целостную картину развития объекта в будущем.

Адекватность означает максимальное приближение теоретической модели к устойчивым, существенным закономерностям.

Она предполагает учет реальных процессов, т.е. необходимость оценки сложившихся условий и возможных отклонений от господствующих тенденций, определение возможной области рассеивания, что эквивалентно оценке вероятностной реализации выявленной тенденции. На практике это означает, что построенные модели должны быть сначала проверены с точки зрения их способности имитировать уже сложившиеся тенденции.

При переходе от модельной имитации сложившихся закономерностей к предвидению процессов и тенденций развития возникает необходимость построения альтернатив, т.е. определения возможных путей развития исследуемого объекта. *Альтернативность* прогнозирования связана с возможностью развития экономического объекта по разным траекториям, при разных взаимосвязях и структурных отношениях. Главная проблема практической реализации этого принципа состоит в том, чтобы отделить те варианты развития, которые могут быть осуществлены, от вариантов, которые при сложившихся условиях не могут быть реализованы. Среди всех вариантов наибольшую вероятность реализации имеет обычно экстраполярный, поскольку он исходит из сохранения сложившихся устойчивых условий и тенденций. Построение экстраполяционных вариантов позволяет установить “точку отсчета” для других вариантов.

Названные принципы лежат в основе конкретных методов ЭП.

7.2. Методы прогнозирования

ОПРЕДЕЛЕНИЕ. *Методом прогнозирования* называется способ исследования объекта прогнозирования, направленный на разработку прогнозов.

Модель прогнозирования представляет собой модель исследуемого объекта, записанную в математической форме. Она должна позволить получить информацию о возможных состояниях объекта в будущем и (или) путях и сроках их достижения.

При построении прогнозной модели может быть использован один или несколько методов. Например, при построении линейной модели (уравнение прямой) могут быть использованы *методы средних, двух точек, наименьших квадратов, экспоненциального сглаживания, гармонических весов, эволюции* и др.

В большинстве методов, используемых в настоящее время в практической деятельности, заложены следующие основные предположения:

основные наблюдавшиеся в прошлом факторы, тенденции и зависимости сохраняются, либо можно будет предвидеть и определить направление их изменения в прогнозируемом периоде;

развитие экономических процессов может быть представлено в виде плавной траектории;

экономические процессы имеют вероятностный характер, а развитие исследуемого объекта определяется суммарным влиянием закономерности и случайности.

Последнее предположение обуславливает целесообразность использования *статистических методов прогнозирования*, которые при необходимости могут дополняться и другими методами (*аналогий, экспертных оценок* и т.д.). Так, например, в банковской деятельности ежедневные денежные потоки формируются под влиянием определенных закономерностей (запланированные платежи), а также необязательных, порой непредвиденных поступлений или платежей.

Процесс прогнозирования, опирающийся на статистические методы, распадается на два основных этапа. *Первый этап* (индуктивный) — обобщение данных, наблюдаемых за достаточно продолжительный период, и представление статистических закономерностей в виде модели, которая выражается либо аналитической функцией тенденции развития, либо в виде зависимости от нескольких факторов-аргументов. *Второй этап* — собственно прогноз — дедуктивный. На основе выявленных закономерностей определяют ожидаемые значения прогнозируемого показателя, которые должны быть осмысленны с содержательной точки зрения. Указанные этапы конкретизируются в определенной последовательности шагов:

1. Обоснование теоретических предположений, являющихся исходными для составления прогноза, и выбор системы показателей, адекватно отражающих развитие объекта. Каждый показатель должен иметь экономическое содержание и быть количественно измеримым.

2. Разработка системы моделей, отражающих развитие отдельных сторон и показателей исследуемого объекта, а также взаимосвязей между показателями.

3. Сбор и предварительный анализ данных.

4. Количественная оценка взаимосвязей показателей и параметров моделей. В зависимости от выбираемого критерия и численного метода оценки получаются разные результаты.

5. Определение прогнозных значений на основе построенной модели. Экстраполяция выявленных тенденций (продление на будущее) позволяет получить точечный прогноз. Однако вероятность точного попадания в эту точку практически равна нулю. Отсюда следует необходимость вычисления перспективных оценок в виде “вилки” через доверительные интервалы.

Методы ЭП, в частности статистического прогнозирования, “осмысливают” лишь часть формализованной информации о прошлом развитии, в то время как специалист-аналитик владеет еще и значительным объемом слабо формализуемой, но очень важной информации о текущем и будущем развитии исследуемого объекта, а также его внешней среды. Поэтому полученную прогностическую информацию следует рассматривать как дополнительную, которая призвана помочь аналитику прояснить ситуацию и принять правильное решение.

Таким образом, ЭП требует глубокого анализа и учета объективных экономических законов общественного развития, конкретного экономического анализа исследуемого объекта. Кроме того, необходимо иметь достоверную количественную информацию о прошлом развитии, технические и программные средства реализации методов прогнозирования, а также практические навыки по их использованию.

7.3. Статистический анализ временных рядов

Основной формой представления информации о динамике экономических показателей являются *временные ряды* наблюдений. Временной ряд, состоящий из N уровней $Y(1)$, $Y(2)$, ..., $Y(N)$, может быть записан в компактной форме: $Y(t)$, $t = 1, 2, \dots, N$,

где t — порядковый номер наблюдения. Формально задача прогнозирования сводится к получению оценок значений ряда на некотором периоде будущего, т.е. к получению $Y(t)$ в момент времени $t = N + 1, N + 2, \dots$. При использовании методов экстраполяции исходят из предположения о сохранении закономерностей прошлого развития в течение периода прогнозирования.

Простейший способ прогнозирования основан на использовании фактически достигнутого уровня $Y(N)$ и *среднего абсолютного прироста* (САП). В этом случае прогноз на k шагов вперед на момент $t = N + k$ получается по формулам:

$$Y(n + k) = Y(N) + k \cdot \text{САП}, \quad (5.33)$$

где $\text{САП} = [Y(N) - Y(1)]/(N - 1)$.

Этот способ очень привлекателен из-за простоты и легкости реализации. Однако он обладает существенными недостатками: например, невозможно сформировать интервал, внутрь которого попадает прогнозируемая величина. Поэтому данный подход используется обычно в условиях очень малого объема наблюдений, когда статистические методы применить невозможно.

Статистические методы исследования основаны на предположении о возможности представления уровней ряда в виде суммы компонент, отражающих закономерность и случайность развития:

$$Y(t) = f(t) + E(t), \quad (5.34)$$

где $f(t)$ — *тренд* (долговременная тенденция) развития; $E(t)$ — остаточная компонента.

Основная цель статистического анализа временных рядов — изучение соотношения закономерности и случайности в формировании значений уровней ряда с оценкой количественной меры их влияния. Закономерности, объясняющие динамику показателя в прошлом, могут быть использованы для прогнозирования его значений в будущем, а учет случайности позволяет определить вероятность отклонения от закономерного развития и их возможную величину.

7.4. Кривые роста

Для отражения тенденций изменения исследуемого показателя, найденных с помощью кривых роста, используются разнообраз-

разные математические функции, в которых задействован только один фактор — время t . Из большого числа кривых роста воспользуемся простейшей линейной моделью вида

$$Y_p(t) = a_0 + a_1 t \quad (t = 1, 2, \dots, N). \quad (5.35)$$

Параметры кривых роста оцениваются по методу наименьших квадратов, т.е. подбираются таким образом, чтобы кривая роста располагалась на минимальном удалении от точек исходных данных. Математически критерий оценки параметров линейной модели записывается так:

$$\begin{aligned} \sum \{Y(t) - (a_0 + a_1 t)\}^2 &\rightarrow \min, \\ \begin{cases} a_0 N + a_1 \sum t = \sum Y(t), \\ a_0 \sum t + a_1 \sum t^2 = \sum (Y(t)t). \end{cases} \end{aligned}$$

Суммирование производится по всем уровням ряда. Для упрощения расчетов используется центрирование переменных относительно их среднего значения. Решением системы являются значения, вычисляемые по формулам

$$\begin{aligned} a_1 &= \sum [(t - t_{cp}) \times (Y(t) - Y_{cp})] / \sum (t - t_{cp})^2, \\ a_0 &= Y_{cp} - a_1 t_{cp}, \end{aligned} \quad (5.36)$$

где t_{cp} — среднее значение фактора времени; Y_{cp} — среднее значение исследуемого показателя.

▷ П Р И М Е Р 5.46. Построить линейную модель для данных о курсе акций за девять недель, приведенных в табл. 5.2. Проверить модель на адекватность и составить точечный прогноз на два шага.

Р е ш е н и е. Рассмотрим табл. 5.2. В ней приведены промежуточные вычисления и результаты использования линейной модели. В нижней строке записаны суммы значений в графах.

По данным первых двух колонок, легко определить: $Y_{cp} = 504/9 = 56$, $t_{cp} = 45/9 = 5$. Подставляя рассчитанные значения в формулу (5.36), получаем

$$a_1 = 428/60 = 7,13,$$

Таблица 5.2

t	$Y(t)$	$t - t_{cp}$	$(t - t_{cp})^2$	$Y(t) - Y_{cp}$	$(t - t_{cp})(Y(t) - Y_{cp})$	$Y_p(t)$	$E(t)$
1	25	-4	16	-31	124	27,5	-2,5
2	34	-3	9	-22	66	34,6	-0,6
3	42	-2	4	-14	28	41,7	0,3
4	51	-1	1	-5	5	48,9	2,1
5	55	0	0	-1	0	56,0	-1,0
6	67	1	1	11	11	63,1	3,9
7	73	2	4	17	34	70,3	2,7
8	76	3	9	20	60	77,4	-1,4
9	81	4	16	25	100	84,5	-3,5
45	504	0	60	0	428	504	0

$$a_0 = 56 - 7,13 \cdot 5 = 20,35.$$

Таким образом, линейная модель имеет вид:

$$Y_p(t) = 20,35 + 7,13t \quad (t = 1, 2, \dots, 9). \quad (5.37)$$

Последовательно подставляя в модель вместо фактора t его значения от 1 до N , получаем расчетные значения уровней $Y_p(t)$:

$$Y_p(t) = 20,35 + 7,13 \cdot 1 = 27,5 \quad (t = 1),$$

$$Y_p(t) = 20,35 + 7,13 \cdot 2 = 34,6 \quad (t = 2) \text{ и т. д.}$$

Вычислим отклонения расчетных значений от значений, полученных при наблюдениях, по формуле $E(t) = Y(t) - Y_p(t)$ ($t = 1, 2, \dots, 9$):

$$E(1) = Y(1) - Y_p(1) = 25,0 - 27,5 = -2,5,$$

$$E(2) = Y(2) - Y_p(2) = 34,0 - 34,6 = -0,6 \text{ и т. д.}$$

Отражает ли эта модель закономерность изменения исследуемого показателя — иными словами, можно ли полученные значения $Y_p(t)$ рассматривать как тенденцию? Для ответа на этот вопрос оценим качество модели, или ее адекватность исследуемому процессу. Адекватность характеризуется выполнением определенных статистических свойств и точностью, т.е. степенью близости к фактическим данным. Модель считается хорошей со статистической точки зрения, если она адекватна.

Модель является адекватной, если математическое ожидание значений остаточного ряда близко или равно нулю и если значе-

ния остаточного ряда случайны, независимы и подчинены нормальному закону распределения.

Исследуем полученную модель на адекватность.

1. Проверка равенства математического ожидания уровней остаточного ряда нулю осуществляется с использованием *t-критерия Стьюдента*:

$$t = \frac{|\bar{E}|}{S_E} \sqrt{N},$$

где $\bar{E}$ — среднее значение уровней остаточного ряда; S_E — среднее квадратическое отклонение уровней остаточного ряда. Значение $\bar{E}$ берется по модулю, без учета знака.

Среднее квадратическое отклонение

$$S_E = \sqrt{\frac{\sum_{t=1}^N E(t)^2 - \left[\sum_{t=1}^N E(t) \right]^2}{N(N-1)}}.$$

Если $\sum_{t=1}^N E(t) = 0$, то

$$S_E = \sqrt{\frac{\sum_{t=1}^N E(t)^2}{N(N-1)}}.$$

Гипотеза отклоняется, если $t > t_{\text{табл}}$ с заданным уровнем доверительной вероятности p .

При $N = 1$ и $p = 70\%$ получаем $t_{\text{табл}} = 1,10$.

В рассматриваемом примере $\sum_{t=1}^N E(t) = 0$ (см.табл. 5.2), поэто-

му гипотеза о равенстве математического ожидания значений остаточного ряда нулю выполняется.

2. Проверка случайностей уровней остаточного ряда проводится на основе *критерия поворотных точек*. В соответствии с ним каждый уровень ряда сравнивается с двумя рядом стоящими уровнями.

Если он больше или меньше их, то эта точка считается поворотной. Далее подсчитывается сумма поворотных точек s . В случайном ряду чисел должно выполняться строгое неравенство

$$s > [2(N-2)/3 - 2\sqrt{(16N-29)/90}]. \quad (5.38)$$

Квадратные скобки здесь означают, что берется целая часть числа (в отличие от процедуры округления). При $N = 9$ в правой части неравенства имеем:

$$[2 \times (9-2)/3 - 2\sqrt{(16 \cdot 9 - 29)/90}] = [2,4] = 2.$$

В табл. 5.3 в четвертой графе для первого и последнего наблюдения проставим прочерк, ноль — если точка неповоротная, и единицу, если она поворотная. В нашем примере количество поворотных точек равно трем ($s = 3$), неравенство (5.38) выполняется, следовательно, свойство случайности выполняется.

3. При проверке независимости (отсутствия автокорреляции) определяется отсутствие в остаточном ряду систематической составляющей. Это проверяется с помощью *d-критерия Дарбина-Уотсона*, в соответствии с которым вычисляется коэффициент d :

$$d = \sum_{t=2}^N [E(t) - E(t-1)]^2 / \sum_{t=1}^N E(t)^2. \quad (5.39)$$

Вычисленное значение этого критерия сравнивается с двумя табличными уровнями (нижним d_1 и верхним d_2). Если d (или d') находится в интервале от нуля до d_1 , то уровни остаточного ряда сильно автокоррелированы, а модель неадекватна. Если его значение попадает в интервал от d_2 до 2, то уровни ряда независимы. Если d превышает 2, то это свидетельствует об отрицательной корреляции и перед вводом в таблицу этот результат надо преобразовать:

$$d' = 4 - d.$$

В нашем примере $d = 63,94/48,82 = 1,31$.

Для линейной модели при девяти наблюдениях можно взять в качестве критических табличных уровней величины $d_1 = 1,08$, $d_2 = 1,36$. Следовательно, расчетное значение d попало в зону между d_1 и d_2 , поэтому однозначного вывода сделать нельзя. В этом

случае следует применять другие критерии, например первый коэффициент автокорреляции $r(1)$, вычисляемый по формуле

$$r(1) = \sum_{t=2}^N E(t)E(t-1) / \sum_{t=2}^N E(t)^2. \quad (5.40)$$

Если $|r(1)| > r_{\text{табл}}$ (при $N < 15$, $r_{\text{табл}} = 0,36$), то присутствие в остаточном ряду существенной автокорреляции подтверждается. Для рассматриваемого примера $r(1) = 7,60/48,82 = 0,16$. Следовательно, по этому критерию выполняется свойство независимости уровней остаточной компоненты.

4. Соответствие остаточного ряда нормальному закону распределения определим при помощи RS -критерия:

$$RS = [E_{\max} - E_{\min}] / S_E, \quad (5.41)$$

где $E_{\max}$ и $E_{\min}$ — максимальный и минимальный уровни остаточного ряда.

Если значение этого критерия попадает между табулированными границами с заданным уровнем вероятности, то гипотеза о нормальном распределении остаточного ряда принимается. Для $N = 9$ и 5%-ного уровня значимости этот интервал равен (2,7; 3,7).

В нашем примере: $E_{\max} = 3,9$ и $E_{\min} = -3,5$, а размах 7,4.

Среднее квадратическое отклонение

$$S_E = \sqrt{\sum_{t=1}^N E(t)^2 / (N-1)} = \sqrt{48,82/(9-1)} = \sqrt{6,10} = 2,47,$$

откуда

$$RS = [3,9 - (-3,5)] / 2,47 = 3,0.$$

Расчетное значение попадает в интервал. Следовательно, свойство нормальности распределения выполняется, что позволяет строить доверительный интервал прогноза.

Результаты исследования (оценка) адекватности полученной модели отражены в табл. 5.3.

