РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ

В.Л. Клюшин

ВЫСШАЯ МАТЕМАТИКА ДЛЯ ЭКОНОМИСТОВ

Учебное пособие

Допущено Министерством образования и науки
Российской Федерации в качестве учебного пособия
для студентов высших учебных заведений,
обучающихся по экономическим специальностям

МОСКВА
ИНФРА-М
2009

УДК 51(075.8)
ББК 22.1я73
К52

Рецензенты:

кафедра высшей и прикладной математики
Академии труда и социальных отношений
(зав. кафедрой д-р физ.-мат. наук, проф. *Л.Б. Шапиро*);
д-р экон. наук, проф. *В.И. Данилин*

К52 **Клюшин В. Л.** Высшая математика для экономистов: Учеб. пособие. — М.: ИНФРА-М, 2009. — 448 с. — (Учебники РУДН).

ISBN 5-16-002752-1

Изложены основы линейной алгебры, математического анализа, теории дифференциальных уравнений, теории рядов. Основные теоретические положения учебного материала сопровождаются решением задач. Везде, где это возможно, раскрывается экономический смысл математических понятий, рассматриваются экономические приложения и простейшие модели. В основу книги положены лекции, которые автор читает на протяжении многих лет.

Для студентов экономических факультетов вузов, экономистов-практиков, а также лиц, занимающихся самообразованием.

ББК 22.1я73

ISBN 5-16-002752-1

ПРЕДИСЛОВИЕ

Эта книга представляет собой учебное пособие по высшей математике для студентов экономических специальностей и написана в соответствии с требованиями государственных образовательных стандартов в области математики для специалистов с высшим образованием по экономическим специальностям.

Пособие соответствует программе по высшей математике для экономистов, рассчитанной приблизительно на 200 часов аудиторных занятий. Оно предназначено студентам-первокурсникам, а также студентам, изучающим общий курс высшей математики на протяжении первых трех семестров обучения в вузе. В книге представлены элементы линейной алгебры и аналитической геометрии, математический анализ, основы теории дифференциальных уравнений и рядов. Следует заметить, что это только фундаментальная часть математики для экономистов, существует еще прикладная часть, не вошедшая в эту книгу: «Теория вероятностей и математическая статистика», «Оптимизация и основы теории принятия решений».

Излагаемый теоретический материал сопровождается, где это возможно, экономическими приложениями, дается экономический смысл математических понятий, приводятся математические формулировки некоторых экономических законов. Некоторые приложения выделены в отдельные главы. Все экономические приложения несложны, поскольку студенты первого курса еще не располагают серьезными экономическими знаниями.

Автор надеется, что книга будет полезна также студентам-заочникам и экономистам-практикам.

Раздел I
ЭЛЕМЕНТЫ ЛИНЕЙНОЙ АЛГЕБРЫ

Глава 1
ВЕКТОРЫ И ДЕЙСТВИЯ С НИМИ. ЛИНЕЙНЫЕ ПРОСТРАНСТВА

1.1. ЛИНЕЙНЫЕ ОПЕРАЦИИ НАД ВЕКТОРАМИ

Из школьного курса геометрии известно, что если на плоскости задана прямоугольная система координат, то каждый вектор $\overline{a}$ характеризуется своими координатами — парой чисел a_1, a_2: $\overline{a}=(a_1,\ a_2)$. В трехмерном пространстве вектору соответствует тройка координат: $\overline{a}=(a_1,\ a_2,\ a_3)$.

Определение. Любой упорядоченный набор n действительных чисел $(a_1,\ a_2,\ ...,\ a_n)$ называется ***n*-мерным вектором** $\overline{a}$:

$$\overline{a}=(a_1,\ a_2,\ \ldots,\ a_n). \tag{1.1}$$

Числа a_1, a_2, ..., a_n называются **координатами**, или **компонентами**, вектора $\overline{a}$. Например, $\overline{a}$ = (4, 2, 5, 0, –2) — пятимерный вектор; в частности, его третья компонента равна 5, а пятая равна –2.

Заметим, что координаты вектора $\overline{a}$ можно расположить либо в строку [см. (1.1)], либо в столбец:

$$\overline{a}=\begin{pmatrix} a_1 \\ a_2 \\ \ldots \\ a_n \end{pmatrix}. \tag{1.2}$$

Запись вида (1.1) называется вектором-строкой, а (1.2) — вектором-столбцом.

Число координат вектора называют **размерностью вектора**.

Два n-мерных вектора $\overline{a} = (a_1, a_2, \ldots, a_n)$ и $\overline{b} = (b_1, b_2, \ldots, b_n)$ называются **равными**, если их соответствующие координаты равны: $a_1 = b_1, a_2 = b_2, ..., a_n = b_n$. В этом случае пишем $\overline{a} = \overline{b}$.

Суммой двух *n*-мерных векторов $\overline{a} = (a_1, a_2, \ldots, a_n)$ и $\overline{b} = (b_1, b_2, \ldots, b_n)$ называется вектор

$$\overline{a} + \overline{b} = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

Вектор, все координаты которого равны нулю, называется **нулевым**:

$$\overline{0} = (0, 0, \ldots, 0).$$

Вектор $(-a_1, -a_2, ..., -a_n)$ называется **противоположным** вектору $\overline{a} = (a_1, a_2, \ldots, a_n)$ и обозначается $-\overline{a}$:

$$-\overline{a} = (-a_1, -a_2, ..., -a_n).$$

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

Произведением вектора $\overline{a} = (a_1, a_2, ..., a_n)$ **на число** k называется вектор

$$k\overline{a} = (ka_1, ka_2, \ldots, ka_n).$$

Сложение векторов и умножение вектора на число называются **линейными операциями**.

Отметим следующие **свойства линейных операций** (их легко проверить):

1. $\overline{a} + \overline{b} = \overline{b} + \overline{a}$.

2. $(\overline{a} + \overline{b}) + \overline{c} = \overline{a} + (\overline{b} + \overline{c})$.

3. $\overline{a} + \overline{0} = \overline{a}$.

4. $\overline{a} + (-\overline{a}) = \overline{0}$.

5. $k(\overline{a} + \overline{b}) = k\overline{a} + k\overline{b}$.

6. $(k_1 + k_2)\overline{a} = k_1\overline{a} + k_2\overline{a}$.

7. $k_1(k_2\overline{a}) = (k_1k_2)\overline{a}$.

8. $1 \cdot \overline{a} = \overline{a}$.

Определение. Множество всех n-мерных векторов, в котором определены операции сложения векторов и умножения вектора на число, называется ***n*-мерным векторным пространством** и обозначается $\mathbf{R}^n$.

Пространство $\mathbf{R}^n$ называется также **линейным пространством**.

1.2. СКАЛЯРНОЕ ПРОИЗВЕДЕНИЕ ВЕКТОРОВ

Скалярным произведением двух векторов $\overline{a}=(a_1,\ a_2,\ \ldots,\ a_n)$ и $\overline{b}=(b_1,\ b_2,\ \ldots,\ b_n)$ называется число

$$\left(\overline{a},\overline{b}\right)=a_1b_1+a_2b_2+\ldots+a_nb_n. \qquad (1.3)$$

Проиллюстрируем скалярное произведение простыми примерами.

Пример **1.1.** Хозяйка покупает 0,5 кг хлеба, 5 кг картофеля, 3 кг огурцов, 2 кг помидоров и 1,5 кг мяса по ценам соответственно 12, 11, 15, 30, 80 руб. за килограмм. Если рассмотреть вектор товаров $\overline{a}$ = (0,5; 5; 3; 2; 1,5) и вектор цен $\overline{b}$ = (12; 11; 15; 30; 80), то сумма денег, затраченных на эту покупку, выражается скалярным произведением:

$$\left(\overline{a},\overline{b}\right)=0{,}5\cdot 12+5\cdot 11+3\cdot 15+2\cdot 30+1{,}5\cdot 80=286\text{ руб.}$$

Пример **1.2.** Сумма в 300 000 руб. помещается под проценты на год в четыре банка: 50 000 – под 12%, 50 000 – под 15%, 100 000 – под 10% и 100 000 – под 20%.

Здесь вектор вкладов $\overline{a}$ = (50 000, 50 000, 100 000, 100 000), вектор процентных ставок $\overline{b}$ = (0,12; 0,15; 0,10; 0,20).

Первоначальная сумма возрастет на величину, выражаемую скалярным произведением:

$$\left(\overline{a},\overline{b}\right)=50\,000\cdot 0{,}12+50\,000\cdot 0{,}15+100\,000\cdot 0{,}10+$$
$$+\,100\,000\cdot 0{,}20=43\,500\text{ руб.}$$

Перечислим **основные свойства скалярного произведения**:

1. $\left(\overline{a},\overline{b}\right)=\left(\overline{b},\overline{a}\right)$.

2. $\left(k\overline{a},\overline{b}\right)=k\left(\overline{a},\overline{b}\right)$.

3. $\left(\overline{a},\overline{b}+\overline{c}\right)=\left(\overline{a},\overline{b}\right)+\left(\overline{a},\overline{c}\right)$.

4. $(\overline{a},\ \overline{a})\geq 0$; при этом $(\overline{a},\ \overline{a})=0$ тогда и только тогда, когда $\overline{a}$ – нулевой вектор.

1.3. ЛИНЕЙНАЯ ЗАВИСИМОСТЬ ВЕКТОРОВ

Определение. Вектор $\overline{a}$ называется **линейной комбинацией** векторов $\overline{a}_1,\overline{a}_2,\ \ldots,\ \overline{a}_s$ из $\mathbf{R}^n$, если

$$\overline{a}=\lambda_1\overline{a}_1+\lambda_2\overline{a}_2+\ldots+\lambda_s\overline{a}_s,$$

где λ_1, λ_2, ..., λ_s – какие угодно действительные числа. В этом случае говорят, что вектор $\overline{a}$ *линейно выражается* через векторы $\overline{a}_1,\overline{a}_2,\ \ldots,\ \overline{a}_s$.

Пусть, например, $\overline{a}_1 = (2, 1, 3, -2)$, $\overline{a}_2 = (3, 1, 2, 2)$, $\overline{a}_3 = (2, 1, 2, 0)$, тогда

$$3\overline{a}_1 + 2\overline{a}_2 - 4\overline{a}_3 = (6, 3, 9, -6) + (6, 2, 4, 4) - (8, 4, 8, 0) = \\ = (4, 1, 5, -2).$$

Вектор $\overline{a} = (4, 1, 5, -2)$ есть линейная комбинация векторов $\overline{a}_1, \overline{a}_2, \overline{a}_3$:

$$\overline{a} = 3\overline{a}_1 + 2\overline{a}_2 - 4\overline{a}_3.$$

Всякий набор векторов из $\mathbf{R}^n$ будем называть **системой векторов**. В рассмотренном примере система состоит из четырех векторов: $\overline{a}_1$, $\overline{a}_2$, $\overline{a}_3$ и $\overline{a}$. При этом один из них (вектор $\overline{a}$) есть линейная комбинация остальных векторов этой системы.

Определение. Система векторов $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_m$ называется **линейно зависимой**, если существуют такие числа $\lambda_1, \lambda_2, \ldots, \lambda_m$, не равные одновременно нулю, что

$$\lambda_1\overline{a}_1 + \lambda_2\overline{a}_2 + \ldots + \lambda_m\overline{a}_m = \overline{0}. \tag{1.4}$$

В противном случае векторы $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_m$ называются **линейно независимыми**. Иначе говоря, векторы $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_m$ линейно независимы, если из равенства (1.4) следует $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$.

Докажем, что *система, содержащая более одного вектора $\overline{a}_1$, $\overline{a}_2, \ldots, \overline{a}_m$, линейно зависима тогда и только тогда, когда хотя бы один из них есть линейная комбинация остальных.*

1. Пусть выполняется равенство (1.4) и хотя бы один из коэффициентов отличен от нуля, например $\lambda_m \neq 0$. Тогда

$$\overline{a}_m = -\frac{\lambda_1}{\lambda_m}\overline{a}_1 - \frac{\lambda_2}{\lambda_m}\overline{a}_2 - \ldots,$$

т.е. вектор $\overline{a}_m$ есть линейная комбинация остальных векторов.

2. Пусть теперь один из векторов, например $\overline{a}_2$, есть линейная комбинация остальных:

$$\overline{a}_2 = \lambda_1\overline{a}_1 + \ldots + \lambda_m\overline{a}_m.$$

Тогда

$$\lambda_1\overline{a}_1 + (-1)\cdot\overline{a}_2 + \ldots + \lambda_m\overline{a}_m = \overline{0},$$

и в этом равенстве имеется коэффициент, отличный от нуля ($\lambda_2 = -1$). Следовательно, система векторов $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_m$ линейно зависима.

Отметим некоторые **свойства векторов пространства $\mathbf{R}^n$** .

1. *Если система $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_m$ содержит нулевой вектор, то она линейно зависима.*

Чтобы убедиться в этом, достаточно в левой части равенства (1.4) взять нулевой вектор с коэффициентом 1, а остальные коэффициенты равными нулю.

2. *Если часть векторов системы $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_m$ линейно зависимы**, *то и все эти векторы линейно зависимы.*

В самом деле, пусть, например, в системе $\overline{a}_1, \overline{a}_2, \overline{a}_3, \overline{a}_4, \overline{a}_5$ векторы $\overline{a}_2, \overline{a}_3, \overline{a}_5$ линейно зависимы: $\lambda_2\overline{a}_2 + \lambda_3\overline{a}_3 + \lambda_5\overline{a}_5 = \overline{0}$ и, например, $\lambda_3 \neq 0$. Тогда

$$0 \cdot \overline{a}_1 + \lambda_2\overline{a}_2 + \lambda_3\overline{a}_3 + 0 \cdot \overline{a}_4 + \lambda_5\overline{a}_5 = \overline{0},$$

и в этом равенстве по меньшей мере один коэффициент отличен от нуля ($\lambda_3 \neq 0$). Следовательно, эта система $\overline{a}_1, \overline{a}_2, \overline{a}_3, \overline{a}_4, \overline{a}_5$ линейно зависима.

Приведем без доказательства следующую важную теорему.

***Теорема* 1.1.** Любая система, содержащая $(n + 1)$ векторов пространства $\mathbf{R}^n$, линейно зависима.

1.4. БАЗИС И РАНГ СИСТЕМЫ ВЕКТОРОВ

Дана система векторов

$$\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_k. \qquad (1.5)$$

Всякая подсистема системы (1.5) называется **базисом** этой системы, если выполняются следующие условия:

1) эта подсистема линейно независима;

2) любой вектор системы (1.5) линейно выражается через векторы этой подсистемы.

Система векторов может иметь несколько базисов. Можно показать, что *все базисы данной системы векторов состоят из одного и того же числа векторов*. Это число (число векторов базиса) называется **рангом** системы.

Пространство $\mathbf{R}^n$, очевидно, является системой, содержащей все n-мерные векторы. Понятие базиса распространяется и на $\mathbf{R}^n$.

Определение. Система векторов называется **базисом пространства $\mathbf{R}^n$**, если:

1) эта система линейно независима;

* Иначе говоря, данная система векторов содержит линейно зависимую подсистему.

2) всякий вектор пространства $\mathbf{R}^n$ линейно выражается через векторы данной системы.

Примером базиса в $\mathbf{R}^n$ является система, состоящая из n «единичных» векторов:

$$\bar{e}_1 = (1, 0, \ldots, 0),$$

$$\bar{e}_2 = (0, 1, \ldots, 0),$$

$$\ldots\ldots\ldots$$

$$\bar{e}_n = (0, 0, \ldots, 1).$$

Действительно, с одной стороны, эта система линейно независима (так как из $\lambda_1\bar{e}_1 + \lambda_2\bar{e}_2 + \ldots + \lambda_n\bar{e}_n = \bar{0}$ следует $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$), с другой стороны, всякий вектор $\bar{a} = (a_1, a_2, \ldots, a_n)$ представим в виде

$$\bar{a} = a_1\bar{e}_1 + a_2\bar{e}_2 + \ldots + a_n\bar{e}_n,$$

т.е. является линейной комбинацией векторов $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$.