Теперь исследуем модель на точность.

Для характеристики точности воспользуемся средней относительной ошибкой:

Таблица 5.3

t	$Y(t)$	$E(t)$	Точки поворота	$E(t)^2$	$E(t) - E(t-1)$	$[E(t) - E(t-1)]^2$	$E(t)E(t-1)$	$ E(t) /Y(t) \cdot 100$
1	25	-2,5	—	6,25	—	—	—	10,0
2	34	-0,6	0	0,36	1,9	3,61	1,5	1,76
3	42	0,3	0	0,09	0,9	0,81	-0,18	0,71
4	51	2,1	1	4,41	1,8	3,24	0,63	4,12
5	55	-1,0	1	1,0	-3,1	9,61	-2,1	1,82
6	67	3,9	1	15,21	4,9	24,01	-3,9	2,82
7	73	2,7	0	7,29	-1,2	1,44	10,53	3,70
8	76	-1,4	0	1,96	-4,1	16,81	-3,78	1,84
9	81	-3,5	—	12,25	-2,1	4,41	4,9	4,32
Σ	504	0	3	48,82	4—	63,94	7,6	34,09

$$\bar{E}_{\text{отн}} = \frac{1}{N} \sum_{t=1}^N [|E(t)|/Y(t)] \cdot 100\% = \frac{1}{9} \cdot 34,09\% = 3,79\%. \quad (5.42)$$

Величина менее 5% свидетельствует о хорошем уровне точности модели (ошибка до 15% считается приемлемой).

Итак, мы убедились, что модель адекватна. Теперь составим точечный прогноз.

Точечный прогноз на k шагов вперед получается путем подстановки в модель параметра $t = N + 1, \dots, N + k$. При прогнозировании на два шага имеем:

$$\begin{aligned} Y_p(10) &= 20,35 + 7,13 \cdot 10 = 91,65 & (k = 1, t = 10), \\ Y_p(11) &= 20,35 + 7,13 \cdot 11 = 98,78 & (k = 2, t = 11). \end{aligned}$$

Доверительный интервал прогноза будет иметь следующие границы:

$$\text{Верхняя граница прогноза} = Y_p(N + k) + U(k),$$

$$\text{Нижняя граница прогноза} = Y_p(N + k) - U(k).$$

Величина $U(k)$ для линейной модели имеет вид:

$$U(k) = S_T k_p \sqrt{1 + 1/N + (N + k - t_{\text{cp}})^2 / \sum_{t=1}^N (t - t_{\text{cp}})^2}, \quad (5.43)$$

где S_T — среднее квадратическое отклонение от линии тренда.

Если в качестве тренда используется уравнение прямой, то

$$S_T = \sqrt{\frac{\sum_{t=1}^N [Y(t) - Y_p(t)]^2}{N-2}}, \text{ или } S_T = \sqrt{\frac{\sum_{t=1}^N E(t)^2}{N-2}}.$$

В нашем примере

$$S_T = \sqrt{48,82/(9-2)} = \sqrt{6,974} = 2,64.$$

Коэффициент k_p является табличным значением статистики Стьюдента. Если исследователь задает уровень вероятности попадания прогнозируемой величины внутрь доверительного интервала, равный 70%, то $k_p = 1,1$, и, следовательно, по формуле (5.43)

$$U(1) = 2,64 \cdot 1,1 \cdot \sqrt{1 + 1/9 + (9+1-5)^2/60} = 3,60,$$

$$U(2) = 2,64 \cdot 1,1 \cdot \sqrt{1 + 1/9 + (9+2-5)^2/60} = 3,80.$$

В табл. 5.4 сведены результаты расчетов прогнозных оценок по формуле (5.37).

Таблица 5.4

Время t	Шаг k	Прогноз $Y_p(t)$	Нижняя граница	Верхняя граница
10	1	91,65	88,05	95,25
11	2	98,78	94,48	102,58

Если построенная модель адекватна, то с выбранной пользователем вероятностью можно утверждать, что при сохранении сложившихся закономерностей развития прогнозируемая величина попадает в интервал, образованный нижней и верхней границами. В нашем случае модель полностью адекватна исследуемому процессу и, следовательно, такое утверждение правомерно. ►

Контрольные вопросы

1. Какие события называются несовместными, равновероятными?
2. Какие основные формулы комбинаторики вам известны?
3. В чем состоит свойство устойчивости относительной частоты?

4. Зачем вводят геометрические вероятности?
5. Чему равна сумма вероятностей событий, образующих полную группу?
6. Что такое противоположные события и условная вероятность?
7. В чем суть теоремы об умножении вероятностей для независимых событий?
8. Каковы характеристики нормального распределения?
9. Что такое математическое ожидание?
10. Что такое дисперсия дискретной случайной величины?
11. Как найти функцию распределения по известной плотности распределения?
12. В чем состоят основные задачи математической статистики?
13. Как измерить риск портфеля, состоящего из двух акций?

Упражнения

1. В портфеле инвестора, склонного к риску, 50 разных облигаций, среди которых 5 облигаций с высокой степенью неисполнения обязательств по ним. Инвестор решает наугад исключить из портфеля одну облигацию. Найти вероятность того, что он исключил из портфеля только “бросовую” облигацию.

2. Брошена игральная кость. Найти вероятность того, что выпадет нечетное число очков.

3. Участники жеребьевки тянут из ящика жетоны с номерами от 1 до 100. Найти вероятность того, что номер первого наудачу извлеченного жетона не содержит цифры 2.

4. В мешочке имеется 5 одинаковых кубиков. На всех гранях каждого кубика написана одна из следующих букв: о, п, р, с, т. Найти вероятность того, что на вынутых по одному и расположенных “в одну линию” кубиках можно будет прочесть слово “спорт”.

5. На каждой из шести одинаковых карточек напечатана одна из следующих букв: н, к, а, р, т, о. Карточки тщательно перемешаны. Найти вероятность того, что на четырех, вынутых по одной и расположенных “в одну линию” карточках можно будет прочесть слово “крот”.

6. Куб, все грани которого окрашены, распилен на тысячу кубиков одинакового размера, которые затем тщательно перемешаны. Найти вероятность того, что наудачу извлеченный кубик будет иметь окрашенных граней: а) одну; б) две; в) три.

7. Из тщательно перемешанного полного набора 28 костей домино наудачу извлечена кость. Найти вероятность того, что вторую наудачу извлеченную кость можно приставить к первой, если первая кость: а) оказалась дублем; б) не оказалась дублем.

8. Восемь различных карандашей раскладывают случайным образом в коробку. Найти вероятность того, что два определенных карандаша окажутся лежащими рядом.

9. Портфель частного инвестора состоит из десяти различных акций, причем пять из них стоят по 4 руб. каждая, три — по 1 руб. и две — по 3 рубля. Найти вероятность того, что выбранные наудачу две акции стоят 5 рублей.

10. Вероятность того, что торговец акциями (day-Trader) при однократной продаже акций компании B закроет свою позицию (купит новые акции) с прибылью, равна $p = 0,9$. Day-Trader произвел три сделки. Найти вероятность того, что все три сделки были успешны.

11. Брошены монета и игральная кость. Найти вероятность совмещения событий «появился “герб”», “появилось 6 очков”.

12. В двух ящиках находятся детали: в первом — 10 (из них 3 стандартных), во втором — 15 (из них 6 стандартных). Из каждого ящика наудачу вынимают по одной детали. Найти вероятность того, что обе детали окажутся стандартными.

13. В студии телевидения 3 телевизионных камеры. Для каждой камеры вероятность того, что она включена в данный момент, равна $p = 0,6$. Найти вероятность того, что в данный момент включена хотя бы одна камера.

14. Чему равна вероятность того, что при бросании трех игральных костей 6 очков появится хотя бы на одной из костей?

15. Предприятие изготавливает 95% изделий стандартных, причем из них 86% — первого сорта. Найти вероятность того, что взятое наудачу изделие, изготовленное на этом предприятии, окажется первого сорта.

16. Монету подбрасывают до тех пор, пока 2 раза подряд она не выпадет одной и той же стороной. Найти вероятность следующих событий: а) опыт окончится до шестого бросания; б) потребуется четное число бросаний.

17. Из цифр 1, 2, 3, 4, 5 сначала выбирается одна, а затем из оставшихся четырех — вторая цифра. Предполагается, что все возможные исходы равновероятны. Найти вероятность того, что будет выбрана нечетная цифра: а) в первый раз; б) во второй раз; в) в оба раза.

18. Вероятность того, что торговец акциями (day-Trader) при однократной покупке акций компании A закроет свою позицию с прибылью, равна 0,6. Сколько сделок должен сделать торговец акциями, чтобы с вероятностью не менее 0,8 получить прибыль хотя бы от одной сделки?

19. Найти математическое ожидание дискретной случайной величины, зная закон ее распределения:

X	6	3	1
p	0,2	0,3	0,5

20. Производится 4 сделки с вероятностями получения прибыли: $p_1 = 0,6$, $p_2 = 0,4$, $p_3 = 0,5$, $p_4 = 0,7$. Найти математическое ожидание общего числа успешных сделок.

21. Дискретные независимые случайные величины заданы законами распределения:

X	1	2
p	0,2	0,8

Y	0,5	1
p	0,3	0,7

Найти математическое ожидание произведения XY двумя способами: а) составив закон распределения XY ; б) пользуясь свойством 3 (см. подпараграф 4.1 главы 5).

22. Дискретные случайные величины X и Y заданы законами распределения, указанными в упражнении 21. Найти математическое ожидание суммы $X + Y$ двумя способами: а) составив закон распределения $X + Y$; б) пользуясь свойством 4 (см. подпараграф 5.2 главы 5).

23. Вероятность неудачи при сдаче студентом зачета по теории вероятностей равна 0,2. Найти математическое ожидание числа студентов, не сдавших зачет, если в группе 10 человек.

24. Найти математическое ожидание произведения числа очков, которые могут выпасть при одном бросании двух игральных костей.

25. Найти математическое ожидание числа лотерейных билетов, на которые выпадут выигрыши, если приобретено 20 билетов, причем вероятность выигрыша по одному билету равна 0,3.

26. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0 & \text{при } x \leq -1, \\ x/3 + 1/3 & \text{при } -1 < x \leq 2, \\ 1 & \text{при } x > 2. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, заключенное в интервале $(0, 1)$.

27. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0 & \text{при } x \leq 2, \\ x/2 - 1 & \text{при } 2 < x \leq 4, \\ 1 & \text{при } x > 4. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, заключенное в интервале $(2, 3)$.

28. Дискретная случайная величина X задана законом распределения

X	2	6	10
p	0,5	0,4	0,1

Построить график функции распределения этой величины.

29. Математическое ожидание нормально распределенной случайной величины X равно 3 и среднее квадратическое отклонение $\sigma = 2$. Написать плотность распределения вероятности X .

30. Написать плотность распределения вероятности нормально распределенной случайной величины X , зная, что $M(X) = 3$, $D(X) = 16$.

31. Нормально распределенная случайная величина X задана плотностью $f(x) = \frac{1}{5\sqrt{2\pi}} e^{-(x-1)^2/50}$. Найти математическое ожидание и дисперсию X .

32. Дана функция распределения нормированного нормального закона $F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$. Найти плотность распределения $f(x)$.

33. Предприниматель рассматривает возможность покупки акций трех предприятий, по каждой из которых известна доходность (отношение полученного дохода за период времени к цене акции). Акции какого предприятия следует считать более доходными, если руководствоваться средним значением (математическим ожиданием) доходности?

Предприятие 1		Предприятие 2		Предприятие 3	
Доходность (%), X	Вероятность, P_X	Доходность (%), Y	Вероятность, P_Y	Доходность (%), Z	Вероятность, P_Z
4	0,2	2	0,1	1	0,1
9	0,4	12	0,4	15	0,25
8	0,3	6	0,2	3	0,4
11	0,1	14	0,3	20	0,25

Акции какого предприятия являются менее рискованными?

34. Стандартное отклонение доходности акций компании A за период времени t составляет 40%, акций компании B за тот же период — 30%. Коэффициент корреляции $\rho = 0,8$. Определить риск портфеля за тот же период времени, если инвестор купил акций компании A на 10 000 руб., B — на 40 000 руб.

Ответы к упражнениям

1. 0,1. 2. 0,5. 3. 0,81. 4. $\frac{1}{120}$. 5. $\frac{1}{360}$. 6. а) 0,384; б) 0,096; в) 0,008.
7. а) $\frac{2}{9}$; б) $\frac{4}{9}$. 8. $\frac{1}{4}$. 9. $\frac{1}{3}$. 10. 0,729. 11. $\frac{1}{12}$. 12. 0,12. 13. 0,936. 14. $\frac{91}{216}$.
15. 0,817. 16. а) $\frac{15}{16}$; б) $\frac{2}{3}$. 17. а) $\frac{3}{5}$; б) $\frac{3}{5}$; в) $\frac{3}{10}$. 18. $n \geq 2$. 19. 2,6.

20. 2,2 сделки. **21.** 1,53. **22.** 2,65. **23.** 2 студента. **24.** 12,25 очка.

25. 6 билетов. **26.** $\frac{1}{3}$. **27.** $\frac{1}{2}$. **29.** $f(x) = \frac{1}{2\sqrt{2\pi}} e^{-(x-3)^2/8}$.

30. $f(x) = \frac{1}{4\sqrt{2\pi}} e^{-(x-3)^2/32}$. **31.** $M(X) = 1$, $D(X) = 25$.

32. $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. **34.** 38%.

Глава 6

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ

§1. Оптимизационные задачи на графах

Существует ряд задач, которые удобно представлять в виде графических структур. Например, графически можно формализовать процесс принятия решения, функционирование производственной системы, транспортировку продукции, передачу информации и т. д. Анализ графических структур позволяет находить оптимальные решения этих задач. Для построения и исследования графических структур разработана система методов, изучаемая в *теории графов*. В данном параграфе приведены основные понятия теории графов и ее применение для выбора оптимального решения по методу ветвей и границ.

1.1. Основные понятия теории графов

Пусть задано некоторое непустое множество X и множество U , состоящее из пар элементов множества X . Пары во множестве U и элементы в парах могут повторяться. Множества X и U задают граф $G(X, Y)$.

Элементы множества X называют *вершинами* графа, элементы множества U — *ребрами* графа.

Если пары во множестве U повторяются, то граф G называют *псевдографом*, или *графом с кратными ребрами*.

Если элементы в парах множества U не упорядочены, то граф G называется *неориентированным*. Если они упорядочены, то граф G является *ориентированным графом*, или *орграфом*, а элементы множества U называют *дугами*.

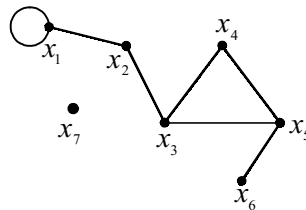


Рис. 6.1

Граф задается в виде точек и линий, их соединяющих. На рис. 6.1 приведен пример неориентированного графа.

Для этого графа множества вершин X и ребер U имеют следующий вид:

$$X = \{x_1; x_2; x_3; x_4; x_5; x_6\},$$

$$U = \{(x_1, x_1); (x_1, x_2); (x_2, x_3); (x_3, x_4); (x_3, x_5); (x_4, x_5); (x_5, x_6)\}.$$

Введем ряд основных понятий для *неориентированного* графа. Ребро, начало и конец которого совпадают, называется *петлей*. В данном случае — петля (x_1, x_1) .

Вершины называются *смежными*, или *соседними*, если существует ребро, их соединяющее.

Если вершина является началом или концом ребра, то вершина и ребро называются *инцидентными*.

Степенью вершины называется число инцидентных ей ребер, степень вершины x обозначается $d(x)$. Вершина, степень которой равна нулю, называется *изолированной*, а вершина, степень которой равна единице, — *висячей*, или *тупиковой*.

Для графа, представленного на рис. 6.1, изолированной является вершина x_7 ; $d(x_7) = 0$; тупиковой — вершина x_6 ; $d(x_6) = 1$; $d(x_2) = 2$; $d(x_3) = 3$; $d(x_4) = 2$; $d(x_5) = 3$; $d(x_1) = 3$ (ребро (x_1, x_1) учитывается дважды).

Маршрутом в графе называется последовательность вершин и ребер, в которой конец предыдущего ребра совпадает с началом следующего (это не относится к первому и последнему ребру). Число ребер в маршруте определяет его *длину*. Пример маршрута для графа, представленного на рис. 6.1: $(x_1; x_2; x_3; x_5; x_4; x_3; x_2)$. Маршрут содержит шесть ребер, следовательно, его длина равна 6.

Цепью называется маршрут, в котором все ребра попарно различны. Пример цепи — $(x_4; x_3; x_2; x_1; x_1)$. Длина цепи 4.

Простой называется цепь, в которой все вершины попарно различны. Простая цепь — $(x_1; x_2; x_3; x_4; x_5)$.

Циклом (*простым циклом*) называется цепь (простая цепь), начало и конец которой совпадают. Для графа, представленного на рис. 6.1, простым циклом является последовательность $(x_3; x_4; x_5; x_3)$.

Подграфом графа G называется граф G_1 с множеством вершин X_1 и множеством ребер U_1 , такой, что $X_1 \subset X$, $U_1 \subset U$. Подграф называется *собственным*, если он отличен от самого графа.

Граф называется *связным*, если для любых двух его вершин существует цепь, соединяющая эти вершины.

Расстоянием между вершинами связного графа называется длина самой короткой цепи, соединяющей вершины.

Диаметром графа называется максимальное расстояние между его вершинами.

Компонентой связности графа называется связный подграф, не являющийся собственным подграфом никакого другого связного подграфа данного графа. Граф, представленный на рис. 6.1, имеет две компоненты связности.

Вершина графа, удаление которой повышает число компонент связности, называется *точкой сочленения*. Под удалением вершины понимается удаление самой вершины и всех инцидентных ей ребер. Для графа, представленного на рис. 6.1, точкой сочленения является, например, вершина x_3 , удаление которой приводит к появлению третьей компоненты связности.

Деревом называется связный граф без циклов (рис. 6.2).

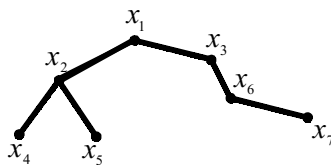


Рис. 6.2

Лесом называется граф без циклов, т.е. совокупность деревьев (рис. 6.3).

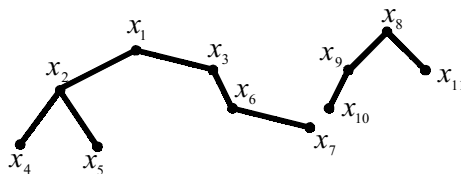


Рис. 6.3

Граф называется *полным*, если любые две его вершины соединены ребром. Полный граф с n вершинами обозначается K_n . На рис. 6.4, а и б изображены графы K_4 и K_5 соответственно.

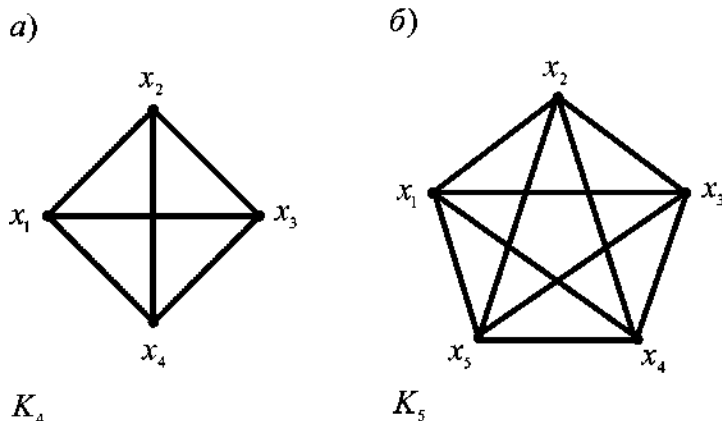


Рис. 6.4

Граф называется *регулярным степени d* , если все его вершины имеют степень d . Так, K_4 — это регулярный граф степени 3; K_5 — это регулярный граф степени 4. Регулярный граф, все вершины которого имеют степень 1, называется *паросочетанием*.

Граф называется *двудольным*, если множество его вершин X может быть разделено на два непересекающихся подмножества таким образом, что каждое ребро графа соединяет вершины из двух разных подмножеств.

Гамильтоновой цепью называется простая цепь, содержащая все вершины графа. Пример гамильтоновой цепи для графа K_5 : $(x_1; x_2; x_3; x_4; x_5)$ (см. рис. 6.4, б). *Гамильтоновым циклом* называется простой цикл, содержащий все вершины графа. Для K_4 простой цикл — $(x_1; x_2; x_3; x_4; x_1)$ (см. рис. 6.4, а).

В *ориентированном* графе каждая дуга имеет направление, показанное стрелкой (рис. 6.5).

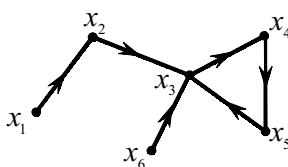


Рис. 6.5

Маршрут в ориентированном графе часто называют *контуром*, а цепь — *путем*. В графе, изображенном на рис. 6.5, путем является, например, последовательность $(x_2; x_3; x_4; x_5)$. Последовательность $(x_2; x_3; x_5)$ путем не является, так как не существует дуги, соединяющей x_3 и x_5 .

В основном все определения, данные для неориентированных графов, применимы для орграфов.

Вместо степени вершины вводится понятие полустепеней исхода и захода. Если вершина является началом дуги, то дуга называется *исходящей* из вершины, если концом — *заходящей*.

Полустепенью исхода вершины x (обозначается $d^-(x)$) называется число дуг, исходящих из этой вершины; *полустепенью захода* $d^+(x)$ — число дуг, заходящих в вершину.

Рассмотрим граф, изображенный на рис. 6.5.

$$d^-(x_1) = 1, \quad d^+(x_1) = 0;$$

$$d^-(x_2) = 1, \quad d^+(x_2) = 1;$$

$$d^-(x_3) = 1, \quad d^+(x_3) = 3; \text{ и т. д.}$$

Как уже отмечалось, с помощью графов можно производить расчет различных структур. Если структура достаточно сложна, то для решения задачи необходимо использовать компьютер. При автоматизированных работах граф неудобно задавать графически. Проще представлять его в виде двумерного массива, или матрицы.

Известно несколько **типов матриц**, позволяющих задавать граф в удобном виде.

1. *Матрицей смежности* графа, содержащего n вершин, называется матрица $A_{n \times n}$ с n строками и n столбцами, каждый элемент которой a_{ij} определяется по следующей формуле:

$$a_{ij} = \begin{cases} 1, & \text{если вершины } i \text{ и } j \text{ соединены ребром или дугой;} \\ 0 & \text{— в противном случае.} \end{cases}$$

Для графа с кратными ребрами (дугами) вместо единицы записывается число ребер (дуг) между вершинами i и j .

Для неориентированного графа, представленного на рис. 6.1, матрица смежности имеет размерность 7×7 и записывается в виде

$$A = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{matrix}.$$

Слева и сверху матрицы выписаны номера вершин.

Для ориентированного графа, изображенного на рис. 6.5, матрица смежности имеет размерность 6×6 и записывается в виде

$$A = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{matrix}.$$

Матрицу смежности чаще применяют для задания неориентированного графа. Для задания ориентированного графа лучше использовать матрицу инцидентности.

2. *Матрицей инцидентности* ориентированного графа с n вершинами и m ребрами называется матрица $B_{n \times m}$ с n строками и m столбцами, каждый элемент которой b_{ij} определяется по следующей формуле:

$$b_{ij} = \begin{cases} 1, & \text{если вершина } i \text{ является началом ребра } j; \\ -1, & \text{если вершина } i \text{ является концом ребра } j; \\ 2, & \text{если вершина } i \text{ — начало и конец ребра } j; \\ 0, & \text{если вершина } i \text{ и ребро } j \text{ не инцидентны.} \end{cases}$$

Перенумеруем дуги графа, представленного на рис. 6.5 (рис. 6.6). Тогда матрица инцидентности будет иметь вид

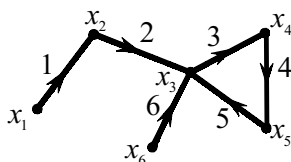


Рис. 6.6

$$B = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}.$$

1.2. Метод ветвей и границ. Задача коммивояжера

Пусть критерий эффективности задачи не может быть представлен в виде функции от управляющих переменных или является достаточно сложной функцией, для которой не существует методов решения соответствующей задачи. В этом случае для нахождения оптимального решения необходимо перебрать все варианты решения задачи. Однако такой перебор может занять много (в некоторых случаях бесконечно много) времени. Например, если требуется составить расписание обработки партии из m деталей на n станках, для выбора оптимального по некоторому критерию расписания надо просмотреть $(m!)^n$ вариантов. Для партии, состоящей из 9 деталей, обрабатываемых на 6 станках, надо перебрать порядка 10^{30} вариантов. Это, конечно, очень большое число.

Для сокращения числа вариантов перебора может быть применен метод ветвей и границ.

Метод ветвей и границ — это метод направленного перебора множества вариантов решений задачи. Графически перебор можно представить в виде *дерева*, т. е. связного графа, не содержащего циклов. *Корень* этого дерева — все множество вариантов, а вершины дерева — подмножества частично упорядоченных вариантов решения.

Допустим, требуется составить расписание обработки четырех деталей.

Для определенности присвоим деталям номера 1, 2, 3, 4. Тогда решением данной задачи будет последовательность номеров, соответствующая расписанию, предпочтительному перед другими по выбранному критерию эффективности.

Корнем дерева решений R является множество всех вариантов расписаний; вершины R_i соответствуют расписаниям, в которых первой обрабатывается деталь с номером i , а вершины R_{ij} — расписаниям, в которых первыми обрабатываются детали с номерами i и j (рис. 6.7).

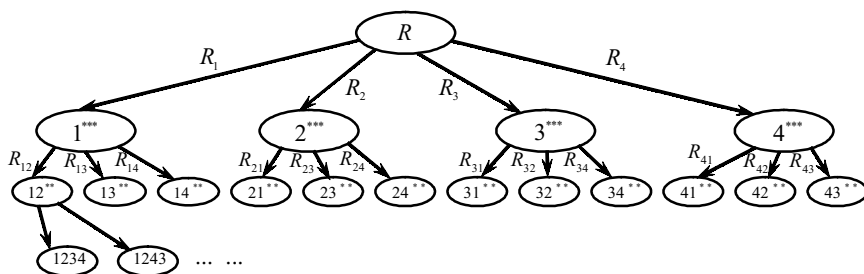


Рис. 6.7

Для нахождения оптимального варианта применяем метод ветвей и границ. Его идея состоит в том, что ветвятся не все вершины. Сначала вершины просматриваются и каждая из них оценивается. Ветвится та вершина, которая получает лучшую оценку. Трудность состоит в получении этой оценки.

Каждой вершине соответствует множество вариантов решений. Каждому варианту решения x соответствует определенное значение

ние критерия эффективности $f(x)$. Лучшее из этих значений (минимальное или максимальное) удобно взять в качестве оценки вершины. Однако подсчитать точно значение f , не перебрав всех вариантов, невозможно. Поэтому используется не точное значение f , а его оценка снизу (при минимизации) или сверху (при максимизации). Оценка снизу называется *оценкой нижней границы множества вариантов*, оценка сверху — *оценкой верхней границы множества вариантов*.

Оценка вершины должна удовлетворять следующим условиям:

1. Она не должна быть больше (при минимизации) или меньше (при максимизации) соответственно минимального или максимального значения функции f для данного подмножества вариантов.

2. Значение оценки для подмножеств нижнего уровня не должно быть меньше (при минимизации) или больше (при максимизации) значения оценки для подмножеств более высокого уровня.

3. Оценка единственного варианта решения на последнем уровне не точно совпадает со значением функции f для этого решения.

Алгоритм метода ветвей и границ:

1. Построить вершины первого уровня. Для каждой вершины подсчитать оценку нижней (верхней) границы. Продолжить ветвление той вершины, которой соответствует лучшая (минимальная или максимальная) оценка.

2. Для всех вершин i -го уровня ($i = 2, 3, \dots$) подсчитать оценку. Продолжить ветвление той висячей вершины уровня $i, i - 1, \dots, 1$, которой соответствует лучшая (минимальная или максимальная) оценка.

3. Действия, указанные в пункте 2, повторять до тех пор, пока не будет получено точное решение на последнем уровне. Для него подсчитать точное значение f . Если это значение не хуже оценок оставшихся висячих вершин, то найдено оптимальное решение. Если это значение строго лучше, то оптимальное решение единственно. Если значение функции f для вершин последнего уровня не лучше значения оценок оставшихся висячих вершин, то переходят к пункту 2.

Метод ветвей и границ не исключает того, что в ходе решения задачи не будет произведен полный перебор вариантов.

Классическим примером применения метода ветвей и границ является **задача коммивояжера**: *коммивояжер, или торговый агент, реализует товар в нескольких населенных пунктах, расположенных на определенном расстоянии друг от друга. Коммивояжер должен объехать все эти пункты по кратчайшему маршруту и вернуться обратно.*

С точки зрения теории графов задача коммивояжера формулируется как *задача нахождения гамильтонова цикла минимальной длины.*

▷ ПРИМЕР 6.1. Решить задачу коммивояжера для пяти пунктов: $m = 5$. Матрица попарных расстояний между пунктами имеет вид

	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
<i>A</i>	0	70	120	110	130
<i>B</i>	70	0	∞	20	50
<i>C</i>	120	∞	0	30	120
<i>D</i>	110	20	30	0	50
<i>E</i>	130	50	120	50	0

Расстояние, равное ∞ , соответствует тому, что не существует маршрута, соединяющего пункты *B* и *C*.

Решение. Путь состоит из пяти звеньев. Длина каждого звена не меньше 20, следовательно, в качестве оценки снизу L можно взять длину минимального звена, умноженную на общее число звеньев:

$$L = 5 \cdot 20 = 100,$$

и на каждом шаге ветвления к длине известного пути добавлять число оставшихся звеньев, умноженное на 20.

Дерево представлено на рис. 6.8. Корнем дерева является множество вариантов путей, исходящих из пункта *A* (так как надо объехать все пункты, то начальный пункт может быть любым). Вершинами дерева являются подмножества частично упорядоченных путей. Буквы, стоящие в каждой вершине, соответствуют пунктам объезда, известным на данном шаге ветвления.

Наименьшую длину (320 единиц) имеют пути

$$A \rightarrow B \rightarrow E \rightarrow D \rightarrow C \rightarrow A$$

и

$$A \rightarrow C \rightarrow D \rightarrow E \rightarrow B \rightarrow A.$$

Вершины $A \rightarrow C \rightarrow E$ и $A \rightarrow E \rightarrow C$ не подверглись ветвлению, так как добавление еще хотя бы одного пункта к данным маршрутам увеличит их длину до значения, большего 320. Из рис. 6.8 видно, что в процессе решения было перебрано достаточно большое число вариантов. Этого можно было бы избежать, взяв более точную оценку нижней границы. Например, оценивая вершину на каждом шаге ветвления, можно было добавлять к уже известной длине пути длину минимального из оставшихся звеньев, умноженную на их число. ►

Аналогичным образом можно решать задачу коммивояжера, если в качестве исходных данных вместо расстояний рассматривать стоимость проезда от одного населенного пункта к другому и выбирать путь, минимизирующий суммарную стоимость проезда.

Рассмотрим постановку **задачи минимизации затрат при производстве и продаже оборудования**. Обозначим:

n — количество изделий;

T — транспортные расходы на перевозку n изделий из первой точки во вторую;

p_1 — цена работ в первой точке, где находится материал;

p_2 — цена работ во второй точке, причем $p_2 < p_1$.

Таким образом, получим:

np_1 — затраты на обработку материала в первой точке;

$np_2 + T$ — затраты на обработку материала во второй точке.

Вычислим минимальную партию, при которой выгоднее будет производить обработку материала во второй точке.

Для экономии средств необходимо выполнение неравенства

$$np_2 + T < np_1.$$

Отсюда получаем нужный нам объем партии: $n > \frac{T}{p_1 - p_2}$.