В предыдущем примере базис состоял из n векторов. Это не случайно. Справедлива следующая теорема (приведем ее без доказательства).

Теорема **1.2.** Линейно независимая система векторов в $\mathbf{R}^n$ является базисом тогда и только тогда, когда число этих векторов равно n.

Число векторов базиса пространства, т.е. максимальное число его линейно независимых векторов, называется **размерностью пространства**. Пространство $\mathbf{R}^n$, ранее названное n-мерным исходя из других соображений, является также n-мерным в том смысле, что его размерность равна n.

1.5. РАЗЛОЖЕНИЕ ВЕКТОРА ПО БАЗИСУ

Пусть система векторов

$$\bar{a}_1 = (a_{11}, a_{12}, \ldots, a_{1n}),\ \bar{a}_2 = (a_{21}, a_{22}, \ldots, a_{2n}),\ \ldots,$$
$$\bar{a}_m = (a_{m1}, a_{m2}, \ldots, a_{mn}) \qquad (1.6)$$

является базисом* и вектор $\bar{x}$ разложен по векторам (1.6):

$$\bar{x} = x_1\bar{a}_1 + x_2\bar{a}_2 + \ldots + x_m\bar{a}_m. \qquad (1.7)$$

---

* Здесь компоненты векторов приходится снабжать двойными индексами: первый – указывает на номер вектора, второй – на номер компоненты.

Возникает вопрос: однозначно ли определяются коэффициенты $x_1, x_2, \ldots, x_m$ разложения (1.7)?

***Теорема* 1.3.** Разложение вектора $\bar{x}$ по векторам базиса единственно.

Доказательство. Предположим, что вектор $\bar{x}$ представлен в виде линейной комбинации векторов (1.6) двумя способами:

$$\bar{x} = x_1\bar{a}_1 + x_2\bar{a}_2 + \ldots + x_m\bar{a}_m,$$
$$\bar{x} = x_1'\bar{a}_1 + x_2'\bar{a}_2 + \ldots + x_m'\bar{a}_m.$$

Вычитая одно равенство из другого, получим

$$\left(x_1 - x_1'\right)\bar{a}_1 + \left(x_2 - x_2'\right)\bar{a}_2 + \ldots + \left(x_m - x_m'\right)\bar{a}_m = \bar{0}.$$

Однако система (1.6) линейно независима, следовательно,

$$x_1 - x_1' = 0, \quad x_2 - x_2' = 0, \quad \ldots, \quad x_m - x_m' = 0.$$

Отсюда

$$x_1 = x_1', \quad x_2 = x_2', \quad \ldots, \quad x_m = x_m'.$$

Итак, разложение (1.7) единственно. Коэффициенты разложения (1.7) называются *координатами вектора $\bar{x}$ в базисе* (1.6).

1.6. ЛИНЕЙНЫЕ НОРМИРОВАННЫЕ ПРОСТРАНСТВА. ЕВКЛИДОВО ПРОСТРАНСТВО

Определение. **Линейным пространством** называется всякое множество V произвольных элементов, называемых векторами, для которых определены операции сложения и умножения на действительные числа, т.е. для любых двух векторов $\bar{u}_1$ и $\bar{u}_2$ из V определен вектор $\bar{u}$, называемый *суммой векторов* $\bar{u}_1$ и $\bar{u}_2$ и обозначаемый $\bar{u}_1 + \bar{u}_2$, а для любого вектора $\bar{u}$ и любого действительного числа λ определен вектор $\lambda\bar{u}$, называемый *произведением вектора $\bar{u}$ на число* λ, причем выполняются следующие условия:

1) $\bar{u}_1 + \bar{u}_2 = \bar{u}_2 + \bar{u}_1$;

2) $\left(\bar{u}_1 + \bar{u}_2\right) + \bar{u}_3 = \bar{u}_1 + \left(\bar{u}_2 + \bar{u}_3\right)$;

3) в множестве V имеется элемент $\bar{0}$, называемый *нулевым элементом*, удовлетворяющий для любого $\bar{u}$ условию

$$\bar{u} + \bar{0} = \bar{u};$$

4) ко всякому вектору $\bar{u}$ имеется вектор $-\bar{u}$, называемый *противоположным* вектору $\bar{u}$ и удовлетворяющий условию

$$\bar{u} + \left(-\bar{u}\right) = \bar{0};$$

5) $\lambda\left(\bar{u}_1+\bar{u}_2\right)=\lambda\bar{u}_1+\lambda\bar{u}_2$;

6) $\left(\lambda_1+\lambda_2\right)\bar{u}=\lambda_1\bar{u}+\lambda_2\bar{u}$;

7) $\lambda_1\left(\lambda_2\bar{u}\right)=\left(\lambda_1\lambda_2\right)\bar{u}$;

8) $1\cdot\bar{u}=\bar{u}$.

(Перечисленные условия называют **аксиомами линейного пространства**.)

Следует подчеркнуть, что множество V может состоять из элементов любой природы.

Примерами линейных пространств являются, в частности:

n-мерное векторное пространство;

множество всех многочленов $P_n(x)$ степени не выше n с обычным сложением и умножением на числа*.

Определение. Линейное пространство V называется **нормированным**, если для любого вектора $\bar{u}$ определена его *норма* $\|\bar{u}\|$, обладающая следующими свойствами:

1) $\|\bar{0}\|=0$;

2) для любого $\bar{u}\neq\bar{0}$ справедливо неравенство $\|\bar{u}\|>0$;

3) для всякого действительного числа λ справедливо равенство

$$\|\lambda\bar{u}\|=|\lambda|\cdot\|\bar{u}\|;$$

4) для любых $\bar{u}$ и $\bar{v}$ из V справедливо неравенство треугольника:

$$\|\bar{u}+\bar{v}\|\leq\|\bar{u}\|+\|\bar{v}\|.$$

Заметим, в частности, что если V есть множество векторов обычной плоскости, то V есть нормированное пространство с нормой $\|\bar{u}\|=|\bar{u}|$, где $|\bar{u}|$ – длина вектора.

Одним из **способов задания нормы в пространстве** V является задание в нем скалярного произведения.

Будем говорить, что *в пространстве V задано скалярное произведение*, если каждой паре векторов $\bar{u}$ и $\bar{v}$ поставлено в соответствие число $\left(\bar{u},\bar{v}\right)$ так, что выполняются следующие условия:

1) $\left(\bar{u},\bar{v}\right)=\left(\bar{v},\bar{u}\right)$;

2) $\left(\lambda\bar{u},\bar{v}\right)=\lambda\left(\bar{u},\bar{v}\right)$;

---

* Убедитесь самостоятельно, что множество всех многочленов, степень которых в точности равна n, не является линейным пространством относительно обычных операций сложения и умножения на числа.

3) $\left(\overline{u}, \overline{v}_1 + \overline{v}_2\right) = \left(\overline{u}, \overline{v}_1\right) + \left(\overline{u}, \overline{v}_2\right)$;

4) $\left(\overline{u}, \overline{u}\right) \geq 0$; при этом $\left(\overline{u}, \overline{u}\right) = 0$ тогда и только тогда, когда $\overline{u}$ – нулевой вектор: $\overline{u} = \overline{0}$.

Если задано скалярное произведение, то норма определяется следующим образом:

$$\|\overline{u}\| = \sqrt{\left(\overline{u}, \overline{u}\right)}. \tag{1.8}$$

Убедимся в том, что норма, определяемая равенством (1.8), обладает всеми четырьмя свойствами, перечисленными в определении. Первые три свойства очевидны, необходимо проверить только неравенство треугольника. Для этого предварительно д о к а ж е м **неравенство Коши – Буняковского**:

$$\left(\overline{u}, \overline{v}\right)^2 \leq \left(\overline{u}, \overline{u}\right)\left(\overline{v}, \overline{v}\right). \tag{1.9}$$

Рассмотрим вектор $\overline{w} = t\overline{u} + \overline{v}$, где t – какое-либо число. Имеем

$$\left(\overline{w}, \overline{w}\right) = \left(t\overline{u} + \overline{v}, t\overline{u} + \overline{v}\right) = t^2\left(\overline{u}, \overline{u}\right) + 2t\left(\overline{u}, \overline{v}\right) + \left(\overline{v}, \overline{v}\right).$$

Обозначим $\left(\overline{u}, \overline{u}\right) = \alpha$, $\left(\overline{u}, \overline{v}\right) = \beta$, $\left(\overline{v}, \overline{v}\right) = \gamma$. Получаем

$$\left(\overline{w}, \overline{w}\right) = \alpha t^2 + 2\beta t + \gamma.$$

Так как $\left(\overline{w}, \overline{w}\right) \geq 0$, то $\alpha t^2 + 2\beta t + \gamma \geq 0$ при всех значениях t. Следовательно, дискриминант этого квадратного трехчлена меньше или равен нулю. Поэтому $\beta^2 - \alpha\gamma \leq 0$, т.е. $\beta^2 \leq \alpha\gamma$, или, что то же самое,

$$\left(\overline{u}, \overline{v}\right)^2 \leq \left(\overline{u}, \overline{u}\right)\left(\overline{v}, \overline{v}\right).$$

Д о к а ж е м теперь неравенство треугольника:

$$\sqrt{\left(\overline{u} + \overline{v}, \overline{u} + \overline{v}\right)} \leq \sqrt{\left(\overline{u}, \overline{u}\right)} + \sqrt{\left(\overline{v}, \overline{v}\right)}.$$

Используя неравенство Коши – Буняковского, получаем

$$\begin{aligned}
&\|\overline{u} + \overline{v}\|^2 = \left(\overline{u} + \overline{v}, \overline{u} + \overline{v}\right) = \\
&= \left(\overline{u}, \overline{u}\right) + \left(\overline{v}, \overline{v}\right) + 2\left(\overline{u}, \overline{v}\right) \leq \|\overline{u}\|^2 + \|\overline{v}\|^2 + 2\|\overline{u}\|\|\overline{v}\| = \\
&= \left(\|\overline{u}\| + \|\overline{v}\|\right)^2,
\end{aligned}$$

что и требовалось доказать.

Ранее (см. § 1.2) мы рассматривали скалярное произведение n-мерных векторов $\bar{a} = (a_1, a_2, \ldots, a_n)$ и $\bar{b} = (b_1, b_2, \ldots, b_n)$, определяемое по формуле (1.3):

$$\left(\bar{a}, \bar{b}\right) = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n.$$

Определение. n-мерное векторное пространство $\mathbf{R}^n$, в котором задано скалярное произведение векторов, называется **евклидовым пространством**.

Длина (норма) вектора $\bar{a} = (a_1, a_2, \ldots, a_n)$ в евклидовом пространстве $\mathbf{R}^n$ есть корень из его скалярного квадрата:

$$\left|\bar{a}\right| = \left(\bar{a}, \bar{a}\right) = \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}. \tag{1.10}$$

Угол между векторами $\bar{a}$ и $\bar{b}$ определяется равенством

$$\cos\varphi = \frac{\left(\bar{a}, \bar{b}\right)}{\left|\bar{a}\right| \cdot \left|\bar{b}\right|}. \tag{1.11}$$

Из неравенства Коши – Буняковского следует, что $\cos\varphi$, определяемый по формуле (1.11), удовлетворяет условию $\left|\cos\varphi\right| \le 1$, поэтому наше определение угла между векторами вполне корректно.

Векторы $\bar{a}$ и $\bar{b}$ называются **ортогональными**, если $\left(\bar{a}, \bar{b}\right) = 0$. Из определения следует, что если два ненулевых вектора ортогональны, то угол между ними равен $\frac{\pi}{2}$.

Будем говорить, что векторы $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_m$ образуют *ортонормированную систему* в n-мерном евклидовом пространстве $\mathbf{R}^n$, если эти векторы попарно ортогональны, т.е. $\left(\bar{e}_i, \bar{e}_j\right) = 0$ при $i \ne j$, $(i, j = 1, 2, \ldots, m)$, и норма каждого из них равна единице: $\left(\bar{e}_i, \bar{e}_i\right) = 1$.

Убедимся в том, что всякая ортонормированная система л и н е й н о н е з а в и с и м а, т.е. что из равенства

$$\lambda_1 \bar{e}_1 + \lambda_2 \bar{e}_2 + \ldots + \lambda_m \bar{e}_m = \bar{0} \tag{1.12}$$

следует $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$.

Пусть i – любое число, удовлетворяющее условию $1 \le i \le m$. Умножим скалярно равенство (1.12) на $\bar{e}_i$:

$$\lambda_1\left(\bar{e}_1, \bar{e}_i\right) + \lambda_2\left(\bar{e}_2, \bar{e}_i\right) + \ldots + \lambda_m\left(\bar{e}_m, \bar{e}_i\right) = 0.$$

Так как $\left(\bar{e}_i, \bar{e}_j\right) = 0$ при $i \ne j$, то последнее равенство равносильно равенству

$$\lambda_i\left(\bar{e}_i, \bar{e}_i\right) = 0.$$

Отсюда, так как $\left(\overline{e}_i, \overline{e}_i\right) \neq 0$, получаем $\lambda_i = 0$. Итак, $\lambda_i = 0$ при всех $i = 1, 2, \ldots, m$.

Из того, что всякая ортонормированная система векторов линейно независима, следует, что ортонормированная система, содержащая n векторов, образует базис в пространстве $\mathbf{R}^n$, называемый *ортонормированным*. Имеет место следующее утверждение.

***Теорема* 1.4.** Во всяком n-мерном евклидовом пространстве существует ортонормированный базис.

Примером ортонормированного базиса является система n единичных векторов: $\overline{e}_1 = (1, 0, \ldots, 0)$, $\overline{e}_2 = (0, 1, \ldots, 0)$, ..., $\overline{e}_n = (0, 0, \ldots, 1)$.

Вопросы

1. Могут ли быть равными два вектора, один из которых – четырехмерный, а другой – пятимерный?
2. Какие векторы получаются из вектора $\overline{a}$ умножением на числа 0 и –1?
3. Какие векторы называются линейно независимыми?
4. Будет ли система векторов $\overline{a} = (1, 2, 3)$, $\overline{b} = (2, 3, 4)$, $\overline{c} = (3, 4, 5)$, $\overline{d} = (4, 5, 6)$ линейно независимой?
5. Образуют ли векторы $\overline{e}_1 = (1, 0, 0, 0)$, $\overline{e}_2 = (0, 2, 0, 0)$, $\overline{e}_3 = (0, 0, 3, 0)$, $\overline{e}_4 = (0, 0, 0, 4)$ базис пространства $\mathbf{R}^4$?
6. Какие числа называются координатами вектора в данном базисе?
7. При каком значении y скалярное произведение векторов $\overline{a} = (1, 2, 3)$ и $\overline{b} = (1, y, 3)$ равно нулю?
8. При каких значениях y векторы $\overline{a} = (1, 2, 3)$ и $\overline{b} = (-3, y, -9)$ образуют линейно независимую систему?

Глава 2
МАТРИЦЫ И ДЕЙСТВИЯ С НИМИ

2.1. ОСНОВНЫЕ ПОНЯТИЯ

Определение. **Матрицей** (числовой матрицей) размера $m \times n$ называется прямоугольная таблица чисел, содержащая m строк и n столбцов:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \qquad (2.1)$$

Числа, составляющие матрицу, называются ее **элементами**.

Для обозначения элементов матрицы используются буквы с двойной индексацией: a_{ij}, где i — номер строки, j — номер столбца. Матрицу записывают также в сокращенном виде:

$$A = (a_{ij}), \quad i = 1, 2, ..., m; \quad j = 1, 2, ..., n. \qquad (2.2)$$

В случае когда число строк матрицы равно числу ее столбцов, т.е. $m = n$, она называется **квадратной матрицей** порядка n.

Матрица может состоять из одной строки или из одного столбца:

$$A = \left(a_{11}, a_{12}, \ldots, a_{1n}\right), \quad B = \begin{pmatrix} b_{11} \\ b_{12} \\ \cdots \\ b_{1m} \end{pmatrix}.$$

Поэтому *вектор-строка* или *вектор-столбец* — это частные случаи матриц.

Элементы матрицы a_{ij}, у которых номер столбца совпадает с номером строки, т.е. $i = j$, называются **диагональными**. Для квадратной матрицы элементы a_{11}, a_{22}, ..., a_{nn} образуют **главную диагональ**.

Матрица называется **симметричной**, если ее элементы, симметричные друг другу относительно главной диагонали, равны между собой:

$$a_{ij} = a_{ji}.$$

Квадратная матрица называется **диагональной**, если все ее элементы вне главной диагонали равны нулю:

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Диагональная матрица называется **единичной**, если все ее диагональные элементы – единицы:

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

Матрица любого размера $m \times n$ называется **нулевой**, если все ее элементы равны нулю:

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

Две матрицы $A = (a_{ij})$ и $B = (b_{ij})$ называются **равными** ($A = B$), если они имеют одинаковые размеры и их соответствующие элементы равны: $a_{ij} = b_{ij}$.

2.2. ЛИНЕЙНЫЕ ОПЕРАЦИИ НАД МАТРИЦАМИ. ТРАНСПОНИРОВАНИЕ МАТРИЦ

Для матриц определены операции сложения и умножения на число.

Суммой матриц $A = (a_{ij})$ и $B = (b_{ij})$ одинакового размера называется матрица $C = (c_{ij})$, элементы которой имеют вид $c_{ij} = a_{ij} + b_{ij}$, $i = 1, 2, ..., m$; $j = 1, 2, ..., n$. В этом случае пишем $C = A + B$.

Пример **2.1.** Пусть $A=\begin{pmatrix}1 & 2\\ 3 & 0\\ 4 & 5\end{pmatrix}$, $B=\begin{pmatrix}3 & 4\\ 2 & 6\\ 5 & 3\end{pmatrix}$. Тогда

$C=A+B=\begin{pmatrix}4 & 6\\ 5 & 6\\ 9 & 8\end{pmatrix}$.

Произведением матрицы $A=(a_{ij})$ **на действительное число** λ называется матрица $\lambda A=(\lambda a_{ij})$.

Пример **2.2.** Пусть $A=\begin{pmatrix}2 & 3 & 4\\ 3 & 2 & 1\end{pmatrix}$, $\lambda=4$. Тогда

$\lambda A=\begin{pmatrix}8 & 12 & 16\\ 12 & 8 & 4\end{pmatrix}$.

Сложение матриц и умножение матриц на число называются **линейными операциями**.

Свойства линейных операций (непосредственно вытекают из их определения):

1. $A+B=B+A$.

2. $(A+B)+C=A+(B+C)$.

3. $\lambda(A+B)=\lambda A+\lambda B$.

4. $(\lambda_1+\lambda_2)A=\lambda_1 A+\lambda_2 A$.

5. $(\lambda_1\lambda_2)A=\lambda_1(\lambda_2 A)=\lambda_2(\lambda_1 A)$.

6. $A+0=A$ (0 – нулевая матрица).

7. Если $\lambda=0$, то $\lambda A=0$ – нулевая матрица.

Транспонированием матрицы называется замена строк матрицы на ее столбцы с сохранением их порядка.

Обозначая через A' матрицу, полученную транспонированием матрицы $A=(a_{ij})$, можем написать: $A'=(a_{ji})$.

В частности, если матрица A есть вектор-строка, то матрица A' – вектор-столбец, и наоборот.

Пример **2.3.** Если $A=\begin{pmatrix}2\\ 5\\ 8\end{pmatrix}$, то $A'=(2 \quad 5 \quad 8)$.

Если $A = \begin{pmatrix} 2 & 3 & 4 & 5 \\ 1 & 5 & 7 & 9 \end{pmatrix}$, то $A' = \begin{pmatrix} 2 & 1 \\ 3 & 5 \\ 4 & 7 \\ 5 & 9 \end{pmatrix}$.

Отметим очевидные **свойства операции транспонирования**:

1. $A'' = A$.

2. Если A – симметричная матрица, то $A' = A$.

2.3. УМНОЖЕНИЕ МАТРИЦ

Умножение матрицы A на матрицу B определено только в тех случаях, когда *число столбцов* матрицы A равно *числу строк* матрицы B.

Произведением матрицы A размера $m \times k$ **на матрицу** B размера $k \times n$ называется матрица $C = (c_{ij})$, элементы которой имеют вид

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

$i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$.

Легко видеть, что элемент c_{ij} матрицы C есть скалярное произведение i-го вектора-строки матрицы A на j-й вектор-столбец матрицы B.

***Пример* 2.4.** Вычислить произведение матриц AB, где

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

Решение. Убеждаемся, что число столбцов матрицы A равно числу строк матрицы B (и равно 3). Следовательно, умножение возможно. Матрица $C = AB$ будет иметь размер 2×3:

$$AB = \begin{pmatrix} 2\cdot 1 + 0\cdot 0 + 1\cdot(-2) & 2\cdot 2 + 0\cdot 1 + 1\cdot 0 & 2\cdot 1 + 0\cdot 2 + 1\cdot 3 \\ -1\cdot 1 + 3\cdot 0 + 2\cdot(-2) & -1\cdot 2 + 3\cdot 1 + 2\cdot 0 & -1\cdot 1 + 3\cdot 2 + 2\cdot 3 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 4 & 5 \\ -5 & 1 & 11 \end{pmatrix}.$$

Перечислим **свойства произведения матриц**. Пусть A, B и C – матрицы таких размеров, что произведение этих матриц определено. Тогда:

1. $(AB)C = A(BC)$.

2. $(A+B)C = AC + BC$.

3. $A(B+C) = AB + AC$.

4. $\lambda(AB) = (\lambda A)B$.

5. $AE = EA = A$.

Заметим, что среди свойств произведения матриц нет свойства коммутативности ($AB = BA$). Более того, если произведение AB существует, то перестановка сомножителей не всегда возможна, т.е. произведение BA может и не существовать. В случае же, когда и AB, и BA существуют, эти произведения могут не совпадать (и даже могут быть матрицами разных размеров).

***Пример* 2.5.** Пусть $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$. Тогда

$$AB = \begin{pmatrix} 4 & 4 \\ 8 & 4 \end{pmatrix}, \quad BA = \begin{pmatrix} 2 & 4 \\ 7 & 6 \end{pmatrix},$$

т.е. $AB \neq BA$.

***Пример* 2.6.** Пусть $A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & -1 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 2 & 0 \\ 1 & -1 \end{pmatrix}$. Тогда

$$AB = \begin{pmatrix} 4 & 0 \\ 1 & -2 \end{pmatrix}, \quad BA = \begin{pmatrix} 1 & -1 & 3 \\ 4 & 2 & 4 \\ 1 & 2 & -1 \end{pmatrix},$$

т.е. $AB \neq BA$, AB и BA разных размеров.

Рассмотрим еще одно **свойство произведения матриц, связанное с операцией транспонирования**.

Если матрицы A и B таковы, что их произведение определено, то имеет место следующее равенство:

6. $(AB)' = B'A'$.

Иначе говоря, матрица, полученная транспонированием произведения, равна произведению матриц, получаемых транспонированием сомножителей, взятых в обратном порядке.

Доказательство. Прежде всего убедимся в том, что если произведение AB определено, то будет определено и произведение $B'A'$. Действительно, если определено произведение AB, то число столбцов матрицы A равно числу строк матрицы B. Но число строк матрицы B равно числу столбцов матрицы B', а число столбцов матрицы A равно числу строк матрицы A', следовательно, произведение $B'A'$ определено. Далее, элемент матрицы $(AB)'$, стоящий в ее i-й строке и j-м столбце, есть элемент матрицы AB, стоящий в ее j-й строке и i-м столбце. Поэтому он равен скалярному произведению j-й строки матрицы A и i-го столбца матрицы B, т.е. сумме произведений соответствующих элементов j-го столбца матрицы A' и i-й строки матрицы B', а это значит, что элемент матрицы $B'A'$, стоящий в ее i-й строке и j-м столбце, также равен скалярному произведению j-й строки матрицы A и i-го столбца матрицы B. Равенство доказано.

2.4. ОБРАТНАЯ МАТРИЦА

Операции деления матриц не существует. Однако при некоторых условиях для квадратных матриц можно определить действие, обратное умножению. Прежде чем сделать это, введем некоторые понятия, которые понадобятся нам в дальнейшем.

Всякую матрицу можно рассматривать как систему ее векторов-строк или векторов-столбцов. Можно доказать, что *ранг системы векторов-строк матрицы равен рангу системы ее векторов-столбцов* (т.е. максимальное число линейно независимых векторов-строк матрицы равно максимальному числу ее линейно независимых векторов-столбцов).

Определение. **Рангом** матрицы называется ранг системы ее векторов-строк (или векторов-столбцов).

Квадратная матрица A порядка n называется **невырожденной**, если ее строки линейно независимы (т.е. ранг A равен n). В противном случае матрица называется **вырожденной**.

Прежде чем определить понятие обратной матрицы, заметим, что для всякого числа $a \neq 0$ существует число, ему обратное: $a^{-1} = \dfrac{1}{a}$, такое, что $aa^{-1} = 1$.

Определение. Пусть A – квадратная матрица. Матрица A^{-1} называется **обратной** по отношению к матрице A, если их произведение равно единичной матрице:

$$AA^{-1} = E.$$

При этом легко убедиться, что умножение матриц A и A^{-1} коммутативно:

$$AA^{-1} = A^{-1}A = E.$$

В дальнейшем покажем, что обратная матрица существует только для невырожденной квадратной матрицы.

Для практического вычисления обратной матрицы применяется, в частности, так называемый *метод элементарных преобразований*.

Элементарными преобразованиями матрицы называются:

- перестановка строк (столбцов);
- умножение строки (столбца) на число, отличное от нуля;
- прибавление к элементам строки (столбца) соответствующих элементов другой строки (столбца), умноженных на некоторое число.

Нетрудно показать, что в результате элементарных преобразований невырожденной матрицы снова получим невырожденную матрицу.

Нахождение обратной матрицы методом элементарных преобразований:

1. Пусть A – невырожденная квадратная матрица. Припишем к ней единичную матрицу E того же размера – получим «сдвоенную» матрицу $(A|E)$.

2. Далее проделаем такие элементарные преобразования над строками матрицы $(A|E)$, чтобы на месте A получилась единичная матрица E. Тогда на месте первоначально приписанной матрицы E возникнет матрица A^{-1}.

(Заметим, что при практическом применении этого метода нет необходимости предварительно проверять невырожденность матрицы A. Это следует из самой возможности приведения A к E.)

***Пример* 2.7.** Дана матрица

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

Найти обратную матрицу A^{-1}.

Решение. Составим матрицу $(A|E)$ и применим метод элементарных преобразований. Здесь l_i обозначает i-ю строку ($i = 1, 2, 3$):

$$(A|E)=\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\2&2&3&0&1&0\\3&3&4&0&0&1\end{array}\right)\xrightarrow{l_2-2l_1,\,l_3-3l_1}\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&-2&-3&-2&1&0\\0&-3&-5&-3&0&1\end{array}\right)\longrightarrow$$

$$\xrightarrow{l_1+l_2,\,2l_3-3l_2}\left(\begin{array}{ccc|ccc}1&0&0&-1&1&0\\0&-2&-3&-2&1&0\\0&0&-1&0&-3&2\end{array}\right)\longrightarrow$$

$$\xrightarrow{l_2-3l_3}\left(\begin{array}{ccc|ccc}1&0&0&-1&1&0\\0&-2&0&-2&10&-6\\0&0&-1&0&-3&2\end{array}\right)\longrightarrow\left(\begin{array}{ccc|ccc}1&0&0&-1&1&0\\0&1&0&1&-5&3\\0&0&1&0&3&-2\end{array}\right).$$

Итак, $A^{-1}=\begin{pmatrix}-1&1&0\\1&-5&3\\0&3&-2\end{pmatrix}$.

С помощью обратной матрицы решают матричные уравнения видов: $AX=B$, $XA=B$.

Пример **2.8.** Решить матричные уравнения:

1) $\begin{pmatrix}1&2\\3&4\end{pmatrix}X=\begin{pmatrix}5&6\\11&12\end{pmatrix}$; 2) $\begin{pmatrix}1&0&2\\2&1&4\\1&1&1\end{pmatrix}X=\begin{pmatrix}1&3&5\\4&7&11\\3&3&3\end{pmatrix}$;

3) $X\begin{pmatrix}1&2&-2\\2&5&-4\\-2&-4&5\end{pmatrix}=\begin{pmatrix}-1&-2&3\\1&4&-1\\5&11&-11\end{pmatrix}$.

Решение. 1. Это уравнение вида $AX=B$. Его решение $X=A^{-1}B$. Здесь $A=\begin{pmatrix}1&2\\3&4\end{pmatrix}$. Найдем A^{-1}:

$$\left(\begin{array}{cc|cc}1&2&1&0\\3&4&0&1\end{array}\right)\to\left(\begin{array}{cc|cc}1&2&1&0\\0&-2&-3&1\end{array}\right)\to\left(\begin{array}{cc|cc}1&0&-2&1\\0&-2&-3&1\end{array}\right)\to\left(\begin{array}{cc|cc}1&0&-2&1\\0&1&3/2&-1/2\end{array}\right),$$

т.е. $A^{-1}=\begin{pmatrix}-2&1\\3/2&-1/2\end{pmatrix}$. Тогда

$$X = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}\begin{pmatrix} 5 & 6 \\ 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}.$$

Итак, $X = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$.

2. Это уравнение также имеет вид $AX = B$. Здесь $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 4 \\ 1 & 1 & 1 \end{pmatrix}$. Найдем A^{-1}:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & 1 & 4 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -2 & 2 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array}\right),$$

т.е. $A^{-1} = \begin{pmatrix} 3 & -2 & 2 \\ -2 & 1 & 0 \\ -1 & 1 & -1 \end{pmatrix}$. Тогда

$$X = \begin{pmatrix} 3 & -2 & 2 \\ -2 & 1 & 0 \\ -1 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 5 \\ 4 & 7 & 11 \\ 3 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}.$$

3. Это уравнение имеет вид $XA = B$. Его решение $X = BA^{-1}$. Здесь $A = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$. Найдем A^{-1}:

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

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & -2 & 5 & -2 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 9 & -2 & 2 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \end{array}\right),$$

т.е. $A^{-1}=\begin{pmatrix} 9 & -2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$. Тогда

$$X=\begin{pmatrix} -1 & -2 & 3 \\ 1 & 4 & -1 \\ 5 & 11 & -11 \end{pmatrix}\begin{pmatrix} 9 & -2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

(Обратите внимание на порядок сомножителей при решении последнего уравнения.)

Пример **2.9.** Решить матричное уравнение

$$\begin{pmatrix} 1 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}X\begin{pmatrix} 1 & 0 & 4 \\ 1 & 1 & 5 \\ 0 & 2 & 3 \end{pmatrix}=\begin{pmatrix} 6 & 1 & 24 \\ 15 & 4 & 62 \\ 11 & 1 & -40 \end{pmatrix}.$$

Решение. Это уравнение вида $AXB = C$. Можно решать его, например, так: сначала найти A^{-1} и, умножив эту матрицу слева на обе части уравнения, получить $XB = A^{-1}C$, затем найти B^{-1} и, умножив ее справа на обе части полученного равенства, найти $X = A^{-1}CB^{-1}$. Можно проделать эту процедуру в обратном порядке. Итак,

$$A=\begin{pmatrix} 1 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix},\quad A^{-1}=\begin{pmatrix} 9 & -2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$ (см. пример 2.8);

$$XB=\begin{pmatrix} 9 & -2 & 2 \\ -2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 6 & 1 & 24 \\ 15 & 4 & 62 \\ 11 & 1 & -40 \end{pmatrix}=\begin{pmatrix} 2 & 3 & 12 \\ 3 & 2 & 14 \\ 1 & 3 & 8 \end{pmatrix},$$

$B=\begin{pmatrix} 1 & 0 & 4 \\ 1 & 1 & 5 \\ 0 & 2 & 3 \end{pmatrix}$. Найдем B^{-1}:

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

$$\to\left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 2 & -2 & 1 \end{array}\right)\to\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -7 & 8 & -4 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -2 & 1 \end{array}\right),$$

т.е. $B^{-1} = \begin{pmatrix} -7 & 8 & -4 \\ -3 & 3 & -1 \\ 2 & -2 & 1 \end{pmatrix}$. Тогда

$$X = \begin{pmatrix} 2 & 3 & 12 \\ 3 & 2 & 14 \\ 1 & 3 & 8 \end{pmatrix} \begin{pmatrix} -7 & 8 & -4 \\ -3 & 3 & -1 \\ 2 & -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Вопросы

1. В каком месте матрицы $A = (a_{ij})$ расположен элемент a_{52}?
2. Может ли матрица состоять: а) из одной строки; б) из одного столбца; в) из одной строки и одного столбца?
3. Может ли какой-нибудь элемент a_{ii} диагональной матрицы быть равным нулю?
4. Могут ли быть равными квадратные матрицы, одна из которых третьего порядка, а другая – четвертого?
5. Можно ли найти сумму двух матриц, одна из которых имеет размер 3×4, а другая – размер 4×3?
6. Существует ли произведение матриц AB, если матрица A имеет размер 3×4, а матрица B – размер 4×2? Существует ли для этих матриц произведение BA?
7. Можно ли найти произведение двух матриц, одна из которых – квадратная, а другая не является квадратной?
8. Пусть для матриц A и B существуют произведения AB и BA. Можно ли утверждать, что матрицы AB и BA одного размера?
9. Может ли произведение двух ненулевых матриц быть нулевой матрицей?
10. Что такое квадрат матрицы? Может ли квадрат ненулевой матрицы быть нулевой матрицей?
11. Существует ли обратная матрица A^{-1} для диагональной матрицы $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$? Если A^{-1} существует, то каков ее вид?