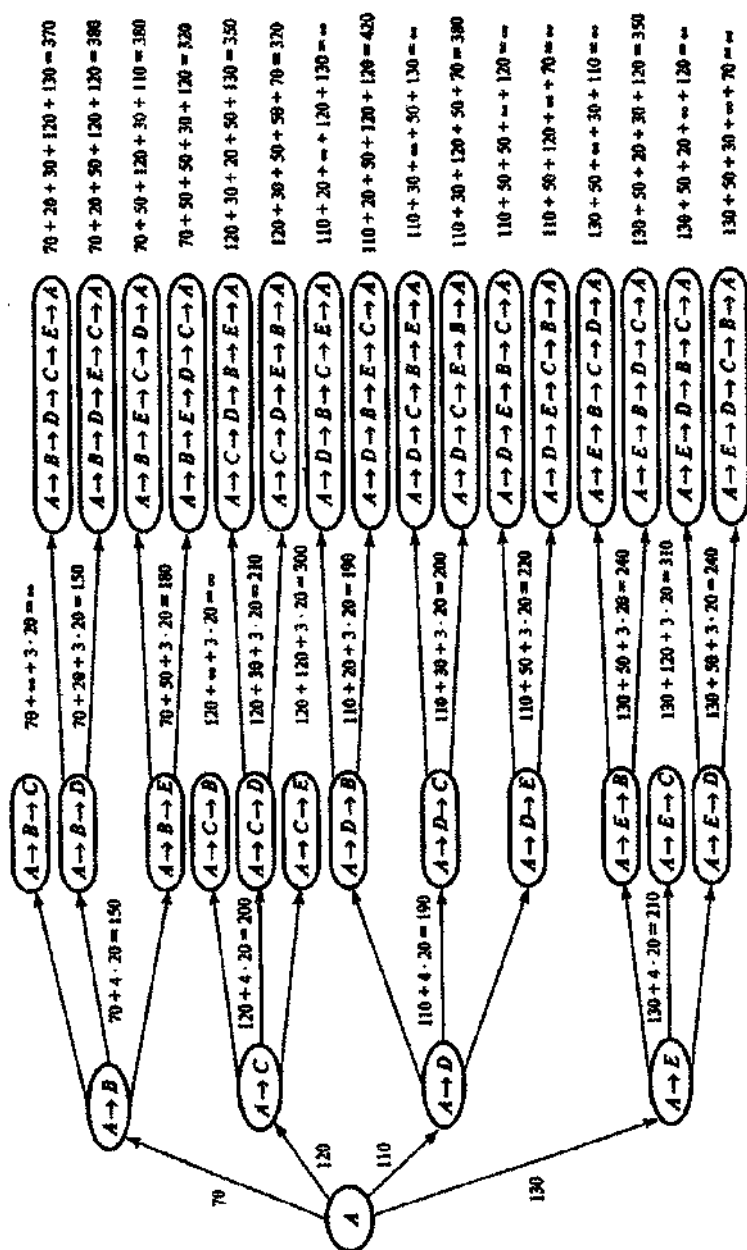
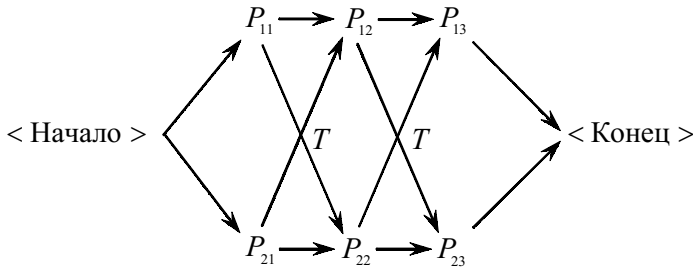


Рис. 6.8

Таким образом, в общем виде наша задача представляет собой задачу минимизации затрат, которые зависят от расценок на обработку материалов в разных точках (зависящих, в свою очередь, от объема партии) и от затрат на транспортировку между всевозможными точками.

Покажем на примере, что не всегда выгодно закупать материал по самой низкой цене, хотя на первый взгляд это противоречит здравому смыслу. В определенных условиях выполнение какой-либо операции по минимальной цене приводит к большим затратам, чем при выполнении этого же действия по более высоким расценкам. Также будет показано, что не всегда целесообразно выполнять какие-нибудь производственные циклы в одном месте; иногда более выгодно потратиться на транспортировку и выполнить эти действия в разных точках.

Рассмотрим граф



где P_{ij} — это цена j -й операции для одной единицы продукции в i -й точке. Для партии из n единиц будем считать затраты равными nP_{ij} (оптовые скидки не учитываются).

Поскольку здесь всего две точки, то транспортные расходы между ними одинаковы. Будем считать, что они равны T .

Зададим конкретные значения:

$P_{11} = 20$ — стоимость материала при $n = 1$ в первой точке;

$P_{12} = 10$ — стоимость сборки при $n = 1$ в первой точке;

$P_{13} = 10$ — стоимость реализации при $n = 1$ в первой точке;

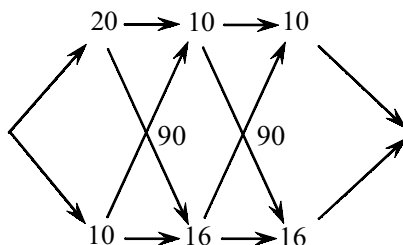
$P_{21} = 10$ — стоимость материала при $n = 1$ во второй точке;

$P_{22} = 16$ — стоимость сборки при $n = 1$ во второй точке;

$P_{23} = 16$ — стоимость реализации при $n = 1$ во второй точке;

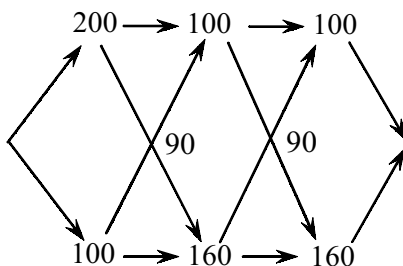
$T = 90$ — затраты на транспортировку продукции из одной точки в другую.

Тогда дерево графа будет выглядеть следующим образом:



Если мы выполним все действия в первой точке, то наши затраты будут равны 40, а если во второй — 42. Любой другой производственный путь приведет к расходам больше 100. Таким образом, в данных условиях выгодно закупать материал по более высокой цене.

Теперь увеличим партию в 10 раз ($n = 10$):



Конечно, если выполнять все операции лишь в одном месте, все равно выгоднее делать это в первой точке. Однако этот вариант не является самым выгодным. Сравним: цена для первой точки равна 400, в то же время, если произвести первую операцию во второй точке, а вторую и третью — в первой, затраты будут 390. Таким образом, не всегда целесообразно выполнять все производственные циклы в одном месте. В данном случае оказалось, что более экономично потратиться на транспортировку и выполнить действия в разных точках.

§2. Моделирование экономических процессов по схеме марковских случайных процессов

2.1. Марковский случайный процесс с дискретными состояниями

Многие операции, которые приходится анализировать при выборе оптимального решения, развиваются как случайные процессы, зависящие от ряда случайных факторов.

Для того чтобы рассчитать числовые параметры, характеризующие эффективность таких операций, нужно построить вероятностную модель явления, учитывающую сопровождающие его случайные факторы.

Для математического описания многих операций, развивающихся в форме случайного процесса, может быть с успехом применен математический аппарат, разработанный в теории вероятностей для так называемых марковских случайных процессов.

Поясним понятие марковского случайного процесса.

Пусть имеется некоторая система S , состояние которой меняется с течением времени (под системой S может пониматься все что угодно: промышленное предприятие, техническое устройство, ремонтная мастерская и т. д.). Если состояние системы S меняется во времени случайным, заранее непредсказуемым образом, говорят, что в системе S протекает *случайный процесс*.

Примеры случайных процессов:

флуктуации цен на фондовом рынке;

обслуживание клиентов в парикмахерской или ремонтной мастерской;

выполнение плана снабжения группы предприятий и т. д.

Конкретное протекание каждого из этих процессов зависит от ряда случайных, заранее непредсказуемых факторов, таких как:

поступление на фондовый рынок непредсказуемых известий о политических изменениях;

случайный характер потока заявок (требований), поступающих со стороны клиентов;

случайные перебои в выполнении плана снабжения и т. д.

ОПРЕДЕЛЕНИЕ. Случайный процесс, протекающий в системе, называется *марковским* (или *процессом без последствия*), если он обладает следующим свойством: для каждого 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 пришла в это состояние (т.е. как развивался процесс в прошлом).

Другими словами, в марковском случайном процессе будущее развитие его зависит только от настоящего состояния и не зависит от “предыстории” процесса.

Рассмотрим пример. Пусть система S представляет собой фондовый рынок, который уже существует какое-то время. Нас интересует, как будет работать система в будущем. Ясно, по крайней мере в первом приближении, что характеристики работы в будущем (вероятности падения цен конкретных акций через неделю) зависят от состояния системы в настоящий момент (здесь могут вмешаться самые различные факторы типа решений правительства или результатов выборов) и не зависят от того, когда и как система достигла своего настоящего состояния (не зависят от характера движения цен на эти акции в прошлом).

На практике часто встречаются случайные процессы, которые, с той или другой степенью приближения можно считать марковскими.

Теория марковских случайных процессов имеет широкий спектр различных приложений. Нас будет интересовать главным образом применение теории марковских случайных процессов к построению математических моделей операций, ход и исход которых существенно зависит от случайных факторов.

Марковские случайные процессы подразделяются на *классы* в зависимости от того, как и в какие моменты времени система S может менять свои состояния.

ОПРЕДЕЛЕНИЕ. Случайный процесс называется *процессом с дискретными состояниями*, если возможные состояния системы $s_1, s_2, s_3, \dots$ можно перечислить (пронумеровать) одно за другим, а сам процесс состоит в том, что время от времени система S скачком (мгновенно) перескакивает из одного состояния в другое.

Например, разработку проекта S осуществляют совместно два отдела, каждый из которых может совершить ошибку. Возможны следующие состояния системы:

- s_1 — оба отдела работают нормально;
- s_2 — первый отдел совершил ошибку, второй работает нормально;
- s_3 — второй отдел совершил ошибку, первый работает нормально;
- s_4 — оба отдела совершили ошибку.

Процесс, протекающий в системе, состоит в том, что она случайным образом в какие-то моменты времени переходит (“перескакивает”) из состояния в состояние. Всего у системы четыре возможных состояния. Перед нами — процесс с дискретными состояниями.

Кроме процессов с дискретными состояниями существуют *случайные процессы с непрерывными состояниями*: для этих процессов характерен постепенный, плавный переход из состояния в состояние. Например, процесс изменения напряжения в осветительной сети представляет собой случайный процесс с непрерывными состояниями.

Мы будем рассматривать только случайные процессы с дискретными состояниями.

При анализе случайных процессов с дискретными состояниями очень удобно пользоваться геометрической схемой — так называемым графом состояний. *Граф состояний* геометрически изображает возможные состояния системы и ее возможные переходы из состояния в состояние.

Пусть имеется система S с дискретными состояниями:

$$s_1, s_2, \dots, s_n.$$

Каждое состояние будем изображать прямоугольником, а возможные переходы (“перескоки”) из состояния в состояние — стрелками, соединяющими эти прямоугольники. Пример графа состояния приведен на рис. 6.9.

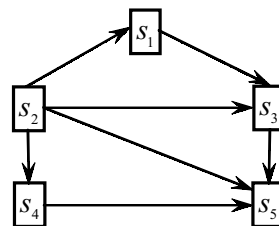


Рис. 6.9

Заметим, что стрелками отмечаются только непосредственные переходы из состояния в состояние; если система может перейти из состояния s_2 в s_3 только через s_1 , то стрелками отмечаются только переходы $s_2 \rightarrow s_1$ и $s_1 \rightarrow s_3$, но не $s_2 \rightarrow s_3$.

Рассмотрим несколько примеров:

1. Система S — фирма, которая может находиться в одном из пяти возможных состояний:

- s_1 — работает с прибылью;
- s_2 — утратила перспективу развития и перестала приносить прибыль;
- s_3 — стала объектом для потенциального поглощения;
- s_4 — находится под внешним управлением;
- s_5 — имущество ликвидируемой фирмы продается на торгах.

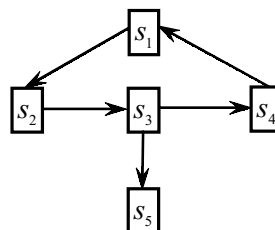


Рис. 6.10

Граф состояний фирмы показан на рис. 6.10.

2. Система S — банк, имеющий два отделения. Возможны следующие состояния системы:

- s_1 — оба отделения работают с прибылью;
- s_2 — первое отделение работает без прибыли, второе работает с прибылью;
- s_3 — второе отделение работает без прибыли, первое работает с прибылью;
- s_4 — оба отделения работают без прибыли.

Предполагается, что улучшение состояния не происходит.

Граф состояний представлен на рис. 6.11. Отметим, что на графе не показан возможный переход из состояния s_1 непосредственно в s_4 , который осуществится, если банк *сразу* будет работать в убыток. Возможностью такого события можно пренебречь, что и подтверждает практика.

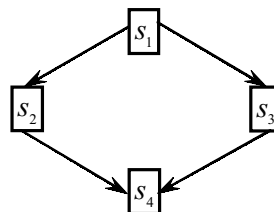


Рис. 6.11

3. Система S — инвестиционная компания, состоящая из двух трейдеров (отделов): I и II; каждый из них может в какой-то

момент времени начать работать в убыток. Если это происходит, то руководство компании немедленно принимает меры для восстановления прибыльной работы отдела.

Возможные состояния системы:

- s_1 — деятельность обоих отделов прибыльна;
- s_2 — первый отдел восстанавливается, второй работает с прибылью;
- s_3 — первый отдел работает с прибылью, второй восстанавливается;
- s_4 — оба отдела восстанавливаются.

Граф состояний системы показан на рис. 6.12.

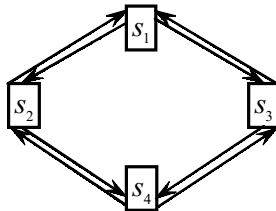


Рис. 6.12

4. В условиях предыдущего примера деятельность каждого трейдера перед тем, как он начнет восстанавливать прибыльную работу отдела, подвергается изучению руководством фирмы в целях принятия мер по ее улучшению.

Состояния системы будем для удобства нумеровать не одним, а двумя индексами; первый будет означать состояния первого трейдера (1 — работает с прибылью, 2 — его деятельность изучается руководством, 3 — восстанавливает прибыльную деятельность отдела); второй — те же состояния для второго трейдера. Например, s_{23} будет означать: деятельность первого трейдера изучается, второй — восстанавливает прибыльную работу.

Возможные состояния системы S :

- s_{11} — деятельность обоих трейдеров приносит прибыль;
- s_{12} — первый трейдер работает с прибылью, деятельность второго изучается руководством компании;

ОПРЕДЕЛЕНИЕ. Случайный процесс называется *процессом с дискретным временем*, если переходы системы из состояния в состояние возможны только в строго определенные, заранее фиксированные моменты времени: $t_1, t_2, \dots$. В промежутки времени между этим моментами система S сохраняет свое состояние.

ОПРЕДЕЛЕНИЕ. Случайный процесс называется *процессом с непрерывным временем*, если переход системы из состояния в состояние возможен в любой, наперед неизвестный, случайный момент t .

Рассмотрим **марковский случайный процесс с дискретными состояниями и дискретным временем**.

Пусть имеется экономическая система S , которая может находиться в состояниях: $s_1, s_2, \dots, s_n$, причем переходы (“перескоки”) системы из состояния в состояние возможны только в моменты: $t_1, t_2, \dots, t_k, \dots$.

Будем называть эти моменты шагами, или этапами, процесса и рассматривать случайный процесс, происходящий в системе S , как функцию целочисленного аргумента: $1, 2, \dots, k, \dots$ (номера шага).

Случайный процесс, происходящий в системе, состоит в том, что в последовательные моменты времени $t_1, t_2, \dots$ система S оказывается в тех или других состояниях, ведя себя, например, следующим образом:

$s_1 \rightarrow s_3 \rightarrow s_5 \rightarrow s_6 \rightarrow s_2 \rightarrow s_3 \rightarrow \dots$

или же

$s_1 \rightarrow s_2 \rightarrow s_4 \rightarrow s_5 \rightarrow s_6 \rightarrow s_2 \rightarrow \dots$ (рис. 6.14).

В общем случае в моменты $t_1, t_2, \dots$ система может не только менять состояние, но и оставаться в прежнем состоянии, например:

$s_1 \rightarrow s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_3 \rightarrow s_3 \rightarrow s_2 \rightarrow s_3 \dots$.

Обозначим $s_i^{(k)}$ — событие, состоящее в том, что после k шагов система находится в состоянии s_i . При любом k события $s_1^{(k)}, s_2^{(k)}, \dots, s_i^{(k)}, \dots, s_n^{(k)}$ образуют полную группу и несовместны (см. подпараграф 1.1 гл. 5).