Глава 3
ОПРЕДЕЛИТЕЛИ

3.1. ОСНОВНЫЕ ПОНЯТИЯ

Существует правило, в соответствии с которым каждой квадратной матрице ставится в соответствие ч и с л о, характеризующее эту матрицу.

Рассмотрим матрицу второго порядка:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{3.1}$$

Число $a_{11}a_{22} - a_{12}a_{21}$ называется **определителем** (или **детерминантом**) матрицы (3.1) и записывается в виде

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

(это определитель второго порядка).

Итак,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{3.2}$$

Например,

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

Понятие определителя связано, в частности, с решением систем линейных уравнений. Рассмотрим систему

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

Чтобы исключить x_2, умножим первое уравнение на a_{22}, а второе – на a_{12} и из первого уравнения вычтем второе:

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12}, \text{ или } dx_1 = d_1,$$

где $d = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $d_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}$.

Если $d \neq 0$, то $x_1 = \dfrac{d_1}{d}$. Аналогично получим

$$dx_2 = d_2, \quad x_2 = \frac{d_2}{d},$$

где $d_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$.

Применение формул $x_1 = \dfrac{d_1}{d}$ и $x_2 = \dfrac{d_2}{d}$ для решения системы уравнений называется **правилом Крамера**.

Пример **3.1.** Решить систему

$$\begin{cases} 2x_1 + 3x_2 = 7, \\ x_1 + 4x_2 = 6. \end{cases}$$

Решение:

$$d = \begin{vmatrix} 2 & 3 \\ 1 & 4 \end{vmatrix} = 2 \cdot 4 - 3 \cdot 1 = 5, \quad d_1 = \begin{vmatrix} 7 & 3 \\ 6 & 4 \end{vmatrix} = 7 \cdot 4 - 3 \cdot 6 = 10,$$

$$d_2 = \begin{vmatrix} 2 & 7 \\ 1 & 6 \end{vmatrix} = 2 \cdot 6 - 7 \cdot 1 = 5,$$

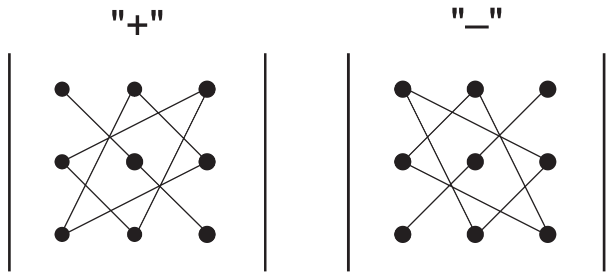
$$x_1 = \frac{d_1}{d} = \frac{10}{5} = 2, \quad x_2 = \frac{d_2}{d} = \frac{5}{5} = 1.$$

Аналогичное правило для систем с любым числом неизвестных будет рассмотрено далее.

Для квадратной матрицы третьего порядка определитель (или детерминант) третьего порядка есть число, определяемое формулой

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \tag{3.3}$$

Определитель третьего порядка есть алгебраическая сумма шести произведений элементов, взятых по одному из каждой строки и каждого столбца. Эти произведения называются *членами определителя*. Формула (3.3) схематически может быть изображена в следующем виде:



Со знаком **«плюс»** берутся произведения, сомножители которых находятся на главной диагонали и в вершинах равнобедренных треугольников с основаниями, параллельными главной диагонали; со знаком **«минус»** – на побочной диагонали и в вершинах равнобедренных треугольников с основаниями, параллельными побочной диагонали.

Введем теперь понятие определителя любого порядка.

Рассмотрим натуральные числа от 1 до n. Эти n чисел могут быть записаны в том или ином порядке. Всякое расположение чисел 1, 2, ..., n называется их **перестановкой**. Например,

$$3,\ 1,\ 5,\ 4,\ 2 \tag{3.4}$$

есть перестановка из пяти чисел.

Нетрудно доказать, что число различных перестановок из n чисел равно произведению $1 \cdot 2 \cdots n$, обозначаемому $n!$ (читается: «эн факториал»).

Рассмотрим произвольную перестановку $\alpha_1, \alpha_2, ..., \alpha_n$, составленную из n первых натуральных чисел. Выберем в этой перестановке два произвольных числа α_i и α_j. Если $i < j$, но $\alpha_i > \alpha_j$, то говорят, что числа α_i и α_j образуют **инверсию**. (Иначе говоря, если в перестановке большее число предшествует меньшему, то эти два числа образуют инверсию.) В частности, в перестановке (3.4) числа $\alpha_3 = 5$ и $\alpha_5 = 2$ образуют инверсию. Если же при $i < j$ выполняется неравенство $\alpha_i < \alpha_j$ (т.е. меньшее число предшествует большему), то числа α_i и α_j инверсию не образуют. Например, в перестановке (3.4) числа $\alpha_2 = 1$ и $\alpha_4 = 4$ не образуют инверсию.

Перестановка называется **четной**, если в ней четное число инверсий. В противном случае перестановка называется **нечетной**.

Рассмотрим квадратную матрицу:

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

Составим какое-нибудь произведение n ее элементов, взятых по одному из каждой строки и каждого столбца:

$$a_{1\alpha_1} a_{2\alpha_2} \cdots a_{n\alpha_n}. \tag{3.5}$$

В произведении (3.5) сомножители записаны в порядке возрастания их первых индексов – номеров строк. Вторые индексы – номера столбцов – образуют перестановку $\alpha_1, \alpha_2, ..., \alpha_n$. Произведение (3.5) называется *членом определителя* матрицы A. Определим **правило знаков**: если вторые индексы образуют четную перестановку, то произведение (3.5) берется со знаком «плюс», если нечетную – со знаком «минус». Таких произведений (3.5), очевидно, $n!$ – столько, сколько различных перестановок образуют их вторые индексы.

Определение. **Определителем** квадратной матрицы порядка n (определителем n-го порядка) называется составленная в соответствии с правилом знаков алгебраическая сумма $n!$ членов, каждый из которых является произведением n элементов матрицы, взятых по одному из каждой строки и каждого столбца.

Определитель n-го порядка обозначается так:

$$d = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{3.6}$$

В дальнейшем будем говорить об элементах, строках и столбцах определителя, имея в виду элементы, строки и столбцы соответствующей матрицы.

Следует отметить, что вычислять определитель n-го порядка при $n > 3$, основываясь непосредственно на определении, было бы затруднительно. (Например, чтобы найти определитель шестого порядка, надо вычислить сумму 720 слагаемых, каждое из которых – произведение шести элементов). Поэтому для того, чтобы вычислить определитель, его предварительно упрощают, преобразовывая с учетом его свойств.

3.2. СВОЙСТВА ОПРЕДЕЛИТЕЛЕЙ

1. *Определитель не меняется при транспонировании*: $|A| = |A'|$.

Из свойства 1 вытекает, что всякое утверждение о строках определителя справедливо и для столбцов и наоборот. В этом смысле *строки и столбцы определителя равноправны*. Поэтому далее будем формулировать свойства для строк, имея в виду при этом, что они справедливы и для столбцов.

2. *Если одна из строк определителя состоит из нулей, то определитель равен нулю.*

Доказывается это свойство просто. Пусть все элементы i-й строки определителя равны нулю. Каждый член определителя содержит множитель, являющийся элементом этой строки. Поэтому каждый член определителя равен нулю, а следовательно, и сам определитель равен нулю.

3. *При перестановке двух строк определитель меняет знак.*

(Иначе говоря, если поменять местами две строки в матрице A, получим матрицу B, такую, что $|A| = -|B|$.)

4. *Определитель, содержащий две одинаковые строки, равен нулю.*

Для доказательства поменяем местами эти две одинаковые строки. Определитель не изменится, но по свойству 3 он поменяет знак: $d = -d$. Следовательно, $d = 0$.

5. *Если все элементы какой-либо строки определителя умножить на число k, то определитель умножится на это число k.*

Действительно, в этом случае каждый член определителя умножится на число k, следовательно, весь определитель умножится на это число.

Из свойства 5 следует, что *общий множитель любой строки определителя можно вынести за знак определителя*.

6. *Определитель, содержащий две пропорциональные строки, равен нулю.*

Докажем это утверждение. Пусть i-я и j-я строки пропорциональны: элементы j-й строки получаются умножением элементов i-й строки на число k. Вынесем k за знак определителя и получим определитель, содержащий две одинаковые строки. По свойству 4 он равен нулю.

7. *Если элементы одной из строк определителя d имеют вид*

$$a_{ik} = b_{ik} + c_{ik} \quad (k = 1, 2, ..., n),$$

то определитель равен сумме двух определителей d_1 *и* d_2, *у которых все строки, кроме указанной, совпадают с соответствующими строками определителя d, а на месте указанной строки определитель* d_1 *содержит строку, состоящую из элементов* b_{ik}, *а определитель* d_2 *— строку, состоящую из элементов* c_{ik} $(k = 1, 2, ..., n)$:

$$d = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{i1}+c_{i1} & b_{i2}+c_{i2} & \cdots & b_{in}+c_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = d_1 + d_2 =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{i1} & c_{i2} & \cdots & c_{in} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Это утверждение легко следует из того, что всякий член определителя d может быть представлен как сумма двух слагаемых, одно из которых есть член определителя d_1, а другое — член определителя d_2.

8. *Если одна из строк определителя есть линейная комбинация других его строк, то определитель равен нулю.*

Свойство 8 есть обобщение свойства 6.

9. *Определитель не меняется, если к элементам любой его строки прибавить соответствующие элементы другой строки, умноженные на одно и то же число.*

Это свойство есть следствие свойств 4–7.

10. *Определитель произведения двух квадратных матриц равен произведению их определителей*:

$$|AB| = |A| \cdot |B|.$$

В заключение еще раз напомним, что все утверждения, сформулированные здесь для строк определителя, остаются верными и для его столбцов. (Имеются в виду строки и столбцы соответствующей матрицы.)

3.3. МИНОРЫ И АЛГЕБРАИЧЕСКИЕ ДОПОЛНЕНИЯ

Рассмотрим определитель n-го порядка. Выделим в нем какой-либо элемент a_{ij} и вычеркнем i-ю строку и j-й столбец, на пересечении которых расположен этот элемент. Получится определитель $(n - 1)$-го порядка, который называется **минором** M_{ij} **элемента** a_{ij}.

Например, возьмем определитель четвертого порядка:

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

Минор M_{23} элемента a_{23} получается вычеркиванием второй строки и третьего столбца, на пересечении которых стоит элемент $a_{23} = 2$:

$$M_{23} = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 2 & 3 \\ 1 & 0 & 3 \end{vmatrix} = 4.$$

Определение. **Алгебраическим дополнением** элемента a_{ij} определителя (3.6) называется число

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

В частности, в приведенном примере алгебраическое дополнение

$$A_{23} = (-1)^5 \cdot 4 = -4.$$

Миноры и алгебраические дополнения играют важную роль в линейной алгебре и ее приложениях. Одним из таких приложений является следующее утверждение.

***Теорема* 3.1.** Определитель равен сумме произведений элементов любой его строки на их алгебраические дополнения:

$$d = a_{i1}A_{i1} + a_{i2}A_{i2} + \ldots + a_{in}A_{in}. \qquad (3.7)$$

(Примем эту теорему без доказательства.)

Формула (3.7) называется **разложением** определителя по i-й строке. Аналогичное утверждение справедливо и для разложения определителя по любому столбцу. Формула (3.7) сводит вычисление определителя n-го порядка к вычислению n определителей $(n-1)$-го порядка.

З а м е ч а н и е. Сумма попарных произведений элементов i-й строки (столбца) определителя на соответствующие алгебраические дополнения к элементам j-й строки (столбца) при $i \neq j$ равна нулю, т.е.

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = 0,$$

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \ldots + a_{ni}A_{nj} = 0$$

при $i \neq j$.

Д о к а ж е м, например, последнее из этих двух равенств. Разложим определитель

$$d = \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

по j-му столбцу:

$$d = a_{1j}A_{1j} + a_{2j}A_{2j} + \ldots + a_{nj}A_{nj}.$$

Теперь заменим элементы j-го столбца элементами i-го столбца (оставляя i-й столбец неизменным). Получим определитель

$$d' = \begin{vmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2i} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{ni} & \cdots & a_{ni} & \cdots & a_{nn} \end{vmatrix},$$

содержащий два одинаковых столбца на i-м и j-м местах и, очевидно, равный нулю: $d' = 0$. Его разложение по j-му столбцу имеет вид

$$d' = a_{1i}A_{1j} + a_{2i}A_{2j} + \ldots + a_{ni}A_{nj}.$$

Следовательно,

$$a_{1i}A_{1j} + a_{2i}A_{2j} + \ldots + a_{ni}A_{nj} = 0,$$

что и требовалось доказать.

Обычно перед вычислением определитель предварительно преобразуют с учетом его свойств. Нередко его приводят к треугольному виду, так как справедливо следующее ***утверждение***:

Если все элементы определителя, расположенные по одну сторону от главной диагонали, равны нулю, то этот определитель равен произведению элементов, стоящих на главной диагонали.

Докажем это утверждение, применив метод математической индукции (см. с. 142).

Для определителя второго порядка это утверждение очевидно. Предположим, что оно доказано для определителя $(n-1)$-го порядка, и рассмотрим определитель n-го порядка:

$$d = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

Разложим его по первому столбцу:

$$d = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

В правой части полученного равенства – определитель $(n-1)$-го порядка. Для него справедливо равенство

$$\begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{22}a_{33}\cdots a_{nn}.$$

Следовательно, $d = a_{11}a_{22}a_{33}\cdots a_{nn}$.

***Пример* 3.2.** Вычислить следующие определители:

$$1)\ d_1 = \begin{vmatrix} 1 & 2 & 4 & 3 \\ 0 & 3 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 3 & 2 & 1 \end{vmatrix};\quad 1)\ d_2 = \begin{vmatrix} 1 & 2 & 3 & 2 \\ 1 & 3 & 4 & 5 \\ 1 & 2 & 5 & 3 \\ 1 & 2 & 7 & 8 \end{vmatrix}.$$

Р е ш е н и е. 1. Разложить определитель d_1 можно по любой строке. Однако минимальный объем вычислений получается при разложении его по строке с наибольшим числом нулей. Разлагаем d_1 по третьей строке, а затем A_{32} по второй строке:

$$d_1 = -2 \cdot \begin{vmatrix} 1 & 4 & 3 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{vmatrix} = -2 \cdot 1 \cdot \begin{vmatrix} 1 & 3 \\ 1 & 1 \end{vmatrix} = 4.$$

2. Вычтем первую строку определителя d_2 из всех остальных, а затем вычтем удвоенную третью строку из четвертой:

$$d_2 = \begin{vmatrix} 1 & 2 & 3 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 4 & 6 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 2 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 \end{vmatrix} = 1 \cdot 1 \cdot 2 \cdot 4 = 8.$$

Мы воспользовались тем, что полученный «треугольный» определитель равен произведению элементов главной диагонали.

3.4. ПРИМЕНЕНИЕ ОПРЕДЕЛИТЕЛЕЙ

О п р е д е л е н и е. Квадратная матрица A называется **невырожденной**, если ее определитель отличен от нуля: $|A| \neq 0$. В противном случае матрица называется **вырожденной**.

***Т е о р е м а* 3.2.** Для матрицы A обратная матрица A^{-1} существует тогда и только тогда, когда матрица A невырожденная.

Д о к а з а т е л ь с т в о. *Необходимость*. Пусть матрица A имеет обратную матрицу A^{-1}. Тогда $A^{-1}A = AA^{-1} = E$. Так как $|E| = 1 \neq 0$ и так как определитель произведения матриц равен произведению их определителей, то $|A^{-1}A| = |A^{-1}| \cdot |A| = |E| \neq 0$, следовательно, $|A^{-1}| \neq 0$ и $|A| \neq 0$.

Достаточность. Пусть дана невырожденная матрица

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

и ее определитель $|A| = d \neq 0$.

Транспонируем матрицу A и затем заменим ее элементы их алгебраическими дополнениями:

$$A^* = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}. \qquad (3.8)$$

Матрица A^* называется **присоединенной** к матрице A.

Найдем произведение AA^*. Учитывая разложение (3.7) и следующее за ним замечание, получим

$$AA^* = \begin{pmatrix} d & 0 & \cdots & 0 \\ 0 & d & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d \end{pmatrix} = dE. \qquad (3.9)$$

(Распишите это произведение AA^* подробно и убедитесь в справедливости равенства (3.9) самостоятельно.)

Нетрудно убедиться также в том, что $AA^* = A^*A$.

Итак, $AA^* = A^*A = dE$. Отсюда

$$A\frac{1}{d}A^* = \frac{1}{d}A^*A = E.$$

Следовательно,

$$A^{-1} = \frac{1}{d}A^*,$$

или, более подробно,

$$A^{-1} = \frac{1}{d}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}. \qquad (3.10)$$

Теорема доказана.

Вычисление обратной матрицы методом присоединенной матрицы:

1. Находим определитель исходной матрицы $d = |A|$. Если $d = 0$, т.е. матрица A вырожденная, то обратной матрицы не существует. Если же $d \neq 0$, продолжаем процесс.

2. Находим алгебраические дополнения элементов матрицы A и составляем из них присоединенную матрицу A^*.

3. Находим обратную матрицу по формуле (3.10). В ряде случаев полезно сделать проверку, т.е. вычислить произведения $A^{-1}A$ и AA^{-1} (или одно из них) и убедиться, что получилась единичная матрица E.

Пример **3.3.** Найти матрицу, обратную матрице A:

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

Решение. 1. Вычисляем определитель:

$$d=\left|A\right|=(3-2)-2(2-1)+3(4-3)=2.$$

2. Находим алгебраические дополнения:

$$A_{11}=(-1)^{1+1}\begin{vmatrix}3 & 1\\ 2 & 1\end{vmatrix}=1,\qquad A_{21}=-\begin{vmatrix}2 & 3\\ 2 & 1\end{vmatrix}=4,\qquad A_{31}=-7,$$

$$A_{12}=(-1)^{1+2}\begin{vmatrix}2 & 1\\ 1 & 1\end{vmatrix}=-1,\qquad A_{22}=\begin{vmatrix}1 & 3\\ 1 & 1\end{vmatrix}=-2,\qquad A_{32}=5,$$

$$A_{13}=(-1)^{1+3}\begin{vmatrix}2 & 3\\ 1 & 2\end{vmatrix}=1,\qquad A_{23}=-\begin{vmatrix}1 & 2\\ 1 & 2\end{vmatrix}=0,\qquad A_{33}=-1.$$

Составляем присоединенную матрицу:

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

3. Вычисляем обратную матрицу:

$$A^{-1}=\frac{1}{2}\begin{pmatrix}1 & 4 & -7\\ -1 & -2 & 5\\ 1 & 0 & -1\end{pmatrix}=\begin{pmatrix}1/2 & 2 & -7/2\\ -1/2 & -1 & 5/2\\ 1/2 & 0 & -1/2\end{pmatrix}.$$

При желании проверить себя легко убеждаемся в том, что $A^{-1}A=$ $=AA^{-1}=E$.

Заметим, что для нахождения обратных матриц высокого порядка обычно применяют другой метод – метод элементарных преобразований (см. § 2.4).

3.5. РАНГ МАТРИЦЫ

Пусть A – матрица размера $m \times n$. Выделим в ней k строк и k столбцов произвольным образом. Элементы, находящиеся на пересечении выделенных строк и столбцов, образуют квадратную матрицу k-го порядка; ее определитель называется минором k-го порядка матрицы A. При этом, очевидно, $k \leq \min(m, n)$.

Определение. Наибольший порядок отличных от нуля миноров матрицы A называется **рангом** матрицы A.

***Пример* 3.4.** Вычислить ранг матрицы

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

Решение. Легко видеть, что ранг матрицы A равен двум: $\operatorname{rg} A = 2$. Действительно, минор второго порядка, стоящий в левом верхнем углу, отличен от нуля:

$$\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = 1 \neq 0,$$

однако всякий минор третьего порядка содержит либо нулевую строку, либо две пропорциональные строки и, следовательно, равен нулю.

В § 1.4 мы определили ранг системы векторов как наибольшее число ее линейно независимых векторов. В связи с этим естественно определить **ранг матрицы** как *наибольшее число ее линейно независимых строк*. Это определение равносильно предыдущему. Можно доказать (это делается в курсе алгебры), что наибольшее число линейно независимых строк матрицы равно наибольшему числу ее линейно независимых столбцов, а также наибольшему порядку отличных от нуля миноров.

Приведем **основные методы вычисления ранга матрицы**.

1. Метод окаймляющих миноров. Пусть в матрице A найден минор M порядка k, отличный от нуля. Рассматриваем миноры $(k + 1)$-го порядка, но не все, а только те, которые содержат в себе («окаймляют») минор M. Если все они равны нулю, то ранг матрицы равен k. В противном случае процедура продолжается.

П р и м е р **3.5.** Найти ранг матрицы

$$A = \begin{pmatrix} 2 & 3 & 0 & 2 & 4 \\ 5 & 8 & 0 & 2 & 5 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix}.$$

Р е ш е н и е. Фиксируем отличный от нуля минор второго порядка:

$$M_2 = \begin{vmatrix} 2 & 3 \\ 5 & 8 \end{vmatrix} = 1 \neq 0.$$

Один из окаймляющих миноров третьего порядка также отличен от нуля:

$$M_3 = \begin{vmatrix} 2 & 3 & 0 \\ 5 & 8 & 0 \\ 0 & 0 & 2 \end{vmatrix} = 2 \neq 0.$$

Однако оба минора четвертого порядка, окаймляющие M_3, равны нулю (так как у каждого из них третья и четвертая строки пропорциональны). Следовательно, ранг матрицы равен трем: $\mathrm{rg}\, A = 3$.

2. Метод элементарных преобразований. Элементарные преобразования не меняют ранга матрицы. Применяя элементарные преобразования, можно привести матрицу к такому виду, когда все элементы, кроме a_{11}, a_{22}, ..., a_{rr}, равны нулю. Число отличных от нуля элементов преобразованной матрицы, очевидно, равно рангу матрицы.

П р и м е р **3.6.** Найти ранг матрицы

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

Р е ш е н и е:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 0 & -1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & -2 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ранг преобразованной матрицы равен двум, следовательно, $\operatorname{rg} A = 2$.

Вопросы

1. При каких условиях определитель матрицы второго порядка равен нулю?
2. С каким знаком в определитель матрицы четвертого порядка входит слагаемое $a_{11}a_{23}a_{34}a_{42}$?
3. Может ли произведение $a_{12}a_{23}a_{32}a_{41}a_{54}$, взятое с соответствующим знаком, быть членом определителя матрицы пятого порядка?
4. Чем отличается минор M_{54} от алгебраического дополнения A_{54}?
5. Пусть матрица A содержит минор пятого порядка, отличный от нуля. Что можно сказать о ранге матрицы A?
6. Чему равна сумма произведений элементов какой-нибудь строки матрицы на алгебраические дополнения элементов другой строки этой матрицы?
7. Можно ли вычислить определитель произведения двух квадратных матриц, не перемножая эти матрицы?
8. Чему равен определитель треугольной матрицы?
9. Какой метод для вычисления обратной матрицы седьмого порядка предпочтительнее: метод присоединенной матрицы или метод элементарных преобразований?
10. Может ли ранг матрицы A размера 7×3 равняться четырем?

Глава 4
СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

4.1. ОСНОВНЫЕ ПОНЯТИЯ

Система m линейных уравнений с n неизвестными имеет вид

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. \end{cases} \tag{4.1}$$

Матрица, составленная из коэффициентов уравнений системы (4.1), т.е.

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

называется **матрицей системы**.

Если обозначить через X матрицу-столбец неизвестных, а через B матрицу-столбец свободных членов:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix},$$

то система (4.1) запишется в виде одного матричного уравнения:

$$AX = B.$$

Добавив к матрице A столбец свободных членов, получим **расширенную матрицу системы** (4.1):

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

(Обычно столбец свободных членов отделяют вертикальной чертой.)

Расширенная матрица содержит всю информацию о системе.

Решением системы (4.1) называется набор чисел

$$x_1 = \alpha_1,\ x_2 = \alpha_2,\ \ldots,\ x_n = \alpha_n,$$

при подстановке которых в эту систему все уравнения обращаются в тождества.

Система уравнений называется **совместной**, если она имеет хотя бы одно решение. Система, не имеющая ни одного решения, называется **несовместной**. Совместная система, имеющая единственное решение, называется **определенной**. Если же система имеет более одного решения, она называется **неопределенной**.

Решить систему – значит найти множество всех ее решений. Множество всех решений системы называется ее **общим решением**.

Две системы называются **эквивалентными** (или **равносильными**), если они имеют одно и то же множество решений, или, что то же самое, одно и то же общее решение.

Обычно для того, чтобы решить систему, ее сначала преобразовывают. При этом преобразованная система должна быть эквивалентна исходной.

Перечислим **элементарные преобразования** системы (4.1):

- перестановка уравнений;
- умножение обеих частей одного уравнения на любое число, отличное от нуля;
- прибавление к обеим частям одного из уравнений системы соответствующих частей другого уравнения, умноженных на одно и то же число;
- вычеркивание уравнения вида $0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0$.

В результате элементарных преобразований получается система, эквивалентная исходной.

4.2. МЕТОДЫ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

1. Метод Гаусса. Это наиболее удобный метод решения систем вида (4.1). Изложим его суть.

Пусть для определенности в системе (4.1) $a_{11} \neq 0$ (если $a_{11} = 0$, то переставим на первое место другое уравнение с ненулевым первым коэффициентом). Умножим п е р в о е уравнение* на $-\dfrac{a_{21}}{a_{11}}$ и прибавим ко второму. Затем умножим первое уравнение на $-\dfrac{a_{31}}{a_{11}}$ и прибавим к третьему и т.д. Наконец, умножим первое уравнение на $-\dfrac{a_{m1}}{a_{11}}$ и прибавим к последнему уравнению. В результате этих элементарных преобразований получим систему, эквивалентную исходной, однако в новой системе ни одно из уравнений, кроме первого, не содержит неизвестное x_1:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1, \\ \qquad\quad a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n = b'_2, \\ \qquad\quad a'_{32}x_2 + a'_{33}x_3 + \ldots + a'_{3n}x_n = b'_3, \\ \qquad\quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\quad a'_{m2}x_2 + a'_{m3}x_3 + \ldots + a'_{mn}x_n = b'_m. \end{cases} \tag{4.2}$$

Заметим, что преобразованы только коэффициенты и свободные члены, поэтому удобнее записывать преобразование системы как преобразование ее расширенной матрицы:

$$\bar{A}' = \left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} & b'_2 \\ 0 & a'_{32} & a'_{33} & \cdots & a'_{3n} & b'_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a'_{m2} & a'_{m3} & \cdots & a'_{mn} & b'_m \end{array}\right). \tag{4.3}$$

На втором этапе с помощью в т о р о г о уравнения аналогичным образом преобразуем все уравнения, начиная с третьего, или, что то же самое, умножая вторую строку матрицы $\bar{A}'$ на соответ-

* Говоря об умножении уравнения на число, мы, разумеется, имеем в виду умножение всех членов обеих частей этого уравнения на это число.

ствующие числа $\left(-\frac{a'_{32}}{a'_{22}}, \ldots, -\frac{a'_{m2}}{a'_{22}}\right)$ и прибавляя к третьей, ..., m-й строкам, получим

$$\bar{A}'' = \left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} & b'_2 \\ 0 & 0 & a''_{33} & \cdots & a''_{3n} & b''_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & a''_{m3} & \cdots & a''_{mn} & b''_m \end{array}\right).$$

Продолжим этот процесс аналогичным образом: далее будут преобразовываться все строки, кроме первых двух, затем кроме первых трех и т.д.

Мы не исследовали заранее систему на совместность. Тем не менее метод Гаусса позволяет на одном из этапов установить возможность несовместности системы. Действительно, если в результате преобразований получим строку, у которой все члены, кроме последнего, равны нулю, а последний – отличен от нуля, то это соответствует уравнению вида

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = b \neq 0,$$

не имеющему решений. Стало быть, система, содержащая такое уравнение, несовместна.

В процессе применения метода Гаусса могут также появляться строки, целиком состоящие из нулей, что соответствует уравнениям вида

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n = 0.$$

Такое может произойти, если соответствующие уравнения исходной системы являются линейными комбинациями других уравнений системы.

Если система (4.1) определенна, то ее матрица в результате преобразований* примет вид

$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a'_{22} & a'_{23} & \cdots & a'_{2n} & b'_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} & b_n^{(n-1)} \end{array}\right),$$

* Заметим, что метод нахождения обратной матрицы (см. § 2.4) основан на аналогичных преобразованиях.

т.е. система будет иметь треугольный вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n = b_1, \\ \qquad a'_{22}x_2 + a'_{23}x_3 + \ldots + a'_{2n}x_n = b'_2, \\ \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad\qquad a_{nn}^{(n-1)}x_n = b_n^{(n-1)} \end{cases} \tag{4.4}$$

(верхние индексы и штрихи указывают, сколько раз менялись коэффициенты и свободные члены в процессе преобразований).

Из последнего уравнения системы (4.4) сразу найдем $x_n = \dfrac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}$, затем, подставив найденное x_n в предпоследнее уравнение (содержащее только x_n и x_{n-1}), найдем x_{n-1} и т.д. Таким образом, последовательно найдем все остальные неизвестные. (Такой процесс иногда называют *обратным ходом метода Гаусса*.)

***Пример* 4.1.** Решить систему

$$\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 3, \\ 2x_1 + 3x_2 + 5x_3 + 3x_4 = 9, \\ x_1 - x_2 + 2x_3 = -3, \\ 2x_1 + 2x_2 + 6x_3 + 4x_4 = 8. \end{cases}$$

Решение. Составим расширенную матрицу системы и применим метод Гаусса:

$$\left(\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 3 \\ 2 & 3 & 5 & 3 & 9 \\ 1 & -1 & 2 & 0 & -3 \\ 2 & 2 & 6 & 4 & 8 \end{array}\right) \xrightarrow{l_2-2l_1,\, l_3-l_1,\, l_4-2l_1} \left(\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 1 & 3 \\ 0 & -2 & 0 & -1 & -6 \\ 0 & 0 & 2 & 2 & 2 \end{array}\right) \to$$

$$\xrightarrow{l_3+2l_2} \left(\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 1 & 3 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 2 & 2 \end{array}\right) \xrightarrow{l_4-l_3} \left(\begin{array}{cccc|c} 1 & 1 & 2 & 1 & 3 \\ 0 & 1 & 1 & 1 & 3 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right).$$

Полученная расширенная матрица соответствует системе

$$\begin{cases} x_1 + x_2 + 2x_3 + x_4 = 3, \\ \qquad x_2 + x_3 + x_4 = 3, \\ \qquad\qquad 2x_3 + x_4 = 0, \\ \qquad\qquad\qquad x_4 = 2, \end{cases}$$

которая эквивалентна исходной системе. Подставляя $x_4 = 2$ в предпоследнее уравнение, находим $x_3 = -1$; подставляя x_4 и x_3 во второе уравнение, находим $x_2 = 2$; наконец, подставляя найденные x_4, x_3, x_2 в первое уравнение, находим $x_1 = 1$.