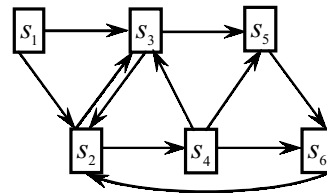


Рис. 6.14

Процесс, происходящий в системе, можно представить как последовательность (цепочку) событий, например: $s_1^{(0)}, s_2^{(1)}, s_1^{(2)}, s_2^{(3)}, s_3^{(4)}, \dots$.

ОПРЕДЕЛЕНИЕ. 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 .

Мы будем описывать марковскую цепь с помощью так называемых вероятностей состояний. Пусть в любой момент времени (после любого, k -го шага) система S может находиться в одном из состояний:

$$s_1, s_2, \dots, s_n,$$

т.е. осуществится одно событие из полной группы несовместных событий:

$$s_1^{(k)}, s_2^{(k)}, \dots, s_n^{(k)}.$$

Обозначим вероятности этих событий:

$p_1(1) = P(s_1^{(1)}); p_2(1) = P(s_2^{(1)}); \dots; p_n(1) = P(s_n^{(1)})$ — вероятности после первого шага;

$p_1(2) = P(s_1^{(2)}); p_2(2) = P(s_2^{(2)}); \dots; p_n(2) = P(s_n^{(2)})$ — вероятности после второго шага;

.....

$p_1(k) = P(s_1^{(k)}); p_2(k) = P(s_2^{(k)}); \dots; p_n(k) = P(s_n^{(k)})$ — вероятности после k -го шага.

Легко видеть, что для каждого номера шага k

$$p_1(k) + p_2(k) + \dots + p_n(k) = 1,$$

так как это вероятности несовместных событий, образующих полную группу.

Будем называть вероятности $p_1(k), p_2(k), \dots, p_n(k)$ *вероятностями состояний*.

Рассмотрим экономический пример.

▷ **ПРИМЕР 6.2.** За неуплату налогов предприятие S подвергается проверке налоговых органов 4 раза: в моменты времени t_1, t_2, t_3, t_4 .

Возможные состояния системы S :

s_1 — предприятие работает, как и раньше;

s_2 — наложен штраф;

s_3 — предприятие находится под временным управлением;

s_4 — предприятие ликвидировано.

Размеченный граф состояний системы показан на рис. 6.15.

В начальный момент времени система находится в состоянии s_1 . Определить вероятности состояний предприятия после четырех проверок.

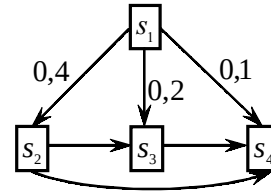


Рис. 6.15

Решение. Обозначим p_{ij} — вероятность перехода системы из состояния i в состояние j . Из графа состояний имеем:

$$p_{12} = 0,4; p_{13} = 0,2; p_{14} = 0,1;$$

и

$$p_{11} = 1 - (p_{12} + p_{13} + p_{14}) = 0,3.$$

Аналогично предыдущему находим:

$$p_{21} = 0; \quad p_{22} = 0,4; \quad p_{23} = 0,4; \quad p_{24} = 0,2;$$

$$p_{31} = 0; \quad p_{32} = 0; \quad p_{33} = 0,3; \quad p_{34} = 0,7;$$

$$p_{41} = 0; \quad p_{42} = 0; \quad p_{43} = 0; \quad p_{44} = 1.$$

Таким образом матрица вероятностей перехода имеет вид:

$$\|p_{ij}\| = \begin{pmatrix} 0,3 & 0,4 & 0,2 & 0,1 \\ 0 & 0,4 & 0,4 & 0,2 \\ 0 & 0 & 0,3 & 0,7 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Так как в начальный момент система S находится в состоянии s_1 , то $p_1(0) = 1$. Вероятности состояний после первой проверки берутся из первой строки матрицы:

$$p_1(0) = 0,3; p_2(1) = 0,4; p_3(1) = 0,2; p_4(1) = 0,1.$$

Вероятности состояний после второй проверки:

$$p_1(2) = p_1(1)p_{11} = 0,3 \cdot 0,3 = \mathbf{0,09};$$

$$p_2(2) = p_1(1)p_{12} + p_2(1)p_{22} = 0,3 \cdot 0,4 + 0,4 \cdot 0,4 = \mathbf{0,28};$$

$$p_3(2) = p_1(1)p_{13} + p_2(1)p_{23} + p_3(1)p_{33} = 0,3 \cdot 0,2 + 0,4 \cdot 0,4 + 0,2 \cdot 0,3 = \mathbf{0,28};$$

$$p_4(2) = p_1(1)p_{14} + p_2(1)p_{24} + p_3(1)p_{34} + p_4(1)p_{44} = 0,3 \cdot 0,1 + 0,4 \cdot 0,2 + 0,2 \cdot 0,7 + 0,1 \cdot 1 = \mathbf{0,35}.$$

Вероятности состояний после третьей проверки:

$$p_1(3) = p_1(2)p_{11} = 0,09 \cdot 0,3 = \mathbf{0,027};$$

$$p_2(3) = p_1(2)p_{12} + p_2(2)p_{22} = 0,03 \cdot 0,4 + 0,28 \cdot 0,4 = \mathbf{0,148};$$

$$p_3(3) = p_1(2)p_{13} + p_2(2)p_{23} + p_3(2)p_{33} = 0,09 \cdot 0,2 + 0,28 \cdot 0,4 + 0,28 \times \\ \times 0,2 = \mathbf{0,214};$$

$$p_4(3) = p_1(2)p_{14} + p_2(2)p_{24} + p_3(2)p_{34} + p_4(2)p_{44} = 0,09 \cdot 0,1 + 0,28 \cdot 0,2 + \\ + 0,28 \cdot 0,7 + 0,35 \cdot 1 = \mathbf{0,611};$$

Вероятности состояний после четвертой проверки:

$$p_1(4) = p_1(3)p_{11} = 0,027 \cdot 0,3 = \mathbf{0,0081};$$

$$p_2(4) = p_1(3)p_{12} + p_2(3)p_{22} = 0,027 \cdot 0,4 + 0,148 \cdot 0,4 = \mathbf{0,0700};$$

$$p_3(4) = p_1(3)p_{13} + p_2(3)p_{23} + p_3(3)p_{33} = 0,027 \cdot 0,2 + 0,148 \cdot 0,4 + 0,214 \times \\ \times 0,3 = \mathbf{0,1288};$$

$$p_4(4) = p_1(3)p_{14} + p_2(3)p_{24} + p_3(3)p_{34} + p_4(3)p_{44} = 0,027 \cdot 0,1 + 0,148 \cdot 0,2 + \\ + 0,214 \cdot 0,7 + 0,611 \cdot 1 = \mathbf{0,7931};$$

Таким образом, нами получены вероятности всех исходов проверок:

предприятие работает, как и раньше: $p_1(4) \approx 0,008$;

наложен штраф: $p_2(4) \approx 0,070$;

предприятие находится под временным управлением: $p_3(4) \approx 0,129$;

предприятие ликвидировано: $p_4(4) \approx 0,793$. ►

Мы рассмотрели *о д н о р о д н у ю* марковскую цепь, для которой вероятности перехода от шага к шагу не меняются.

Рассмотрим теперь общий случай — *н е о д н о р о д н у ю* марковскую цепь, для которой вероятности перехода p_{ij} меняются от шагу к шагу. Обозначим $p_{ij}^{(k)}$ — вероятность перехода системы из состояния s_i в состояние s_j на k -м шаге, т.е. условную вероятность

$$p_{ij}^{(k)} = P(s_j^{(k)} / s_i^{(k-1)}).$$

Предположим, что матрицы вероятностей перехода на каждом шаге заданы. Тогда вероятность того, что система S после k шагов будет находиться в состоянии s_j выразится формулой:

$$p_j^{(k)} = \sum_i p_i(k-1)p_{ij}^{(k)} \quad (i = 1, 2, \dots, n). \quad (6.1)$$

Формула (6.1) отличается от аналогичной формулы для однородной цепи Маркова только тем, что в ней фигурируют веро-

ятности перехода, зависящие от номера шага k . Вычисления по формуле (6.1) ничуть не сложнее, чем в случае одноуровневой цепи.

▷ ПРИМЕР 6.3. Налоговые органы организовали три проверки фирмы, которая может находиться в тех же четырех состояниях s_1, s_2, s_3, s_4 , что и в предыдущем примере, но вероятности перехода для трех последовательных проверок различны и заданы тремя матрицами:

$$\|p_{ij}^{(1)}\| = \begin{vmatrix} 0,3 & 0,4 & 0,2 & 0,1 \\ 0 & 0,4 & 0,4 & 0,2 \\ 0 & 0 & 0,3 & 0,7 \\ 0 & 0 & 0 & 1 \end{vmatrix}; \quad \|p_{ij}^{(2)}\| = \begin{vmatrix} 0,1 & 0,4 & 0,3 & 0,2 \\ 0 & 0,2 & 0,5 & 0,3 \\ 0 & 0 & 0,2 & 0,8 \\ 0 & 0 & 0 & 1 \end{vmatrix};$$

$$\|p_{ij}^{(3)}\| = \begin{vmatrix} 0,05 & 0,3 & 0,4 & 0,25 \\ 0 & 0,1 & 0,6 & 0,3 \\ 0 & 0 & 0,1 & 0,9 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

В начальный момент времени система находится в состоянии s_1 . Найти вероятности состояний после трех проверок налоговых органов.

Решение. Имеем:

$$p_1(1) = 0,3; \quad p_2(1) = 0,4; \quad p_3(1) = 0,2; \quad p_4(1) = 0,1.$$

$$p_1(2) = p_1(1)p_{11}^{(2)} = 0,3 \cdot 0,1 = \mathbf{0,03};$$

$$p_2(2) = p_1(1)p_{12}^{(2)} + p_2(1)p_{22}^{(2)} = 0,3 \cdot 0,4 + 0,4 \cdot 0,2 = \mathbf{0,20};$$

$$p_3(2) = p_1(1)p_{13}^{(2)} + p_2(1)p_{23}^{(2)} + p_3(1)p_{33}^{(2)} = 0,3 \cdot 0,3 + 0,4 \cdot 0,5 + 0,2 \times \\ \times 0,2 = 0,33;$$

$$p_4(2) = p_1(1)p_{14}^{(2)} + p_2(1)p_{24}^{(2)} + p_3(1)p_{34}^{(2)} + p_4(1)p_{44}^{(2)} = \\ = 0,3 \cdot 0,2 + 0,4 \cdot 0,3 + 0,2 \cdot 0,8 + 0,1 \cdot 1 = 0,44;$$

$$p_1(3) = p_1(2)p_{11}^{(3)} = 0,03 \cdot 0,05 \approx 0,002;$$

$$p_2(3) = p_1(2)p_{12}^{(3)} + p_2(2)p_{22}^{(3)} = 0,03 \cdot 0,3 + 0,20 \cdot 0,1 = 0,029;$$

$$\begin{aligned} p_3(3) &= p_1(2)p_{13}^{(3)} + p_2(2)p_{23}^{(3)} + p_3(2)p_{33}^{(3)} = \\ &= 0,03 \cdot 0,4 + 0,02 \cdot 0,6 + 0,33 \cdot 0,1 = \mathbf{0,165}; \end{aligned}$$

$$\begin{aligned} p_4(3) &= p_1(2)p_{14}^{(3)} + p_2(2)p_{24}^{(3)} + p_3(2)p_{34}^{(3)} + p_4(2)p_{44}^{(3)} = \\ &= 0,03 \cdot 0,25 + 0,20 \cdot 0,3 + 0,33 \cdot 0,9 + 0,44 \cdot 1 \approx \mathbf{0,804}. \end{aligned}$$

Итак, вероятности состояний после трех проверок:
 $p_1(3) \approx 0,002$; $p_2(3) \approx 0,029$; $p_3(3) \approx 0,165$; $p_4(3) \approx 0,804$. ►

§ 3. Некоторые аспекты теории игр

3.1. Основные понятия математической теории игр

Классификация игр. Во многих задачах исследования операций приходится сталкиваться с проблемой *принятия решения в условиях неопределенности*.

Неопределенными могут быть как условия выполнения операции, так и сознательные действия противников или других участников, от которых зависит успех операции. Кроме того, неопределенность в той или другой степени может относиться также и к целям (задачам) операции, успех которой далеко не всегда может быть исчерпывающим образом охарактеризован одним единственным числом — показателем эффективности.

Задачами о принятии решений в условиях неопределенности занимается *теория игр и статистических решений*. *Предметом* теории игр являются ситуации, в которых важную роль играют конфликты и совместные действия.

Формализованное описание конфликта представляет собой его математическую модель, которую называют игрой. Математическая модель любого конфликта, в том числе социально-экономического, должна описывать:

а) множество заинтересованных сторон (будем называть их игроками; в литературе по теории игр они именуются также субъектами, лицами, сторонами, участниками);

б) возможные действия каждой из сторон, именуемые также стратегиями или ходами;

в) интересы сторон, представленные функциями выигрыша (платежа) для каждого из игроков.

Различные виды игр можно классифицировать по следующим признакам:

а) число игроков. Различают игры с *двумя, тремя участниками и более*. Возможны также игры с *бесконечным числом игроков*;

б) количество стратегий. Различают конечные и бесконечные игры. В *конечных играх* игроки располагают конечным числом возможных стратегий (например, в игре в орлянку игроки имеют

по два возможных хода — они могут выбрать “орла” или “решку”). Сами стратегии в конечных играх нередко называются *чистыми*. Соответственно, в *бесконечных играх* игроки имеют бесконечное число возможных стратегий (например, в ситуации “продавец — покупатель” каждый игрок может назвать любую цену и количество продаваемого (покупаемого) товара);

в) свойства функций выигрыша (платежных функций). Различают игры с нулевой суммой, с постоянной разностью и с ненулевой суммой. Игры, в которых выигрыш одного из игроков равен проигрышу другого, т.е. налицо прямой конфликт между игроками, называются *играми с нулевой суммой* или *антагонистическими*. Игры в орлянку или очко — типичные примеры антагонистических игр. Прямой противоположностью играм такого типа являются *игры с постоянной разностью*, в которых игроки и выигрывают, и проигрывают одновременно, так что им выгодно действовать сообща. Между этими крайними случаями находится множество *игр с ненулевой суммой*, где имеются и конфликты, и согласованные действия игроков;

г) возможность предварительных переговоров между игроками. Различают кооперативные и некооперативные игры. Игра называется *кооперативной*, если до начала игры игроки образуют коалиции и принимают взаимообязывающие соглашения о своих стратегиях. Игра, в которой игроки не могут координировать свои стратегии подобным образом, называется *некооперативной*. Очевидно, что все антагонистические игры могут служить примером некооперативных игр. Примером кооперативной игры может служить ситуация образования коалиций в парламенте для принятия путем голосования решения, так или иначе затрагивающего интересы участников голосования.

Формальное представление игр. Рассмотрим формальное описание некоторых элементов теории игр. Множество всех игроков, обозначаемое L , в случае конечного их числа может быть задано простым перечислением, например: $L = \{1, 2\}$ при игре в орлянку, $L = \{\text{Продавец, Покупатель}\}$ в ситуации “монополия — монополия”, $L = \{1, 2, \dots, n\}$ в случае анализа результатов голосования в парламенте.

Обозначим через X_l множество стратегий игрока l . При игре в орлянку каждый игрок располагает двумя стратегиями: $X_l = \{\text{Орел}, \text{Решка}\}$; каждый участник голосования имеет выбор на множестве стратегий $\{\text{За}, \text{Против}\}$. В случае взаимодействия на рынке как Продавец, так и Покупатель могут назначать некоторую неотрицательную цену на продаваемый (покупаемый) товар, т.е. множество стратегий каждого из них X_l ; $p_l > 0$.

Каждый игрок выбирает некоторую стратегию X_l , в результате чего складывается набор стратегий $X = \{X_1, X_2, \dots, X_n\}$, называемый *ситуацией*. Так, ситуацию в парламенте описывает список $\{\text{За}, \text{За}, \text{Против}, \text{За}, \dots\}$, полученный в итоге проведенного голосования.

Заинтересованность игроков в ситуациях проявляется в том, что каждому игроку l в каждой ситуации X приписывается число, выражающее степень удовлетворения его интересов в данной ситуации. Это число называется *выигрышем* игрока l и обозначается через $h_l(x)$, а соответствие между набором ситуаций и выигрышем игрока l называется *функцией выигрыша* (платежной функцией) этого игрока H_l .