Итак, система имеет единственное решение:

$$x_1 = 1, x_2 = 2, x_3 = -1, x_4 = 2.$$

Можно решать такую систему, не приводя ее к треугольному виду, а превращая в так называемую *разрешенную систему*. Поясним это на примере, а затем опишем процесс в общем виде.

Пример **4.2.** Решить систему

$$\begin{cases} x_1 + 2x_2 + x_3 + 4x_4 = 7, \\ 3x_1 + 2x_2 + x_3 + x_4 = 3, \\ x_1 + x_2 + 2x_3 + 2x_4 = 4, \\ 2x_1 + 5x_2 + 3x_3 + 9x_4 = 15. \end{cases}$$

Р е ш е н и е. Составим расширенную матрицу и преобразуем ее таким образом, чтобы каждая строка и каждый столбец преобразованной матрицы системы содержали один элемент, равный единице, а остальные – равные нулю:

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 7 \\ 3 & 2 & 1 & 1 & 3 \\ 1 & 1 & 2 & 2 & 4 \\ 2 & 5 & 3 & 9 & 15 \end{array}\right) \xrightarrow{l_2-3l_1,\, l_3-l_1,\, l_4-2l_1} \left(\begin{array}{cccc|c} 1 & 2 & 1 & 4 & 7 \\ 0 & -4 & -2 & -11 & -18 \\ 0 & -1 & 1 & -2 & -3 \\ 0 & 1 & 1 & 1 & 1 \end{array}\right) \to$$

$$\xrightarrow{l_3+l_4,\, l_2+4l_4,\, l_1-2l_4} \left(\begin{array}{cccc|c} 1 & 0 & -1 & 2 & 5 \\ 0 & 0 & 2 & -7 & -14 \\ 0 & 0 & 2 & -1 & -2 \\ 0 & 1 & 1 & 1 & 1 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 0 & -1 & 2 & 5 \\ 0 & 0 & 2 & -7 & -14 \\ 0 & 0 & -2 & 1 & 2 \\ 0 & 1 & 1 & 1 & 1 \end{array}\right) \to$$

$$\xrightarrow{l_4-l_3,\, l_2+7l_3,\, l_1-2l_3} \left(\begin{array}{cccc|c} 1 & 0 & 3 & 0 & 1 \\ 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & -2 & 1 & 2 \\ 0 & 1 & 3 & 0 & -1 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 0 & 3 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 1 & 2 \\ 0 & 1 & 3 & 0 & -1 \end{array}\right) \to$$

$$\xrightarrow{l_1-3l_2,\, l_3+2l_2,\, l_4-3l_2} \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & -1 \end{array}\right).$$

Итак, получена система

$$\begin{cases} x_1 = 1, \\ x_3 = 0, \\ x_4 = 2, \\ x_2 = -1, \end{cases}$$

т.е. решение $x_1 = 1, x_2 = -1, x_3 = 0, x_4 = 2$, или $(1, -1, 0, 2)$.

Очевидно, система (4.1) приводится к виду (4.4), когда ранг ее матрицы совпадает с числом неизвестных: $\operatorname{rg} A = n$.

В случае же, когда система совместна и $\operatorname{rg} A = r < n$, расширенная матрица $\overline{A}$ методом Гаусса преобразовывается к виду

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & \cdots & a_{1r} & \cdots & a_{1n} & b_1 \\ 0 & a'_{22} & \cdots & a'_{2r} & \cdots & a'_{2n} & b'_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{rr}^{(r-1)} & \cdots & a_{rn}^{(r-1)} & b_r^{(r-1)} \end{array}\right),$$

где $a_{11} \neq 0,\ a'_{22} \neq 0,\ \ldots,\ a_{rr}^{(r-1)} \neq 0$. Соответствующая система имеет трапециевидную форму:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r + \ldots + a_{1n}x_n = b_1, \\ a'_{22}x_2 + \ldots + a'_{2r}x_r + \ldots + a'_{2n}x_n = b'_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{rr}^{(r-1)}x_r + \ldots + a_{rn}^{(r-1)}x_n = b_r^{(r-1)}. \end{cases} \tag{4.5}$$

В этом случае объявляем неизвестные $x_{r+1}, \ldots, x_n$ **свободными** и переносим в правые части уравнений.

Пример **4.3.** Решить систему

$$\begin{cases} x_1 + x_2 - 2x_3 + x_4 = 1, \\ 2x_1 + 3x_2 - 3x_3 = 5, \\ x_1 + 4x_2 + 2x_3 - 6x_4 = 11, \\ x_1 - x_2 - 5x_3 + 6x_4 = -6. \end{cases}$$

Решение. Преобразуем расширенную матрицу системы:

$$\left(\begin{array}{cccc|c} 1 & 1 & -2 & 1 & 1 \\ 2 & 3 & -3 & 0 & 5 \\ 1 & 4 & 2 & -6 & 11 \\ 1 & -1 & -5 & 6 & -6 \end{array}\right) \xrightarrow{l_2-2l_1,\, l_3-l_1,\, l_4-l_1} \left(\begin{array}{cccc|c} 1 & 1 & -2 & 1 & 1 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 3 & 4 & -7 & 10 \\ 0 & -2 & -3 & 5 & -7 \end{array}\right) \to$$

$$\xrightarrow{l_3-3l_2,\, l_4+2l_2} \left(\begin{array}{cccc|c} 1 & 1 & -2 & 1 & 1 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 1 & -1 \end{array}\right) \xrightarrow{l_4+l_3} \left(\begin{array}{cccc|c} 1 & 1 & -2 & 1 & 1 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Получаем систему

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

Неизвестные x_4 объявляем свободными: $x_4 = c$. Из последнего уравнения получаем: $x_3 = 1 + c$. Подставляем это значение во второе уравнение, получаем $x_2 = 2 + c$. Подставляя найденные x_3 и x_2 в первое уравнение, получаем $x_1 = 1$.

Система имеет бесконечное множество решений:

$$\bar{x} = (1,\ 2 + c,\ 1 + c,\ c),$$

где c принимает любые числовые значения. Это общее решение системы.

Рассмотрим теперь **метод Гаусса** в несколько ином виде. Метод состоит из нескольких шагов. Предположим, что первые $k - 1$ шагов сделаны, и дадим описание очередного k-го шага.

1. Проверяем, есть ли в системе (полученной после предыдущих $k - 1$ шагов) хотя бы одно противоречивое уравнение. Если такое уравнение в системе есть, то она несовместна – работа с ней прекращается.

2. Если в системе есть *тривиальные* уравнения $0 = 0$, удаляем их.

3. Пусть противоречивых уравнений в системе нет. Тогда одно из них выбирается в качестве *разрешающего уравнения*, а одно из неизвестных объявляется *разрешающим неизвестным*. При этом должны быть выполнены следующие условия:

1) на предыдущих шагах это уравнение не было разрешающим;

2) в разрешающем уравнении коэффициент при разрешающем неизвестном должен быть отличен от нуля* (этот коэффициент называют иногда *разрешающим элементом*);

3) из всех уравнений, кроме разрешающего, исключаем разрешающее неизвестное. Для этого к каждому из этих уравнений прибавляем разрешающее уравнение, умноженное на соответствующее число.

Через конечное число шагов процесс остановится, и будет либо установлена несовместность системы, либо получено общее решение этой системы. Это произойдет, когда все уравнения побывают в роли разрешающих.

Рассмотрим примеры. Как обычно, будем преобразовывать не сами системы, а их расширенные матрицы.

***Пример* 4.4.** Найти общее решение и одно частное решение системы уравнений

$$\begin{cases} x_1 + x_2 - 4x_3 - x_4 + 3x_6 = -1, \\ 2x_1 + x_2 - 5x_3 + x_6 = 7, \\ 3x_1 + 4x_3 + x_5 - 3x_6 = -2, \\ x_1 - x_3 + x_4 - 2x_6 = 8. \end{cases}$$

Решение. Запишем расширенную матрицу системы:

$$\left(\begin{array}{cccccc|c} 1 & 1 & -4 & -1 & 0 & 3 & -1 \\ 2 & 1 & -5 & 0 & 0 & 1 & 7 \\ 3 & 0 & 4 & 0 & 1 & -3 & -2 \\ 1 & 0 & -1 & 1 & 0 & -2 & 8 \end{array}\right).$$

---

* С помощью элементарных преобразований можно сделать разрешающее уравнение таким, чтобы коэффициент при разрешающем неизвестном стал равным единице.

Шаг 1. Убеждаемся, что данная система не содержит противоречивых и тривиальных уравнений. Выбираем первое уравнение в качестве разрешающего уравнения, а коэффициент $a_{12} = 1$ – в качестве разрешающего элемента. Совершаем преобразование $l_2 - l_1$:

$$\left(\begin{array}{cccccc|c} 1 & 1 & -4 & -1 & 0 & 3 & -1 \\ 1 & 0 & -1 & 1 & 0 & -2 & 8 \\ 3 & 0 & 4 & 0 & 1 & -3 & -2 \\ 1 & 0 & -1 & 1 & 0 & -2 & 8 \end{array}\right).$$

Шаг 2. Полученная после первого шага матрица есть расширенная матрица системы, не содержащей противоречивых и тривиальных уравнений. Берем в качестве разрешающего уравнения второе уравнение, а в качестве разрешающего элемента – коэффициент $a_{24} = 1$. Совершаем преобразования $l_1 + l_2$, $l_4 - l_2$:

$$\left(\begin{array}{cccccc|c} 2 & 1 & -5 & 0 & 0 & 1 & 7 \\ 1 & 0 & -1 & 1 & 0 & -2 & 8 \\ 3 & 0 & 4 & 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Шаг 3. Система уравнений, полученная после второго шага, содержит тривиальное уравнение – удаляем его (удаляем строку, состоящую из нулей):

$$\left(\begin{array}{cccccc|c} 2 & 1 & -5 & 0 & 0 & 1 & 7 \\ 1 & 0 & -1 & 1 & 0 & -2 & 8 \\ 3 & 0 & 4 & 0 & 1 & -3 & -2 \end{array}\right).$$

В третьем уравнении коэффициент $a_{35} = 1$. Его можно взять в качестве разрешающего, однако остальные уравнения уже не содержат разрешающего неизвестного x_5. Получаем систему

$$\begin{cases} 2x_1 + x_2 - 5x_3 + x_6 = 7, \\ x_1 - x_3 + x_4 - 2x_6 = 8, \\ 3x_1 + 4x_3 + x_5 - 3x_6 = -2, \end{cases}$$

у которой неизвестные x_2, x_4, x_5 выражаются через свободные неизвестные x_1, x_3, x_6:

$$\begin{cases} x_2 = 7 - 2x_1 + 5x_3 - x_6, \\ x_4 = 8 - x_1 + x_3 + 2x_6, \\ x_5 = -2 - 3x_1 - 4x_3 + 3x_6. \end{cases}$$

Полагая $x_1 = c_1$, $x_3 = c_2$, $x_6 = c_3$, получаем общее решение

$$\overline{x} = (c_1,\ 7 - 2c_1 + 5c_2 - c_3,\ c_2,\ 8 - c_1 + c_2 + 2c_3,\ -2 - 3c_1 - 4c_2 + 3c_3,\ c_3).$$

Взяв, например, $c_1 = 1$, $c_2 = 0$, $c_3 = 3$, получим одно из частных решений:

$$x_1 = 1,\ x_2 = 2,\ x_3 = 0,\ x_4 = 13,\ x_5 = 4,\ x_6 = 3.$$

***Пример* 4.5.** Найти общее решение системы

$$\begin{cases} x_1 + 2x_2 + x_3 + 4x_4 + x_5 = -4, \\ 2x_1 + 5x_2 + 3x_3 + 8x_4 + 3x_5 = -10, \\ x_2 + 2x_3 + 2x_4 + 6x_5 = -1, \\ 4x_1 + 4x_2 + 16x_4 = -3. \end{cases}$$

Решение:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 1 & -4 \\ 2 & 5 & 3 & 8 & 3 & -10 \\ 0 & 1 & 2 & 2 & 6 & -1 \\ 4 & 4 & 0 & 16 & 0 & -3 \end{array}\right) \xrightarrow{l_2 - 2l_1,\ l_4 - 4l_1} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 4 & 1 & -4 \\ 0 & 1 & 1 & 0 & 1 & -2 \\ 0 & 1 & 2 & 2 & 6 & -1 \\ 0 & -4 & -4 & 0 & -4 & 13 \end{array}\right) \to$$

$$\xrightarrow{l_1 - 2l_2,\ l_3 - l_2,\ l_4 + 4l_2} \left(\begin{array}{ccccc|c} 1 & 0 & -1 & 4 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 & -2 \\ 0 & 0 & 1 & 2 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right).$$

Система несовместна.

2. Метод обратной матрицы. Пусть в системе уравнений (4.1) число уравнений равно числу неизвестных: $m = n$. Система имеет вид

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n. \end{cases} \tag{4.6}$$

Запишем систему (4.6) в матричной форме:

$$AX = B. \tag{4.7}$$

Пусть матрица A является невырожденной: $|A| \neq 0$, тогда существует обратная матрица A^{-1}.

Умножим обе части уравнения (4.7) на A^{-1} слева:

$$A^{-1}AX = A^{-1}B,$$

получаем решение системы (4.6)

$$X = A^{-1}B. \tag{4.8}$$

***Пример* 4.6.** Решить системы уравнений:

$$1)\begin{cases} x_1+2x_2+3x_3=4, \\ 2x_1+2x_2+3x_3=5, \\ 3x_1+3x_2+4x_3=7; \end{cases} \quad 2)\begin{cases} x_1+2x_2+3x_3=0, \\ 2x_1+2x_2+3x_3=1, \\ 3x_1+3x_2+4x_3=1; \end{cases} \quad 3)\begin{cases} x_1+2x_2+3x_3=14, \\ 2x_1+2x_2+3x_3=15, \\ 3x_1+3x_2+4x_3=21. \end{cases}$$

Р е ш е н и е. Матрица A^{-1} найдена в примере 2.7.

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

2. $X = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -5 & 3 \\ 0 & 3 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$

3. $X = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -5 & 3 \\ 0 & 3 & -2 \end{pmatrix} \begin{pmatrix} 14 \\ 15 \\ 21 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$

3. Правило Крамера. В § 3.1 мы уже рассматривали правило Крамера для решения системы двух уравнений с двумя неизвестными. Обобщим его на случай любого числа неизвестных.

***Теорема* 4.1 (теорема Крамера).** Пусть дана система n линейных уравнений с n неизвестными $AX = B$. Если $|A| \neq 0$, то система имеет единственное решение:

$$x_1 = \frac{|A_1|}{|A|}, \quad x_2 = \frac{|A_2|}{|A|}, \quad \ldots, \quad x_n = \frac{|A_n|}{|A|}, \tag{4.9}$$

где A_i – матрица, полученная из A заменой ее i-го столбца столбцом свободных членов B ($i = 1, 2, \ldots, n$).

Доказательство. Запишем в развернутом виде с учетом (3.10) решение $X=A^{-1}B$ системы $AX=B$:

$$\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} = \frac{1}{|A|} \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix}.$$

В соответствии с правилом умножения матриц получаем

$$x_i = \frac{1}{|A|}(A_{1i}b_1 + A_{2i}b_2 + \ldots + A_{ni}b_n), \quad i = 1,\ 2,\ \ldots,\ n.$$

Однако выражение, стоящее в скобках, есть разложение определителя $|A_i|$ по i-му столбцу: $A_{1i}b_1 + A_{2i}b_2 + \ldots + A_{ni}b_n = |A_i|$.