Рассмотрим игру двух лиц. Функции выигрыша каждого из игроков удобно представлять в виде **матрицы выигрышей**, где строки — это стратегии одного игрока, столбцы — стратегии другого игрока, а в клетках матрицы указаны выигрыши каждого игрока в каждой из образующихся ситуаций. (Данная форма представления конечных игр двух лиц объясняет общее для них название — *матричные игры*.)

При игре в орлянку каждый игрок имеет по две стратегии: “Орел” и “Решка”. Если игроки выбирают одинаковые стратегии, т.е. в случаях, если оба говорят “Орел” или оба говорят “Решка”, Игрок 1 выигрывает 1000 руб., а Игрок 2 проигрывает 1000 руб. В ситуациях, когда оба игрока выбирают различные стратегии, Игрок 1 проигрывает 1000 руб., а Игрок 2 соответственно эту 1000 руб. выигрывает.

Для Игрока 1 матрица выигрышей H_1 выглядит следующим образом:

$$\begin{array}{c}
 \underbrace{\text{Стратегии Игрока 2}} \\
 \text{Орел} \quad \text{Решка} \\
 \text{Стратегии Игрока 1} \left\{ \begin{array}{l} \text{Орел} \begin{pmatrix} 1000 & -1000 \end{pmatrix} \\ \text{Решка} \begin{pmatrix} -1000 & 1000 \end{pmatrix} \end{array} \right.
 \end{array}$$

Соответственно, для Игрока 2 матрица выигрышей H_2 имеет вид

$$\begin{array}{c}
 \underbrace{\text{Стратегии Игрока 2}} \\
 \text{Орел} \quad \text{Решка} \\
 \text{Стратегии Игрока 1} \left\{ \begin{array}{l} \text{Орел} \begin{pmatrix} -1000 & 1000 \end{pmatrix} \\ \text{Решка} \begin{pmatrix} 1000 & -1000 \end{pmatrix} \end{array} \right.
 \end{array}$$

Для игр с нулевой суммой, в которых выигрыш одного игрока равен проигрышу другого, выполняется соотношение $H_1 = -H_2$. Такова и игра в орлянку.

Часто для наглядности матрицы выигрышей обоих игроков совмещают в одну:

$$\begin{array}{c}
 \underbrace{\text{Стратегии Игрока 2}} \\
 \text{Орел} \quad \text{Решка} \\
 \text{Стратегии Игрока 1} \left\{ \begin{array}{l} \text{Орел} \begin{pmatrix} (1000, -1000) & (-1000, 1000) \end{pmatrix} \\ \text{Решка} \begin{pmatrix} (-1000, 1000) & (1000, -1000) \end{pmatrix} \end{array} \right.
 \end{array}$$

Эта матрица дает полное представление о всей игре. В каждой ее клетке слева указан выигрыш Игрока 1, справа — выигрыш Игрока 2.

3.2. Решение матричных антагонистических игр

Всегда в теории игр в качестве основного предположения допускается, что каждый игрок стремится обеспечить себе максимально возможный выигрыш при любых действиях партнера.

Пусть имеется конечная антагонистическая игра с матрицей выигрыша $H_1 = H$ Игрока 1 и, соответственно, матрицей выигрыша $H_2 = -H$ Игрока 2. Пусть Игрок 1 считает, что какую бы стратегию он ни выбрал, Игрок 2 выберет стратегию, максимизирующую собственный выигрыш и тем самым минимизирующую выигрыш Игрока 1. Оптимальная стратегия Игрока 1, которая обеспечит ему максимальный выигрыш вне зависимости от стратегии противника, будет состоять в выборе стратегии с самым большим платежом. Таким образом, Игрок 1 выбирает i -ю стратегию, которая является решением задачи

$$\max_i \min_j h_{ij}.$$

Точно так же Игрок 2 стремится обеспечить себе максимальный выигрыш (или, что эквивалентно, минимальный проигрыш) вне зависимости от стратегии, избранной партнером. Его оптимальной стратегией будет столбец матрицы H_2 с наименьшим значением максимального платежа. Таким образом, Игрок 2 выберет j -ю стратегию, которая является решением задачи

$$\min_j \max_i h_{ij}.$$

Итак, если Игрок 1 придерживается избранной стратегии (называемой *максиминной стратегией*), его выигрыш в любом случае будет не меньше максиминного значения (называемого также *нижней ценой игры*), т.е.

$$h_{ij} \geq \max_i \min_j h_{ij}.$$

Аналогично, если Игрок 2 придерживается своей *минимаксной стратегией*, его проигрыш будет не больше минимаксного значения (называемого *верхней ценой игры*), т.е.

$$h_{ij} \leq \min_j \max_i h_{ij}.$$

Если верхняя цена игры равна нижней, т.е.

$$\min_j \max_i h_{ij} = \max_i \min_j h_{ij} = h^*,$$

то оба игрока получают свои гарантированные платежи, а значение h^* называется *ценой игры*. Элемент h_{ij} матрицы выигрышей, соответствующий максиминной и минимаксной стратегиям,

называется *седловой точкой* матрицы H (это объясняется тем, что график функции выигрыша в точке h_{ij} по форме напоминает седло: функция убывает при изменении одной из переменных и возрастает при изменении другой переменной). В случае если цена антагонистической игры равна нулю, игра называется *справедливой*.

Оптимальные стратегии легко находятся для небольшого числа стратегий; с ростом количества стратегий вычисления становятся достаточно сложными. Для поиска оптимальных стратегий рекомендуется несколько подходов. Опишем один из них.

Для уменьшения размерности игры используется **доминирование строк и столбцов**. Обычно говорят, что k -я строка матрицы H *доминирует* i -ю строку (т.е. одна чистая стратегия доминирует другую), если:

$$\begin{array}{ll} h_{ij} \leq h_{kj} & \text{для всех } j; \\ h_{ij} < h_{kj} & \text{хотя бы для одного } j. \end{array}$$

Аналогично l -й столбец *доминирует* j -й столбец, если:

$$\begin{array}{ll} h_{il} \leq h_{jl} & \text{для всех } i; \\ h_{il} < h_{jl} & \text{хотя бы для одного } i. \end{array}$$

Смысл этого определения состоит в том, что доминирующая стратегия никогда не хуже, а в некоторых случаях даже лучше, чем доминируемая стратегия. Отсюда следует, что игроку нет необходимости использовать доминируемую стратегию. В самом деле, будут существовать оптимальные смешанные стратегии, при которых вероятность использования доминируемых строк и столбцов равна нулю, и при решении игры все доминируемые строки и столбцы могут быть отброшены, что позволяет уменьшить размеры матрицы. (Этот подход может использоваться также при поиске решения игры в чистых стратегиях.)

Рассмотрим, например, игру со следующей матрицей:

$$\begin{pmatrix} 5 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 2 & 4 \end{pmatrix}$$

Третья строка этой матрицы доминирует вторую. Исключение второй строки приводит к матрице

$$\begin{pmatrix} 5 & 2 & 3 \\ 3 & 2 & 4 \end{pmatrix}$$

Третий столбец в этой урезанной матрице доминирует второй, и исключение второго столбца дает

$$\begin{pmatrix} 5 & 3 \\ 3 & 4 \end{pmatrix}$$

В итоге, если можно найти решение для полученной игры, его легко использовать для решения исходной игры, просто приписав исключенным строкам и столбцам нулевые вероятности.

3.3. Кооперативные и некооперативные игры

В игре с ненулевой суммой уже необязательно, чтобы один из участников выигрывал, а другой проигрывал; напротив, игроки могут и выигрывать, и проигрывать одновременно. Поскольку интересы игроков теперь не являются полностью противоположными, их поведение становится более разнообразным. Так, например, если в игре с нулевой суммой каждому игроку невыгодно было сообщать другому свою стратегию (это могло уменьшить его выигрыш), то в игре с ненулевой суммой становится желательным как-то координировать свои действия с партнером или каким-либо способом влиять на его действия.

Игры с ненулевой суммой могут быть кооперативными и некооперативными. В *некооперативных* играх игроки принимают решения независимо друг от друга либо потому, что соглашение невозможно, либо потому, что оно запрещено правилами игры.

Одним из подходов в решении некооперативных игр является определение **точек равновесия**. В общем случае пара стратегий X, Y для Игрока 1 и Игрока 2 называется *точкой равновесия по Нэшу*, если ни одному из игроков невыгодно отклоняться от своей стратегии в одиночку, т.е. если

$$h_1(x, y^*) \leq h_1(x^*, y^*) \text{ и } h_2(x^*, y) \leq h_2(x^*, y^*)$$

для $\forall x \in X$ и $\forall y \in Y$ ($x^* \in X^*, y^* \in Y$).

(Это определение равновесия сохраняется и для игры n лиц.)

Известно, что для любой некооперативной игры с ненулевой суммой (такие игры называются также *биматричными*) всегда существует по крайней мере одна равновесная пара смешанных стратегий. В общем случае равновесное решение может быть неединственным и каждому из них могут соответствовать различные значения выигрыша каждого игрока.

Основная задача в *кооперативной* игре состоит в дележе общего выигрыша между членами коалиции.

В случае игры двух лиц предполагается, что два игрока не могут воздействовать друг на друга, пока не придут к некоторому соглашению.

Таким образом, игра определяется как множество S в пространстве

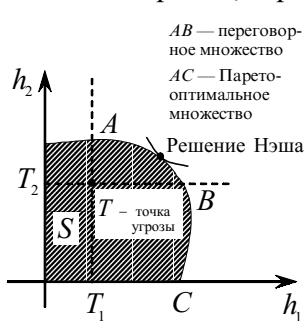


Рис. 6.16

переменных h_1 и h_2 , представляющее общие выигрыши (рис. 6.16). Кроме того, заданы два числа T_1, T_2 , определяющие выигрыш, который каждый из игроков может получить, не вступая в коалицию с партнером. Обычно предполагают, что множество S является замкнутым, выпуклым и ограниченным сверху. Точка T с координатами (T_1, T_2) называется *точкой угрозы* (см. рис. 6.16).

На множестве возможных выигрышей выделяется *множество Парето-оптимальных решений*, т.е. множество точек, принадлежащих S , для которых увеличение выигрыша одного из игроков возможно только за счет уменьшения выигрыша его партнера. Множество таких точек образует *северо-восточная* граница множества S .

Все точки Парето-оптимального множества, находящиеся одновременно *выше и правее* точки угрозы T , образуют так называемое *переговорное множество*. Очевидно, что игрокам нет смысла договариваться относительно решений, не принадлежащих переговорному множеству, либо потому, что положение одного из игроков может быть улучшено при сохранении положе-

ния его партнера, либо потому, что по крайней мере для одного из игроков теряется смысл вступать в коалицию с партнером, так как нехудших результатов он может достичь и в одиночку.

Наконец, на переговорном множестве выделяется *точка решения Нэша* — точка, где достигается максимум произведения превышения выигрышей каждого из игроков над платежами, которые могут быть получены без вступления в коалицию:

$$\max(h_1 - T_1)(h_2 - T_2).$$

В теории игр показано, что если множество возможных платежей S выпукло, замкнуто и ограничено сверху, то точка Нэша существует и единственна. Точка Нэша представляет одно из возможных решений кооперативной игры, от которого нет оснований отказываться ни одному из игроков.

3.4. Связь теории игр с анализом проблем микроэкономики

Для анализа проблем микроэкономики, в том числе для анализа рыночного равновесия как кооперативной игры многих лиц используются основы математической теории игр.

Рассмотрим экономику, в которой имеются два субъекта (Игрок 1 и Игрок 2) и два товара (блага): $(x_1$ и $x_2)$. У каждого игрока имеется своя функция полезности (см. подпараграф 6.3 главы 4), заданная на наборе товаров: $h_1(x_1, x_2)$, $h_2(x_1, x_2)$; предполагается, что обе эти функции непрерывны и монотонны по каждой из переменных и выпуклы (см. подпараграф 2.1 главы 4). В начале игры в экономике имеется общее количество x_1 первого товара и x_2 — второго товара. Предположим, что это начальное количество благ как-то распределено между игроками: Игрок 1 обладает количеством x_1^1 первого товара и x_2^1 — второго, Игрок 2 — количеством x_1^2 и x_2^2 соответственно, так что $x_1^1 + x_1^2 = x_1$ и $x_2^1 + x_2^2 = x_2$.

Теперь возникает вопрос: могут ли игроки путем обмена товарами улучшить свое положение, т.е. увеличить значение функций полезности h_1 и h_2 по сравнению с начальными уровнями $h_1(x_1^1, x_2^1)$ и $h_2(x_1^2, x_2^2)$?

Для наглядного представления экономики с двумя игроками и двумя товарами традиционно используется так называемый *ящик Эджворта* (рис. 6.17). В ящике Эджворта длина горизонтальной оси, соответствующей первому товару, равна общему количеству этого товара x_1 , длина вертикальной оси — общему количеству товара x_2 . Выделенный участок является множеством всех возможных распределений имеющихся товаров между двумя игроками. Нижний левый угол считается началом координат для Игрока 1, верхний правый угол — началом координат для Игрока 2.

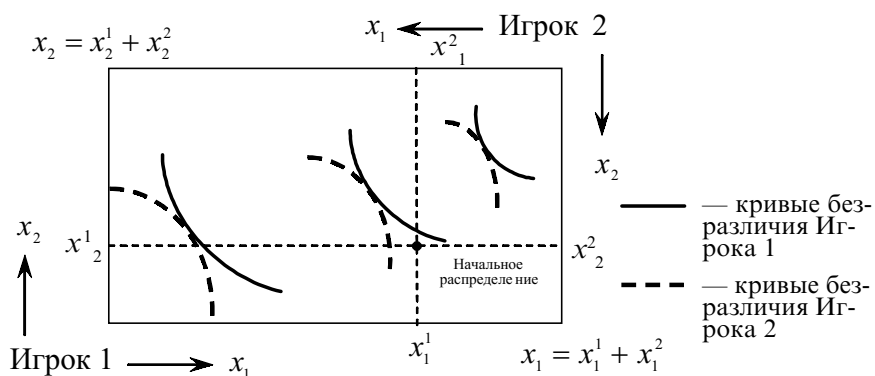


Рис. 6.17

На выделенном участке представлены также два множества кривых безразличия (линий уровня функций выигрыша), принадлежащих каждому из игроков. При этом точка начального распределения товаров имеет координаты (x_1^1, x_2^1) в системе отсчета Игрока 1 (и, соответственно, (x_1^2, x_2^2) в системе отсчета Игрока 2).

Подробно проблему эффективного распределения товаров между игроками мы здесь рассматривать не будем.

Скажем лишь, что в случае двух игроков Парето-оптимальное решение может быть найдено с помощью фиксации уровня полезности одного из игроков (скажем, Игрока 2) и поиска максимума функции полезности другого игрока. Это означает, что необхо-

можно найти такую точку на фиксированной кривой безразличия Игрока 2, в которой Игрок 1 получает максимум своей функции полезности (рис. 6.18).

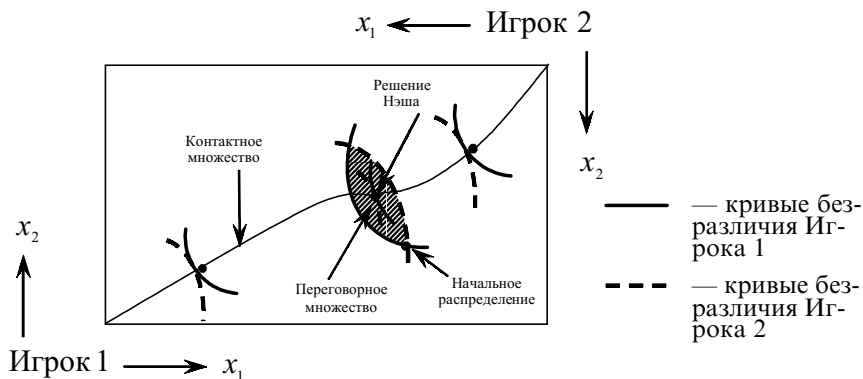


Рис. 6.18

Такой точкой является точка, где кривые безразличия касаются друг друга, так как в противном случае Игрок 1 может, “продвигаясь” вдоль фиксированной линии уровня Игрока 2 внутрь, увеличить значение своей функции полезности. Это свойство можно доказать и математически: максимум функции полезности Игрока 1 при фиксированном уровне полезности Игрока 2 достигается в точке, в которой дифференциалы этих функций равны, т.е. в точке, где кривые безразличия имеют общую касательную.