Итак, $x_i = \dfrac{|A_i|}{|A|}$. Теорема доказана.

Формулы (4.9) называются **формулами Крамера**, а правило решения систем с помощью этих формул – **правилом Крамера**.

Формулы Крамера имеют в основном теоретическое значение. Их применение для решения систем с бóльшим числом неизвестных привело бы к очень громоздким вычислениям. Однако эти формулы имеют весьма важное достоинство: они дают *явное* выражение значений всех неизвестных.

4.3. СОВМЕСТНОСТЬ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ

Рассмотрим матрицу A и расширенную матрицу $\overline{A}$ системы (4.1). Известно, что ранг матрицы равен наибольшему числу ее линейно независимых столбцов. Потому, присоединяя к матрице A столбец свободных членов, либо получаем матрицу, ранг которой на единицу больше ранга A, либо не увеличиваем ранг – если столбец свободных членов есть линейная комбинация остальных столбцов:

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{pmatrix} k_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{pmatrix} k_2 + \ldots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{pmatrix} k_n = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix}. \tag{4.10}$$

Легко видеть, что равенство (4.10) равносильно тому, что система (4.1) имеет решение $x_1 = k_1, x_2 = k_2, \ldots, x_n = k_n$. Вопрос о совместности системы линейных уравнений решается следующей теоремой.

Теорема **4.2 (теорема Кронекера – Капелли).** Система линейных уравнений (4.1) совместна тогда и только тогда, когда ранг ее расширенной матрицы $\overline{A}$ равен рангу ее матрицы A.

Доказательство. 1. Пусть система (4.1) совместна и пусть $k_1, k_2, ..., k_n$ – ее решение. Подставим эти числа вместо неизвестных и получим систему тождеств, которая эквивалентна равенству (4.10), а оно означает, что столбец свободных членов есть линейная комбинация столбцов матрицы A. Отсюда следует, что $\mathrm{rg}\, A = \mathrm{rg}\, \overline{A}$.

2. Пусть теперь дано, что $\mathrm{rg}\, A = \mathrm{rg}\, \overline{A}$. Тогда любая максимальная линейно независимая система столбцов матрицы A остается максимальной линейно независимой системой столбцов и в матрице $\overline{A}$. Следовательно, в частности, столбец свободных членов есть линейная комбинация столбцов этой системы, а отсюда следует, что столбец свободных членов есть линейная комбинация всех столбцов матрицы A, т.е. имеет место равенство вида (4.10). Это, в свою очередь, означает, что числа $k_1, k_2, ..., k_n$ (среди которых, разумеется, некоторые могут быть нулями) составляют решение системы (4.1). Мы доказали, что из равенства $\mathrm{rg}\, A = \mathrm{rg}\, \overline{A}$ следует совместность системы (4.1). Доказательство закончено.

Рассмотрим теперь применение теоремы Кронекера – Капелли. Пусть $\mathrm{rg}\, A = \mathrm{rg}\, \overline{A} = r$. В этом случае будем говорить, что ранг системы (4.1) равен r. Тогда система (4.1) совместна.

Если $r = n$, то система определенна. Ее единственное решение может быть вычислено либо по правилу Крамера, либо путем приведения к виду (4.4).

Если $r < n$, система приводится к виду (4.5). Неизвестные $x_{r+1}, ..., x_n$ объявляем свободными и переносим в правые части уравнений. В левых частях остаются слагаемые, содержащие неизвестные $x_1, x_2, ..., x_r$, называемые **базисными**:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r = b_1 - a_{1,r+1}x_{r+1} - \ldots - a_{1n}x_n, \\ \qquad a'_{22}x_2 + \ldots + a'_{2r}x_r = b'_2 - a_{2,r+1}x_{r+1} - \ldots - a'_{2n}x_n, \\ \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \qquad\qquad a_{rr}^{(r-1)}x_r = b_r^{(r-1)} - a_{r,r+1}^{(r-1)}x_{r+1} - \ldots - a_{rn}^{(r-1)}x_n. \end{cases}$$

Из последнего уравнения системы находим x_r, подставляем его значение (зависящее от свободных неизвестных) в предпоследнее уравнение, находим x_{r-1} и т.д. Таким образом найдем неизвестные $x_1, x_2, \ldots, x_r$, выраженные через свободные неизвестные, которые могут принимать любые значения. Поэтому система (4.1) в данном случае имеет бесконечное множество решений.

Придавая свободным неизвестным произвольные значения и получая соответствующие значения базисных неизвестных, получим все решения системы (4.1).

Пример **4.7.** Решить систему уравнений

$$\begin{cases} x_1 + 2x_2 \qquad + \ \ x_4 = 5, \\ 3x_1 + 5x_2 - x_3 + 2x_4 = 13, \\ x_1 + 3x_2 + x_3 + 2x_4 = 7, \\ x_1 + \ \ x_2 - x_3 \qquad\quad = 3. \end{cases}$$

Решение. Составим расширенную матрицу и применим метод Гаусса:

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 5 \\ 3 & 5 & -1 & 2 & 13 \\ 1 & 3 & 1 & 2 & 7 \\ 1 & 1 & -1 & 0 & 3 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 5 \\ 0 & -1 & -1 & -1 & -2 \\ 0 & 1 & 1 & 1 & 2 \\ 0 & -1 & -1 & -1 & -2 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 5 \\ 0 & -1 & -1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Видим, что $\mathrm{rg}\, A = \mathrm{rg}\, \overline{A} = 2$, так как, в частности, минор, составленный из коэффициентов при x_1 и x_2 в первых двух уравнениях, отличен от нуля:

$$\begin{vmatrix} 1 & 2 \\ 0 & -1 \end{vmatrix} = -1 \neq 0,$$

а все миноры третьего порядка равны нулю. Неизвестные x_1, x_2 берем в качестве базисных, а остальные, т.е. x_3, x_4, объявляем свободными и переносим направо:

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

Придавая свободным переменным произвольные значения $x_3 = c_1$, $x_4 = c_2$, находим бесконечное множество решений системы:

$$x_1 = 1 + 2c_1 + c_2, \quad x_2 = 2 - c_1 - c_2, \quad x_3 = c_1, \quad x_4 = c_2.$$

4.4. ОДНОРОДНЫЕ СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Система линейных уравнений называется **однородной**, если во всех ее уравнениях свободные члены равны нулю:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases} \tag{4.11}$$

Очевидно, такая система всегда имеет нулевое (или тривиальное) решение: $x_1 = 0, x_2 = 0, \ldots, x_n = 0$. Поэтому однородная система всегда совместна. (Очевидно также, что $\mathrm{rg}\, A = \mathrm{rg}\, \overline{A}$.)

Интерес представляет вопрос о существовании ненулевых решений этой системы. Если $m = n$ и определитель системы (4.11) отличен от нуля, то система имеет только нулевое решение (это следует из теоремы Крамера).

Справедливо следующее утверждение (приведем его без доказательства).

***Теорема* 4.3.** Однородная система имеет ненулевое решение тогда и только тогда, когда ранг этой системы меньше числа неизвестных.

В частности, при $m = n$ система имеет ненулевое решение тогда и только тогда, когда ее определитель равен нулю.

Будем записывать всякое решение системы (4.11): $x_1 = c_1$, $x_2 = c_2$, ..., $x_n = c_n$ в виде вектора-строки $\overline{c} = (c_1, c_2, \ldots, c_n)$.

Отметим **свойства решений однородных систем**:

1. Если вектор $\overline{c} = (c_1, c_2, \ldots, c_n)$ – решение системы (4.11), то для любого числа k вектор $k\overline{c} = (kc_1, kc_2, \ldots, kc_n)$ также есть решение этой системы.

2. Если векторы $\overline{c} = (c_1, c_2, \ldots, c_n)$ и $\overline{c}' = (c_1', c_2', \ldots, c_n')$ являются решениями системы (4.11), то их сумма $\overline{c} + \overline{c}'$ также есть решение этой системы.

Справедливость этих свойств проверяется непосредственной подстановкой указанных решений в уравнения системы. (Предлагаем читателю проделать это самостоятельно.)

Из сформулированных свойств следует, что *всякая линейная комбинация решений однородной системы также является решением этой системы*.

Очевидно, если однородная система имеет ненулевое решение, то она имеет бесконечное множество решений. Из множества векторов-решений однородной системы (4.11) можно выбрать базис. Этот базис называется **фундаментальной системой решений однородной системы** (4.11). Система (4.11) обладает в этом случае многими различными фундаментальными системами решений.

Теорема **4.4.** Если ранг r системы линейных однородных уравнений (4.11) меньше числа неизвестных n, то всякая фундаментальная система решений системы (4.11) состоит из $n - r$ решений.

(Теорему 4.4 примем без доказательства.)

Укажем **способ нахождения фундаментальных систем решений системы** (4.11).

Надо взять любую систему из $n - r$ линейно независимых $(n - r)$-мерных векторов, принять компоненты каждого из этих векторов за значения свободных неизвестных $x_{r+1}, ..., x_n$ и найти соответствующие значения для r базисных неизвестных. Получим $n - r$ решений системы уравнений (4.11), составляющих фундаментальную систему.

Обычно удобнее всего в качестве векторов значений для свободных неизвестных брать $(n - r)$-мерные единичные векторы: (1, 0, ..., 0), (0, 1, ..., 0), ..., (0, 0, ..., 1).

Пример **4.8.** Решить систему

$$\begin{cases} x_1+2x_2+x_3+x_4-3x_5=0, \\ 2x_1+3x_2+x_3+2x_4+x_5=0, \\ 2x_1+5x_2+3x_3+2x_4-13x_5=0, \\ x_1+x_2+x_4+4x_5=0, \\ x_1-x_2-2x_3+x_4+18x_5=0. \end{cases}$$

Решение. Преобразуем матрицу системы:

$$\begin{pmatrix} 1 & 2 & 1 & 1 & -3 \\ 2 & 3 & 1 & 2 & 1 \\ 2 & 5 & 3 & 2 & -13 \\ 1 & 1 & 0 & 1 & 4 \\ 1 & -1 & -2 & 1 & 18 \end{pmatrix} \xrightarrow[l_4-l_1,\, l_5-l_1]{l_2-2l_1,\, l_3-2l_1,} \begin{pmatrix} 1 & 2 & 1 & 1 & -3 \\ 0 & -1 & -1 & 0 & 7 \\ 0 & 1 & 1 & 0 & -7 \\ 0 & -1 & -1 & 0 & 7 \\ 0 & -3 & -3 & 0 & 21 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 2 & 1 & 1 & -3 \\ 0 & -1 & -1 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Получаем систему, эквивалентную исходной:

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 - 3x_5 = 0, \\ \quad -x_2 - x_3 \qquad + 7x_5 = 0. \end{cases}$$

Выберем в качестве базисных неизвестных x_1 и x_2 (коэффициенты при них образуют минор, отличный от нуля):

$$\begin{cases} x_1 + 2x_2 = -x_3 - x_4 + 3x_5, \\ \quad -x_2 = \ \ x_3 \qquad - 7x_5. \end{cases}$$

Получаем фундаментальную систему решений:

$\overline{e}_1 = (1,\ -1,\ 1,\ 0,\ 0),\quad \overline{e}_2 = (-1,\ 0,\ 0,\ 1,\ 0),\quad \overline{e}_3 = (-11,\ 7,\ 0,\ 0,\ 1).$

4.5. НЕОДНОРОДНЫЕ СИСТЕМЫ. СТРУКТУРА ОБЩЕГО РЕШЕНИЯ СИСТЕМЫ ЛИНЕЙНЫХ НЕОДНОРОДНЫХ УРАВНЕНИЙ

Рассмотрим систему линейных неоднородных уравнений (4.1)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases}$$

и соответствующую ей однородную систему (4.11)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0, \end{cases}$$

полученную из системы (4.1) заменой свободных членов нулями.

Установим связь между решениями систем (4.1) и (4.11). Справедливы следующие ***у т в е р ж д е н и я***:

1. Сумма любого решения системы (4.1) с любым решением соответствующей однородной системы (4.11) снова есть решение системы (4.1).

2. Разность любых двух решений системы (4.1) есть решение соответствующей однородной системы (4.11).

Доказательства этих утверждений очень просты. Докажем, например, второе.

Пусть $\overline{c} = (c_1, c_2, ..., c_n)$ и $\overline{c}' = (c_1', c_2', ..., c_n')$ — два решения системы (4.1). Подставим разность этих решений $\overline{c} - \overline{c}' = (c_1 - c_1', c_2 - c_2', ..., c_n - c_n')$ в левую часть любого i-го уравнения системы (4.11) и перегруппируем члены:

$$a_{i1}(c_1 - c_1') + a_{i2}(c_2 - c_2') + ... + a_{in}(c_n - c_n') =$$
$$= (a_{i1}c_1 + a_{i2}c_2 + ... + a_{in}c_n) - (a_{i1}c_1' + a_{i2}c_2' + ... + a_{in}c_n') =$$
$$= b_i - b_i = 0.$$

Уравнение превращается в тождество $0 = 0$.

Из этих двух утверждений следует, что, *найдя одно решение системы линейных неоднородных уравнений (4.1) и сложив его с каждым из решений соответствующей однородной системы (4.11), получим все решения системы (4.1)*.

Иначе говоря, *общее решение системы линейных уравнений (4.1) есть сумма любого частного решения этой системы с общим решением соответствующей однородной системы (4.11)*.

***Пример* 4.9.** Решить систему

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 - 3x_5 = 4, \\ 2x_1 + 3x_2 + x_3 + 2x_4 + x_5 = 6, \\ 2x_1 + 5x_2 + 3x_3 + 2x_4 - 13x_5 = 10, \\ x_1 + x_2 + x_4 + 4x_5 = 2, \\ x_1 - x_2 - 2x_3 + x_4 + 18x_5 = -2. \end{cases}$$

Решение. Соответствующая данной системе однородная система рассмотрена в примере 4.8. Преобразовав расширенную матрицу данной системы, придем к системе

$$\begin{cases} x_1 + 2x_2 = 4 - x_3 - x_4 + 3x_5, \\ -x_2 = -2 + x_3 - 7x_5. \end{cases}$$

Найдем одно из частных решений этой системы. Проще всего это сделать, положив $x_3 = x_4 = x_5 = 0$. Получаем решение $c_0 = (0, 2, 0, 0, 0)$. (Частное решение, в котором все значения свободных неизвестных равны нулю, иногда называют **базисным**.)

Фундаментальная система решений соответствующей однородной системы уже найдена в примере 4.8. Воспользуемся ею и получим общее решение данной неоднородной системы:

$$(0, 2, 0, 0, 0) + k_1(1, -1, 1, 0, 0) + k_2(-1, 0, 0, 1, 0) + k_3(-11, 7, 0, 0, 1).$$

Придавая коэффициентам k_1, k_2, k_3 все возможные значения, получим все решения данной системы.

Вопросы

1. Может ли неопределенная система линейных уравнений быть несовместной?
2. Что называется общим решением системы линейных уравнений?
3. Может ли система, содержащая семь уравнений с пятью неизвестными, быть эквивалентной системе четырех уравнений с пятью неизвестными?
4. К какой системе линейных уравнений применимо правило Крамера?
5. Применим ли метод обратной матрицы к неопределенной системе линейных уравнений?
6. Может ли однородная система линейных уравнений быть несовместной?
7. Что называется фундаментальной системой решений однородной системы линейных уравнений?
8. Сколько решений содержит фундаментальная система решений однородной системы уравнений с шестью неизвестными, имеющая ранг 4?
9. Какова структура общего решения системы линейных неоднородных уравнений?

Глава 5
ЛИНЕЙНЫЕ ОПЕРАТОРЫ

5.1. ПОНЯТИЕ ЛИНЕЙНОГО ОПЕРАТОРА

Пусть даны два линейных пространства U и V.

Определение. Если задано некоторое правило A, по которому каждому вектору $\bar{u}$ пространства U ставится в соответствие единственный вектор $\bar{v} = A(\bar{u})$ пространства V, то говорят, что задан **оператор** A, действующий из U в V. Вектор $\bar{v} = A(\bar{u})$ называется **образом** вектора $\bar{u}$, а вектор $\bar{u}$ – **прообразом** вектора $\bar{v}$.