Опираясь на этот факт, можно выяснить, что игроки могут улучшить свое первоначальное положение, обмениваясь товарами, а Игроку 1 выгодно уступить Игроку 2 некоторое количество товара x_1 в обмен на товар x_2 .

На нескольких примерах мы показали, насколько проблемы рыночного взаимодействия близки к проблемам математической теории игр и могут быть эффективно описаны и исследованы с использованием ее аппарата.

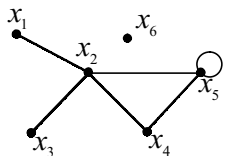
Мы заканчиваем этот раздел математики. Из работ [12, 13] вы можете узнать, какие математические модели можно построить, используя уже пройденные разделы высшей математики.

Контрольные вопросы

1. Какие графы называют регулярными, какие — двудольными?
2. Какие процессы описываются гамильтоновыми цепями?
3. В чем суть метода ветвей и границ? Как решается задача коммивояжера?
4. Что называют марковской цепью, графом состояний?
5. Каковы основные характеристики систем массового обслуживания?
6. Что является предметом теории игр?
7. Какое множество называется Парето-оптимальным?
8. В чем состоит связь между теорией игр и проблемами микроэкономики?

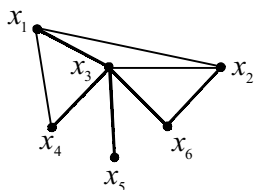
Упражнения

1. Перечислить ребра графа

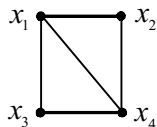


Чему равны степени вершин данного графа?

2. Какая из последовательностей является маршрутом в представленном ниже графе?



3. Составить матрицу смежности для графа



4. Нарисовать граф, соответствующий матрице A :

$$A = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}.$$

5. Банк имеет два филиала. Каждый из них работает независимо друг от друга. Описать все состояния системы, построить граф состояний.

6. В качестве системы S рассмотрим два компьютера I и II. Если операционная система на компьютере работает со сбоями, то ее тестируют на предмет локализации неполадок. Затем она восстанавливается. Описать все состояния системы (состояния помечать двумя цифрами: первая цифра для первого компьютера, вторая — для второго).

7. Дана матрица переходных вероятностей:

$$\|p_{ij}\| = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 1 \end{bmatrix}.$$

Найти вероятности состояний на первом шаге

$$\left\{ p(0) = \left(\frac{4}{5}, \frac{1}{5}, 0 \right) \right\}.$$

8. Имеются два игрока A и B . Игрок A выбирает один из двух шариков (красный или синий), игрок B угадывает цвет. Если игрок A загадал красный шарик и игрок B это угадал, то игрок A платит игроку B 2 руб., в противном случае игрок B платит игроку A 3 руб. Если игрок A загадал синий шарик и игрок B это угадал, то игрок A платит игроку B 3 руб., в противном случае игрок B платит игроку A 2 руб. Составить платежную матрицу относительно игрока A .

Ответы к упражнениям

1. $U = \{(x_1, x_2); (x_2, x_3); (x_2, x_4); (x_2, x_5); (x_4, x_5); (x_5, x_5)\};$

$$d(x_1) = 1, d(x_2) = 4, d(x_3) = 1, d(x_4) = 2, d(x_5) = 4, d(x_6) = 0.$$

2. $(x_1; x_2; x_3; x_6; x_2).$

3.
$$\begin{array}{c} 1 \quad 2 \quad 3 \quad 4 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}.$$

5. s_1 — оба филиала работают;

s_2 — первый филиал закрылся, второй работает;

s_3 — первый филиал работает, второй закрылся;

s_4 — оба филиала закрыты головным банком.

6. s_{11} — оба компьютера работают;

s_{12} — I работает, II тестируется;

s_{13} — I работает, II восстанавливается;

s_{21} — I тестируется, II работает;

s_{22} — оба тестируются;

s_{23} — I тестируется, II восстанавливается;

s_{31} — I восстанавливается, II работает;

s_{32} — I восстанавливается, II тестируется;

s_{33} — оба восстанавливаются.

$$7. \ p(1) = \left(\frac{13}{20}, \frac{3}{10}, \frac{1}{20} \right)$$

$$8. \begin{matrix} B_k & B_c \\ A_k \begin{bmatrix} -2 & 3 \\ 2 & -3 \end{bmatrix} \\ A_c \end{matrix}$$

КОНТРОЛЬНЫЕ РАБОТЫ ДЛЯ СТУДЕНТОВ ЗАОЧНОЙ И ОЧНО-ЗАОЧНОЙ ФОРМ ОБУЧЕНИЯ

Контрольная работа № 1

Матрицы, определители, системы линейных уравнений

№ 1. Решить систему линейных уравнений $Ax = B$: а) методом Крамера; б) методом обратной матрицы (проверить правильность умножением); в) сделать проверку системы (см. подпараграф 5.2 главы 1, пример 1.16):

$$A = \begin{pmatrix} 1 & -1 & 1 \\ n+m & n & m \\ -n & -m & m \end{pmatrix}; \quad B = \begin{pmatrix} -n \\ m+1 \\ n-1 \end{pmatrix}$$

№ 2. Решить систему линейных уравнений $Cx = D$ методом Гаусса или Жордана-Гаусса (см. подпараграф 5.3 главы 1, пример 1.18, 1.19):

$$C = \begin{pmatrix} 0 & n & -m-n & 1 \\ -1 & n+m & -n & 1 \\ 1 & n & -m & 0 \\ -1 & m & m & 0 \end{pmatrix}; \quad D = \begin{pmatrix} n \\ m \\ m-1 \\ n+1 \end{pmatrix}.$$

а) найти ее ранг (см. § 4 главы 1, пример 1.13); б) сделать вывод о ее совместности; в) сделать вывод о ее определенности; г) сделать вывод о количестве решений. Найти решения, сделать проверку (см. подпараграф 5.4 главы 1).

Аналитическая геометрия

№ 3. Найти уравнение множества точек, равноудаленных от точек $P(m; n)$ и $Q(m + 4; n - 2)$ (см. подпараграф 1.1 главы 2, пример 2.1).

№ 4. Составить уравнение прямой, проходящей через точку $P(m; n)$: а) под углом 45 градусов к оси OX ; б) параллельно оси OY . Найти уравнение пучка прямых, проходящих через точку P (см. подпараграф 1.1 главы 2, пример 2.2).

№ 5. Составить уравнение прямой, проходящей через точки $P(-m; n)$ и $Q(-n; m)$ (см. подпараграф 1.2 главы 2, пример 2.3).

№ 6. Составить уравнения двух прямых, проходящих через точку $P(m; n + 1)$, одна из которых параллельна прямой $nx + my - mn = 0$, а другая перпендикулярна той же прямой (см. подпараграф 1.3 главы 2, пример 2.5).

№ 7. Найти координаты центра и радиус окружности $x^2 + y^2 - mx - ny - mn = 0$ (см. подпараграф 2.1 главы 2, пример 2.8.)

№ 8. Построить кривую $y = -mx^2 + nx - mn$ (см. подпараграф 2.2 главы 2, пример 2.11).

№ 9. Составьте уравнение плоскости Γ , которая проходит через точку P и перпендикулярна прямой PQ , если $P = (n; 1; m)$, $Q = (1; n; -m)$. Привести пример двух плоскостей, одна из которых параллельна Γ , другая — перпендикулярна ей.

Контрольная работа № 2

Пределы, непрерывность и разрывы функций

№ 1. Найти пределы функций:

а) $\lim_{x \rightarrow m} \frac{2x + n}{x - 1}$; б) $\lim_{x \rightarrow n} \frac{mx + 4}{x - n}$; в) $\lim_{x \rightarrow m} \frac{x^2 - (m + n)x + nm}{(x - m)^2}$;

г) $\lim_{x \rightarrow n/m} \frac{mnx^2 - (m^2 + n^2)x + mn}{\sqrt{2mx} - \sqrt{mx + n}}$; д) $\lim_{y \rightarrow 0} (1 - my)^{\frac{n}{y}}$;

е) $\lim_{x \rightarrow 0} \frac{\sin mx}{nx}$; ж) $\lim_{x \rightarrow \infty} \left(1 + \frac{m}{x}\right)^{nx}$; з) $\lim_{x \rightarrow 0} \frac{\sqrt{x^2 + n^2} - n}{\sqrt{x^2 + m^2} - m}$;

(см. § 9 главы 4, пример 4.14).

№ 2. Найти асимптоты кривой $y = \frac{x^2 - (m+n)x + nm}{x}$.

(см. § 10 главы 4).

Производные функций

№ 3. Найти производные следующих функций:

а) $y = \left(\frac{1}{n+1} \cdot x^{n+1} + {}^{m+1}\sqrt{x^4} + mn \right)^{m+n}$; б) $y = (\sqrt{x} + n)^m$;

в) $y = \ln \sqrt{\left(\frac{x^2 - m}{x^n + 1} \right)}$; г) $y = \frac{n}{x^3 + mx^2 + nx}$; д) $y = (\sin nx)^m$

(см. подпараграф 1.2.3 главы 4, пример 4.23).

№ 4. По формуле Лейбница найти производную порядка n функции $y = x^m \sin nx$.

(см. подпараграф 13.1 главы 4, теорема 4.24)

№ 5. Найти дифференциалы функций:

а) $y = x^n - (m-n)x + mn$; б) $y = \cos nx + m \sin nx$.

(см. подпараграф 12.4 главы 4)

Приложения производной

№ 6. Составить уравнения касательных к графику функции

$$y = \frac{mx + n}{mx - n},$$

параллельных прямой $2mx + ny + mn = 0$.

№ 7. Найти наибольшее и наименьшее значения функции $f(x) = 2x^3 + 3(m-n)x^2 - 6mx + 1$ на отрезке $[m-n; m+n]$.

№ 8. С помощью методов дифференциального исчисления построить график функции $y = \frac{m+x^2}{n-x^2}$ (см. подпараграф 14.5 главы 4, пример 4.29).

Неопределенный интеграл

№ 9. Найти интегралы:

а) $\int \left(mx^n - \frac{n}{m+1\sqrt{x^{n+1}}} + mn \right)^2 dx$; б) $\int \frac{dx}{x \cdot \sqrt{mx - nx^2}}$;

в) $\int (x+m)^2 \cdot e^{-nx} dx$; г) $\int \frac{dx}{x^3 - (m+n)x^2 + mnx}$; д) $\int nx \cos mx dx$.

(см. примеры в § 18, 19 главы 4)

Несобственные интегралы

№ 10. Вычислить интеграл или установить его расходимость:

$\int_n^{+\infty} \frac{dx}{(n^2 + x^2) \operatorname{arctg}(x/n)}$ (см. подпараграф 21.2 главы 4, пример 4.38).

Применения определенных интегралов

№ 11. Построить схематический чертеж и найти площадь фигуры, ограниченной линиями:

$y = x^2 + mx - n^2, (mn + n^2)x - (m + n)y + m^2n - n^3 = 0$.

(см. подпараграф 21.3 главы 4).

Приближенное вычисление определенных интегралов

№ 12. Для вычисления определенного интеграла

$J = \int_{m-3}^{m+3} \sqrt{x^2 + n} dx$, разбивая отрезок интегрирования на 10 равных

частей, найти приближенные значения J_{10} по формуле трапеций, по формуле Симпсона (см. подпараграф 21.4 главы 4).

Дифференциальные уравнения

№ 13. Выяснить, является ли решением дифференциального уравнения $y'' - my' + ny = 0$ функция $y = x \cdot e^x$.

№ 14. Найти общее решение уравнения: $y'' + 3y' + 2y = m \cos nx$ (см. подпараграф 22.4 главы 4, пример 4.52).

№ 15. Найти общее решение уравнения $y'' - 5y' + 6y = nx + m$.

Ряды

№ 16. Исследовать сходимость знакопеременного ряда. В случае сходимости исследовать на абсолютную и условную сходимость

$$\sum_{k=m+1}^{\infty} \frac{(-1)^{k-1} \cdot n}{nk - m}$$

(см. § 23 главы 4).

№ 17. Разложить функцию $f(x) = e^x$ в ряд по степеням $(x - t)$

(см. подпараграф 23.3 главы 4).

Контрольная работа № 3

1. Сколько трехзначных чисел можно составить из цифр 1, 2, ..., n , если каждая цифра входит в изображение числа только один раз?

2. Сколько можно составить сигналов из $2n + 2$ флажков различного цвета, взятых по 2?

3. Сколькими способами можно выбрать три детали из ящика, содержащего $n + 3$ детали? (см. подпараграф 1.3 главы 5.)

4. В портфеле инвестора, склонного к риску, $4n$ различных облигаций, среди которых $2 + n$ облигации с высокой степенью неисполнения обязательств по ним. Инвестор решает наугад исключить из портфеля n облигаций. Найти вероятность того, что он исключил только “бросовые” облигации. (см. пример 5.14.)

5. При перевозке ящика, в котором содержались $2n + 1$ стандартная и $2m$ нестандартных деталей, утеряна одна деталь, причем неизвестно какая. Наудачу извлеченная (после перевозки) из ящика деталь оказалась стандартной.

Найти вероятность того, что была утеряна: а) стандартная деталь; б) нестандартная деталь.

6. Отдел технического контроля обнаружил 3 нестандартных детали в партии из $3m$ случайно отобранных деталей. Найти относительную частоту появления нестандартных деталей. (см. пример 5.16.)

7. В портфеле инвестора $t + 2$ различных акций, из которых две — с высокой степенью исполнения обязательств по ним, а остальные — с высокой степенью неисполнения обязательств по ним (“бросо-

вые”). Наудачу выбрано 2 акции. Найти вероятность того, что обе акции “бросовые”. (см. пример 5.20.)

8. Вероятность того, что торговец акциями, торгующий во внутри — дневном режиме (day-Trader), при однократной покупке акций компании D закроет свою позицию (продаст акции) с прибылью, равна $0, n$. Сколько сделок должен провести day-Trader, чтобы с вероятностью не менее $0, m$ он получил прибыль хотя бы от одной сделки? (см. пример 5.26.)

9. В урне n белых и m черных шаров. Из урны дважды вынимают по одному шару, не возвращая их обратно. Найти вероятность появления белого шара при втором испытании (событие B) если при первом испытании был извлечен черный шар (событие A).

10. В денежной лотерее выпущено $10m$ билетов. Разыгрывается один выигрыш 100 рублей и 10 выигрышей по десять рублей. Найти закон распределения случайной величины X — стоимости возможного выигрыша для владельца одного лотерейного билета. (см. пример 5.29.)

11. Дискретная случайная величина X задана законом распределения:

X	1	$m + 1$	$m + 3$	$m + 5$
p	0,2	0,1	0,4	0,3

Построить многоугольник распределения. (см. пример 5.30.)

12. Найти математическое ожидание случайной величины X , зная ее закон распределения (см. пример 5.35):

X	$1 + n$	$4 + n$	$7 + n$
p	0,1	0,6	0,3

13. Независимые случайные величины X и Y заданы следующими законами распределения:

X	$5 + n$	$2 + n$	$4 + n$	Y	$7 + m$	$9 + m$
p	0,6	0,1	0,3	p	0,8	0,2

Найти математическое ожидание случайной величины XY . (см. пример 5.37.)

14. Найти дисперсию случайной величины X , которая задана следующим законом распределения (см. пример 5.40):

X	$1 + n$	$3 + n$	$7 + n$
p	0,3	0,5	0,2

15. Найти дисперсию случайной величины X , которая задана следующим законом распределения:

X	$1 + m$	$4 + m$	$6 + m$
p	0,1	0,6	0,3

16. Случайные величины X и Y независимы. Найти дисперсию случайной величины $Z = 3X + 2Y$, если известно, что $D(X) = n$, $D(Y) = m$.

17. Дискретная случайная величина задана таблицей распределения:

X	$1 + n$	$5 + n$	$9 + n$
p	0,3	0,1	0,6

Найти функцию распределения и вычертить ее график.

18. Математическое ожидание нормально распределенной случайной величины X равно $a = n$ и среднее квадратическое отклонение $\sigma = m$. Написать плотность вероятности X . (см. § 5 главы 5.)

19. Написать плотность вероятности нормально распределенной случайной величины X , зная, что $M(X) = n$, $D(X) = m$.

20. Стандартное отклонение доходности акций компании A за период t составляет $10n\%$? Акции компании B за тот же период t — $10m\%$. Определить риск портфеля, если инвестор купил акций компании A на $10\,000m$ руб., B — на $40\,000m$ руб. (см. пример 5.45.)

Рекомендации к выполнению контрольной работы

Контрольная работа выполняется в отдельной тетради в клетку 18 листов аккуратным почерком. Задания переписываются, в них указывается персональная комбинация параметров m и n для каждого студента.

От значений A и B (см. табл. 1, 2) и их комбинации зависит вариант индивидуального задания студента. A — сумма цифр номера в журнале посещаемости, B — сумма цифр номера зачетной книжки. Если полученное число не лежит в интервале $[1; 10]$, то цифры снова складываем. Например, ваш номер в журнале № 27, значит $2 + 7 = 9$, снова складываем: $1 + 1 = 2$, значит, $A = 2$.

Таблица 1

Выбор параметра m

A	10	1	2	3	4	5	6	7	8	9
m	4	3	5	6	3	7	4	9	8	5

Таблица 2

Выбор параметра n

B	10	1	2	3	4	5	6	7	8	9
n	3	6	7	4	5	3	8	5	9	4

В скобках для помощи студентам указан параграф с теоретическим материалом и/или номера аналогичных решенных примеров из учебника.

Итоговую аттестацию (экзаменационную оценку) предлагаем проставлять в соответствии с контролем работы студентов по четырем видам учебной деятельности:

Таблица 3

Вид учебной деятельности	Показатель его значимости для итоговой аттестационной оценки, %
Посещение аудиторных занятий	20
Активность на занятиях	30
Выполнение контрольной работы	20
Итоговый контроль знаний	30

Предлагается проводить итоговый контроль знаний в день экзамена в локальной компьютерной сети института с использованием программы *OprosSystem_mat*, входящей в состав аппаратно-программного комплекса по системе дистанционного обучения и контроля знаний студентов.

По желанию студенты смогут пройти итоговый контроль знаний дистанционно (при условии наличия у них персонального установочного диска) и выслать результат самостоятельной работы преподавателю по Интернету в указанные сроки (предварительно до 25.05.XX или до 25.12.XX).

Итоговая аттестация (экзаменационная оценка) будет представляться следующим образом. Если, например, студент оценивается по указанным выше (см. табл. 3) видам учебной деятельности показателями 2, 2, 5, 4 соответственно, то его итоговая аттестация будет: $0,2 \cdot 2 + 0,3 \cdot 2 + 0,2 \cdot 5 + 0,3 \cdot 4 = 0,4 + 0,6 + 1 + 1,2 = 3,2$.

Для пересчета средней оценки в дискретных “отлично”, “хорошо”, “удовлетворительно” вводится условное соответствие.

Например, при средней оценке от 3 до 3,5 итоговая оценка “удовлетворительно”, от 3,5 до 4,5 — “хорошо”, более 4,5 — “отлично”.

Задать свои вопросы по поводу учебного процесса можно по электронной почте по адресу: **u10228@diapodolsk.ru**

Информацию о лабораторном практикуме *OproSystem_mat* можно найти на сайте:

www.matecon.com

ЛИТЕРАТУРА

1. *Беклемишев Д.В.* Курс аналитической геометрии и линейной алгебры: Учебник для студентов вузов. — М.: Физматлит, 1998.
2. *Бережная Е.В., Бережной В.И.* Математические методы моделирования экономических систем: Учеб. пособие. — М.: Финансы и статистика, 2003.
3. *Виленкин И.В., Гробер В.М.* Высшая математика для студентов экономических, технических, естественно-научных специальностей вузов. — Ростов-н / Д.: Феникс, 2002.
4. *Гмурман В.Е.* Теория вероятностей и математическая статистика. — М.: Высшая школа, 2005.
5. *Горелова Г.В., Кацко И.А.* Теория вероятностей и математическая статистика: Учеб. пособие для студентов вузов, обучающихся по экономическим специальностям. — Ростов-н / Д.: Феникс, 2002.
6. Задачи и упражнения по математическому анализу / Под ред. *В.П. Демидовича*. — М.: Астрель—АСТ, 2003.
7. *Замков О. О., Толстомятенко А.В., Черемных Ю.И.* Математические методы в экономике. — М.: Дело и Сервис, 2001.
8. *Колемаев В.А.* Математическая экономика: Учебник для вузов. — М.: ЮНИТИ, 1998.
9. *Краасс М.С.* Математика для экономистов: Учебник для вузов. — М.: Питер, 2006.
10. *Кремер Н.Ш.* Высшая математика для экономистов: Учебник для вузов. — М.: ЮНИТИ, 2007.
11. *Кремер Н.Ш.* Теория вероятностей и математическая статистика. — М.: ИНФРА-М, 2005.
12. *Кундышева Е.С.* Математическое моделирование в экономике: Учебное пособие / Под научн. ред. проф. *Б.А. Сулакова*. — М.: ИТК «Дашков и К°», 2007.

13. *Кундышева Е.С.* Экономико-математическое моделирование: Учебник для вузов / Под научн. ред. проф. *Б.А. Сулакова*. — ИТК «Дашков и К°», 2007.
14. *Лебедев В.В.* Математика в экономике и управлении: Учеб. пособие по курсу “Высшая математика”. — М.: НВТ-Дизайн, 2004.
15. *Лебедев В.В., Лебедев К.В.* Математическое и компьютерное моделирование экономики: Учеб. пособие. — М.: НВТ-Дизайн, 2002.
16. Общий курс высшей математики для экономистов: Учебник / Под ред. *В.И. Ермакова*. — М.: ИНФРА-М, 2006.
17. Сборник задач по высшей математике для экономистов / Под ред. *В.И. Ермакова*. — М.: ИНФРА-М, 2007.
18. *Солодовников А.С., Бабайцев В.А., Браилов А.В., Шандра И.Г.* Математика в экономике: Учебник для студентов экономических специальностей вузов. — М.: Финансы и статистика, 2006.
19. *Четыркин Е.Н.* Финансовая математика: Учебник для студентов вузов. — М.: Акад. нар. хоз.-ва при Правительстве РФ: Дело, 2002.

ОГЛАВЛЕНИЕ

Введение	3
Глава 1. ВВЕДЕНИЕ В ЛИНЕЙНУЮ АЛГЕБРУ	6
§ 1. Векторные пространства и матрицы	6
1.1. Основные сведения о векторах и векторных пространствах	6
1.2. Основные сведения о матрицах	14
1.3. Операции над матрицами	18
1.4. Собственные значения и собственные векторы матрицы	24
§ 2. Определители квадратных матриц	26
2.1. Основные определения	27
2.2. Свойства определителей	32
§ 3. Обратная матрица	39
§ 4. Ранг матрицы	42
§ 5. Методы решения систем линейных алгебраических уравнений	55
5.1. Основные понятия и определения	55
5.2. Решение системы n линейных уравнений с n неизвестными. Теорема Крамера и метод обратной матрицы	57
5.3. Методы Гаусса и Жордана-Гаусса	62
5.4. Решение систем общего вида	68
5.5. Разложение вектора по базису	72
5.6. Приведенная система. Фундаментальная система решений	76
5.7. Характеристическое уравнение	79
§ 6. Модель Леонтьева межотраслевого баланса	82
<i>Контрольные вопросы</i>	84
<i>Упражнения</i>	85
<i>Ответы к упражнениям</i>	88

Глава 2. ОСНОВЫ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ..... 89

§ 1. Прямые линии и плоскости	89
1.1. Основные понятия об уравнениях	89
1.2. Уравнение прямой на плоскости	92
1.3. Взаимное расположение прямых на плоскости. Расстояние от точки до прямой	98
1.4. Уравнение плоскости	103
1.5. Уравнение прямой линии в пространстве	107
1.6. Взаимное расположение прямой и плоскости в пространстве	109
§ 2. Кривые второго порядка	110
2.1. Эллипс и окружность	110
2.2. Гипербола и парабола	113
2.3. Примеры решения задач	119
2.4. Кривые спроса и предложения. Точка равновесия	124
<i>Контрольные вопросы</i>	127
<i>Упражнения</i>	127
<i>Ответы к упражнениям</i>	128

Глава 3. ОСНОВЫ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ 129

§ 1. Постановка задачи линейного программирования в экономике	129
§ 2. Графический метод решения задачи линейного программирования	134
§ 3. Примеры построения экономических моделей	139
§ 4. Задача линейного программирования с ограничениями-неравенствами. Переход от нее к основной задаче линейного программирования и наоборот	150
§ 5. Симплекс-метод решения задачи линейного программирования	156
§ 6. Отыскание опорного решения основной задачи линейного программирования	168

§ 7. Отыскание оптимального решения основной задачи линейного программирования	176
§ 8. Транспортная задача линейного программирования	185
§ 9. Нахождение опорного плана	191
§ 10. Улучшение плана перевозок. Цикл пересчета	197
§ 11. Двойственность в линейном программировании ...	206
11.1. Понятие двойственности	206
11.2. Несимметричные двойственные задачи	208
11.3. Симметричные двойственные задачи	210
11.4. Виды математических моделей	211
<i>Контрольные вопросы</i>	213
<i>Упражнения</i>	213

Глава 4. ОСНОВЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА

§ 1. Множества и операции над ними	215
1.1. Основные определения	215
1.2. Арифметическое m -мерное пространство	218
§ 2. Бюджетные и производственные множества	220
2.1. Определение бюджетного множества	220
2.2. Производственные множества и их свойства ...	222
§ 3. Числовые последовательности	224
3.1. Числовые последовательности и операции над ними	224
3.2. Ограниченные и неограниченные последовательности	225
3.3. Предел последовательности. Бесконечно большие и бесконечно малые последовательности	226
3.4. Основные свойства бесконечно малых последовательностей	227
§ 4. Сходящиеся последовательности	228
4.1. Понятие сходящейся последовательности	228
4.2. Основные свойства сходящихся последовательностей	229

§ 5. Монотонные последовательности	232
5.1. Определение монотонных последовательностей	232
5.2. Признак сходимости монотонной последовательности. Формула сложных процентов	233
§ 6. Понятие функции	238
6.1. Определение функциональной зависимости. Способы задания функции	238
6.2. Некоторые свойства функций	240
6.3. Производственная функция и функция полезности	241
§ 7. Предел функции	245
7.1. Предел функции в бесконечности	245
7.2. Предел функции в точке	246
7.3. Основные свойства предела функции	247
§ 8. Два замечательных предела	248
8.1. Первый замечательный предел	248
8.2. Второй замечательный предел	249
§ 9. Бесконечно малые и бесконечно большие функции	251
§ 10. Непрерывность функции	259
10.1. Определение непрерывной функции. Точки разрыва. Асимптоты	259
10.2. Обратная функция	265
§ 11. Комплексные числа. Функции комплексного переменного	267
11.1. Основные определения	267
11.2. Геометрическая интерпретация	269
11.3. Тригонометрическая форма комплексного числа	270
11.4. Свойства показательной функции	271
11.5. Степени и корни	271
§ 12. Производная и дифференциал функции	274
12.1. Определение производной	274
12.2. Основные правила дифференцирования	276

12.3. Производная сложной и обратной функций	280
12.4. Дифференциал функции	281
12.5. Геометрический смысл дифференциала функции	282
12.6. Производные комплекснозначных функций действительного аргумента	283
§ 13. Некоторые приложения производной	284
13.1. Производные высших порядков	284
13.2. Дифференциальные теоремы о среднем	286
13.3. Формула Тейлора для многочлена	290
13.4. Разложение произвольной функции	291
§ 14. Построение графиков функций	294
14.1. Правило Лопиталя	294
14.2. Возрастание и убывание функций	295
14.3. Экстремум функции	296
14.4. Выпуклость функции. Точки перегиба	298
14.5. Общая схема исследования функций и построение их графиков	300
14.6. Теорема о пересечении предельных и средних издержек	302
§ 15. Модели экономического взаимодействия на простейших рынках	305
15.1. Паутинообразная модель рынка	305
15.2. Модель Солоу экономического роста	310
§ 16. Понятие функции нескольких переменных. Модель фирмы	313
16.1. Функция m переменных и область ее определения	313
16.2. Предел функции нескольких переменных ...	314
§ 17. Дифференцирование функций нескольких переменных. Производная по направлению градиента	315
17.1. Частные производные	315
17.2. Полное приращение функции	315
17.3. Полный дифференциал	316
17.4. Поверхности уровня	318

17.5. Производная по направлению	319
17.6. Градиент	321
17.7. Производственная функция Кобба-Дугласа	323
17.8. Теория фирмы	324
§ 18. Неопределенный интеграл	328
§ 19. Некоторые свойства неопределенного интеграла	331
19.1. Основные правила вычисления неопределенного интеграла	331
19.2. Интегрирование методом замены переменной	332
19.3. Интегрирование по частям	333
19.4. Метод неопределенных коэффициентов	334
§ 20. Определенный интеграл	336
20.1. Определение определенного интеграла	336
20.2. Вычисление определенного интеграла. Формула Ньютона-Лейбница	338
§ 21. Некоторые приложения интегрального исчисления	341
21.1. Классическая модель Вильсона управления запасами	341
21.2. Несобственные интегралы	343
21.3. Примеры вычисления площадей плоских фигур	346
21.4. Приближенное вычисление определенных интегралов	348
§ 22. Дифференциальные уравнения	352
22.1. Понятие о дифференциальном уравнении и его решении	352
22.2. Линейные уравнения первого порядка	357
22.3. Уравнение Бернулли	358
22.4. Дифференциальные уравнения второго порядка	360
§ 23. Числовые ряды	361
23.1. Понятие ряда. Признаки сходимости рядов	361

23.2. Знакопеременные ряды.	
Их признаки сходимости	376
23.3. Степенной ряд и область его сходимости ...	379
23.4. Современная стоимость денег,	
дисконтирование	384
<i>Контрольные вопросы</i>	391
<i>Упражнения</i>	391
<i>Ответы к упражнениям</i>	395

Глава 5. ВВЕДЕНИЕ В ТЕОРИЮ ВЕРОЯТНОСТЕЙ	
И МАТЕМАТИЧЕСКУЮ СТАТИСТИКУ	398
§ 1. Основные понятия теории вероятностей	398
1.1. Опыт и события	398
1.2. Определение вероятности события	402
1.3. Комбинаторика	404
1.4. Примеры вычисления вероятностей	407
1.5. Относительная частота	411
§ 2. Основные теоремы теории вероятностей	413
2.1. Теорема сложения вероятностей	413
2.2. Теорема умножения вероятностей	415
2.3. Вероятности гипотез	
и формулы полной вероятности	423
§ 3. Случайные величины	427
3.1. Определение случайной величины	427
3.2. Определение функции распределения	427
3.3. Дискретные и непрерывные	
случайные величины	428
3.4. Свойства функции распределения	
вероятностей случайной величины	429
3.5. Закон распределения дискретной	
случайной величины	430
3.6. Плотность распределения вероятностей	
непрерывной случайной величины	434
§ 4. Числовые характеристики случайных величин	438
4.1. Числовые характеристики дискретных	
случайных величин	438

4.2. Числовые характеристики непрерывных случайных величин	447
§ 5. Некоторые специальные распределения случайных величин	449
§ 6. Некоторые аспекты из области математической статистики	460
6.1. Основные задачи математической статистики. Измерение риска, присущего инвестиционному портфелю	460
6.2. Оценивание параметров	474
6.3. Проверка гипотез	480
§ 7. Математическая модель прогнозирования экономических процессов	483
7.1. Экономическое прогнозирование	483
7.2. Методы прогнозирования	485
7.3. Статистический анализ временных рядов	487
7.4. Кривые роста	487
<i>Контрольные вопросы</i>	495
<i>Упражнения</i>	496
<i>Ответы к упражнениям</i>	500

Глава 6. ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ

МЕТОДЫ	502
§ 1. Оптимизационные задачи на графах	502
1.1. Основные понятия теории графов	502
1.2. Метод ветвей и границ. Задача коммивояжера	508
§ 2. Моделирование экономических процессов по схеме марковских случайных процессов	516
2.1. Марковский случайный процесс с дискретными состояниями	516
2.2. Случайные процессы с дискретным и непрерывным временем. Марковская цепь ...	521
§ 3. Некоторые аспекты теории игр	528
3.1. Основные понятия математической теории игр	528

3.2. Решение матричных антагонистических игр ...	531
3.3. Кооперативные и некооперативные игры	534
3.4. Связь теории игр с анализом проблем	
микроэкономики	536
<i>Контрольные вопросы</i>	539
<i>Упражнения</i>	539
<i>Ответы к упражнениям</i>	541
Контрольные работы для студентов заочной	
и очно-заочной форм обучения	543
ЛИТЕРАТУРА	552