Оператор $A(\bar{u})$ называется **линейным**, если он удовлетворяет следующим двум условиям:

1) для любых двух векторов $\bar{u}_1$ и $\bar{u}_2$ пространства U

$$A(\bar{u}_1 + \bar{u}_2) = A(\bar{u}_1) + A(\bar{u}_2);$$

2) для любого вектора $\bar{u}$ из U и любого числа λ

$$A(\lambda\bar{u}) = \lambda A(\bar{u}).$$

Понятие линейного оператора является одним из фундаментальных понятий линейной алгебры.

Пусть теперь $U = \mathbf{R}^n$, $V = \mathbf{R}^m$.

Далее мы увидим, что если в пространстве $\mathbf{R}^n$ задать некоторый базис $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$, а в пространстве $\mathbf{R}^m$ – некоторый базис $\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_m$, то линейный оператор A задается матрицей размера $m \times n$.

Итак, пусть $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$ – какой-нибудь базис пространства $\mathbf{R}^n$. Возьмем произвольный вектор $\bar{x}$ и разложим его по базису:

$$\bar{x} = x_1\bar{e}_1 + x_2\bar{e}_2 + \ldots + x_n\bar{e}_n.$$

Оператор $A(\bar{x})$ – линейный, поэтому

$$A(\bar{x}) = x_1 A(\bar{e}_1) + x_2 A(\bar{e}_2) + \ldots + x_n A(\bar{e}_n). \qquad (5.1)$$

Однако каждый из векторов $A(\overline{e}_i)$ $(i=1, 2, \ldots, n)$ есть вектор из $\mathbf{R}^m$, следовательно, его можно разложить по базису $\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_m$:

$$A(\overline{e}_i) = a_{1i}\overline{f}_1 + a_{2i}\overline{f}_2 + \ldots + a_{mi}\overline{f}_m. \tag{5.2}$$

Подставляя разложение $A(\overline{e}_i)$ $(i=1, 2, \ldots, n)$ из (5.2) в (5.1), получаем

$$A(\overline{x}) = x_1(a_{11}\overline{f}_1 + a_{21}\overline{f}_2 + \ldots + a_{m1}\overline{f}_m) + x_2(a_{12}\overline{f}_1 + a_{22}\overline{f}_2 + \ldots + a_{m2}\overline{f}_m) + \\ + \ldots + x_n(a_{1n}\overline{f}_1 + a_{2n}\overline{f}_2 + \ldots + a_{mn}\overline{f}_m).$$

Группируя члены и собирая коэффициенты при $\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_m$, получаем

$$A(\overline{x}) = (a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n)\overline{f}_1 + (a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n)\overline{f}_2 + \\ + \ldots + (a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n)\overline{f}_m. \tag{5.3}$$

Пусть $y_1, y_2, \ldots, y_m$ – координаты образа вектора $\overline{x}$, т.е. координаты вектора $\overline{y} = A(\overline{x})$ в базисе $\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_m$. Тогда

$$A(\overline{x}) = y_1\overline{f}_1 + y_2\overline{f}_2 + \ldots + y_m\overline{f}_m. \tag{5.4}$$

В силу единственности разложения вектора по базису правые части равенств (5.3) и (5.4) совпадают. Следовательно,

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n. \end{cases} \tag{5.5}$$

Матрица **A** системы (5.5) называется **матрицей оператора** A относительно базисов $\overline{e}_1, \overline{e}_2, \ldots, \overline{e}_n$ и $\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_m$:

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

Итак, *каждому линейному оператору* A: $\mathbf{R}^n \to \mathbf{R}^m$ *соответствует матрица размера* $m \times n$. Очевидно, верно и обратное: *каждой матрице размера* $m \times n$ *соответствует линейный оператор* A: $\mathbf{R}^n \to \mathbf{R}^m$.

Если рассматривать векторы $\overline{x} = (x_1, x_2, ..., x_n)$ и $\overline{y} = A(\overline{x}) = (y_1, y_2, \ldots, y_m)$ как матрицы-столбцы:

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{pmatrix},$$

то равенство $\overline{y} = A(\overline{x})$, или, что то же самое, систему (5.5) можно записать в виде матричного равенства:

$$\mathbf{Y} = \mathbf{AX},$$

где $\mathbf{A}$ – матрица линейного оператора. В частности, когда пространства $\mathbf{R}^n$ и $\mathbf{R}^m$ совпадают, то посредством оператора A пространство $\mathbf{R}^n$ отображается в себя. В этом случае матрица оператора – квадратная матрица порядка n.

5.2. ДЕЙСТВИЯ С ЛИНЕЙНЫМИ ОПЕРАТОРАМИ

Для линейных операторов определены операции сложения и умножения на число.

1. **Суммой двух линейных операторов** A_1 и A_2 называется оператор $(A_1 + A_2)$, определяемый равенством

$$(A_1 + A_2)(\overline{x}) = A_1(\overline{x}) + A_2(\overline{x}).$$

2. **Произведением линейного оператора** A **на число** λ называется оператор λA, определяемый равенством

$$\lambda A(\overline{x}) = \lambda(A(\overline{x})).$$

Известно, что всякий линейный оператор определяется соответствующей квадратной матрицей. Поэтому описанным операциям соответствуют аналогичные операции над матрицами операторов – сложение и умножение на число. Операторы $(A_1 + A_2)$ и λA также являются линейными.

Нулевой оператор $\tilde{0}$ определяется как оператор, переводящий всякий вектор пространства $\mathbf{R}^n$ в нулевой вектор $\overline{0}$.

Очевидно, $(A + \tilde{0}) = A$ для любого оператора A.

Для линейных операторов можно определить также операцию умножения.

3. **Произведением линейных операторов** A_1 и A_2 называется оператор (A_1A_2), определяемый равенством

$$(A_1A_2)(\bar{x}) = A_1(A_2(\bar{x})).$$

Произведение двух линейных операторов также есть линейный оператор.

Тождественный оператор E определяется так:

$$E(\bar{x}) = \bar{x}.$$

Очевидно, $(AE) = (EA) = A$ для любого линейного оператора A.

5.3. СОБСТВЕННЫЕ ВЕКТОРЫ И СОБСТВЕННЫЕ ЗНАЧЕНИЯ ЛИНЕЙНОГО ОПЕРАТОРА

Определение. Ненулевой вектор $\bar{x} \in \mathbf{R}^n$ называется **собственным вектором** линейного оператора A, если существует такое число λ, что

$$A(\bar{x}) = \lambda \bar{x}. \qquad (5.6)$$

При этом число λ называется **собственным значением** оператора A.

Если $\mathbf{A}$ – матрица оператора A, то число λ, удовлетворяющее равенству (5.6), называется *собственным значением* матрицы $\mathbf{A}$, а вектор $\bar{x}$ – *собственным вектором* матрицы $\mathbf{A}$.

Равенство (5.6) можно записать в матричной форме:

$$\mathbf{AX} = \lambda\mathbf{X}, \quad \text{или} \quad \mathbf{AX} = \lambda\mathbf{EX}. \qquad (5.7)$$

Из последнего равенства получаем

$$(\mathbf{A} - \lambda\mathbf{E})\mathbf{X} = 0. \qquad (5.8)$$

Если $\mathbf{A} = (a_{ij}), \quad i, j = 1, 2, ..., n$, то

$$\mathbf{A} - \lambda\mathbf{E} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix}. \qquad (5.9)$$

Уравнение (5.8) эквивалентно системе линейных однородных уравнений

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0, \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \ldots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + (a_{nn} - \lambda)x_n = 0. \end{cases} \qquad (5.10)$$

Так как собственный вектор не является нулевым, то однородная система (5.10) должна иметь ненулевое решение. Как известно, однородная система n уравнений с n неизвестными имеет ненулевое решение тогда и только тогда, когда ее определитель равен нулю. Поэтому

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0, \quad \text{или} \quad \left|\mathbf{A}-\lambda\mathbf{E}\right| = 0. \qquad (5.11)$$

Уравнение (5.11) называется **характеристическим уравнением матрицы**, а левая часть этого уравнения – **характеристическим многочленом** матрицы **A** (или оператора A).

Очевидно, что уравнение (5.11) есть алгебраическое уравнение степени n. Его корни являются собственными значениями матрицы **A**. Подставляя каждый из корней в систему (5.10) и решая ее, получим соответствующий собственный вектор.

***Пример* 5.1.** Найти собственные значения и собственные векторы линейного оператора A, заданного матрицей

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

Решение. Составляем характеристическое уравнение $\left|\mathbf{A}-\lambda\mathbf{E}\right| = 0$ для этой матрицы:

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

Раскрывая определитель, получаем

$$\lambda^2 - 7\lambda + 10 = 0.$$

Корни этого уравнения: $\lambda_1 = 5$, $\lambda_2 = 2$. Подставляя $\lambda_1 = 5$ в систему (5.10) при $n = 2$, получаем

$$\begin{cases} (4-5)x_1 + x_2 = 0, \\ 2x_1 + (3-5)x_2 = 0, \end{cases} \quad \text{т.е.} \quad \begin{cases} -x_1 + x_2 = 0, \\ 2x_1 - 2x_2 = 0. \end{cases} \qquad (*)$$

Собственный вектор, соответствующий собственному значению $\lambda_1 = 5$, является решением этой системы, которая, в свою очередь, равносильна уравнению

$$x_1 - x_2 = 0.$$

Объявляя x_2 свободным неизвестным и полагая $x_2 = c$, получаем первый собственный вектор: $\bar{x}_1 = (c,\ c) = c(1,\ 1)$.

Подставим теперь $\lambda_2 = 2$:

$$\begin{cases} 2x_1 + x_2 = 0, \\ 2x_1 + x_2 = 0. \end{cases} \qquad (**)$$

Эта система также эквивалентна одному уравнению. Полагая $x_2 = d$, получим $\bar{x}_2 = \left(-\dfrac{d}{2},\ d\right) = -\dfrac{d}{2}(1,\ -2)$.

Так как c и d – произвольные числа, то одному собственному значению может соответствовать бесконечное множество векторов. В частности, полагая $c = 1$, $d = -2$, получаем собственные векторы, являющиеся фундаментальными решениями соответствующих однородных систем (*) и (**). Они имеют вид $\bar{x}_1 = (1,\ 1)$, $\bar{x}_2 = (1, -2)$.

Вопросы

1. Чем задается линейный оператор в базисе пространства $\mathbf{R}^n$?
2. Всякая ли квадратная матрица n-го порядка задает в $\mathbf{R}^n$ линейный оператор?
3. Какой линейный оператор называется нулевым?
4. Как выглядит характеристическое уравнение для матрицы

   $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$?
5. Сколько различных собственных значений может иметь матрица третьего порядка?
6. Является ли число $\lambda = 5$ собственным значением матрицы

   $\mathbf{A} = \begin{pmatrix} 5 & 3 \\ 0 & 1 \end{pmatrix}$?
7. Является ли вектор $\bar{x} = (2,\ 3)$ собственным вектором для матрицы

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

Глава 6
КВАДРАТИЧНЫЕ ФОРМЫ

6.1. ОСНОВНЫЕ ПОНЯТИЯ

Определение. **Квадратичной формой** $F(x_1, x_2, ..., x_n)$ от n неизвестных $x_1, x_2, ..., x_n$ называется сумма, каждое слагаемое которой является либо квадратом одной из этих переменных, либо произведением двух разных переменных.

***Пример* 6.1.** Сумма $x^2 - 3xy + 2y$ является квадратичной формой от двух неизвестных: x и y. Сумма $x_1^2 + 2x_1x_2 - 3x_1x_3 + 4x_2^2 - x_2x_3$ является квадратичной формой от трех неизвестных: x_1, x_2, x_3.

(Заметим, что в рассмотренных квадратичных формах уже приведены подобные члены.)

Каждую квадратичную форму можно записать в **стандартном виде**. При этом используется следующая символика.

Сумма $a_k + a_{k+1} + ... + a_n$ записывается в виде

$$a_k + a_{k+1} + \ldots + a_n = \sum_{i=k}^{n} a_i.$$

В частности,

$$a_1 + a_2 + \ldots + a_n = \sum_{i=1}^{n} a_i.$$

Если же рассматривается сумма, слагаемые которой a_{ij} снабжены двумя индексами i и j, причем $i = 1, 2, ..., m$; $j = 1, 2, ..., n$, то можно сначала взять суммы элементов с фиксированным первым индексом, т.е.

$$\sum_{j=1}^{n} a_{1j},\ \sum_{j=1}^{n} a_{2j},\ \ldots,\ \sum_{j=1}^{n} a_{mj},$$

а затем сложить все эти суммы. Тогда для суммы всех элементов a_{ij} получим запись

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}. \qquad (6.1)$$

Можно также складывать сначала слагаемые a_{ij} с фиксированным вторым индексом, а затем уже полученные суммы:

$$\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}. \tag{6.2}$$

Поэтому

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}, \tag{6.3}$$

т.е. в двойной сумме можно менять порядок суммирования.

Суммы (6.1) и (6.3) можно рассматривать как сумму элементов матрицы размера $m \times n$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}.$$

Если в этой матрице сложить элементы каждой строки, а затем сложить полученные суммы, имеем (6.1); если же сначала сложить элементы каждого столбца, а затем сложить то, что получилось, имеем (6.2).

Вернемся теперь к вопросу о стандартном виде квадратичной формы. Считая, что в квадратичной форме $F = F(x_1, x_2, ..., x_n)$ уже приведены подобные члены, введем следующие обозначения: коэффициент квадратичной формы при x_i^2 обозначим через a_{ii}, а коэффициент при произведении $x_i x_j$ для $i \neq j$ – через $2a_{ij}$. Так как, очевидно, $x_i x_j = x_j x_i$, то коэффициент при этом произведении можно было бы обозначить и через $2a_{ji}$, т.е. предполагается, что

$$a_{ij} = a_{ji}. \tag{6.4}$$

Теперь член $2a_{ij}x_i x_j$ можно записать в виде

$$2a_{ij}x_i x_j = a_{ij}x_i x_j + a_{ji}x_j x_i,$$

а вся квадратичная форма $F = F(x_1, x_2, ..., x_n)$ записывается в виде суммы всевозможных членов $a_{ij}x_i x_j$, где i и j независимо друг от друга принимают все значения от 1 до n:

$$F(x_1, x_2, ..., x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j. \tag{6.5}$$

(В частности, если $i = j$, то получается $a_{ii}x_i^2$.) Заметим, что при двойном суммировании нередко обходятся одним знаком суммы. Равенство (6.5) можно записать в виде

$$F(x_1, x_2, ..., x_n) = \sum_{i,j=1}^{n} a_{ij}x_i x_j.$$

Коэффициенты a_{ij} квадратичной формы (6.5) образуют, очевидно, квадратную матрицу $A = (a_{ij})$ порядка n; она называется **матрицей квадратичной формы** $F = F(x_1, x_2, ..., x_n)$, а ранг r матрицы A – **рангом** этой квадратичной формы. Если, в частности, $r = n$, т.е. матрица A – невырожденная, то и квадратичная форма $F = F(x_1, x_2, ..., x_n)$ называется **невырожденной**. Равенство (6.4) означает, что элементы матрицы A, симметричные относительно главной диагонали, равны между собой, т.е. матрица A – симметрическая. Очевидно, для любой симметрической матрицы n-го порядка можно указать вполне определенную квадратичную форму (6.5) от n неизвестных, коэффициентами которых являются элементы матрицы A.

***Пример* 6.2.** Записать квадратичную форму

$$F = F(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 + 2x_2^2 - 3x_2x_3 + x_3^2 + x_3x_2$$

в стандартном виде и найти ее матрицу.

Решение. После приведения подобных членов получим

$$x_1^2 + 4x_1x_2 + 2x_2^2 - 2x_2x_3 + x_3^2 =$$
$$= x_1^2 + 2x_1x_2 + 2x_2x_1 + 2x_2^2 - x_2x_3 - x_3x_2 + x_3^2 =$$
$$= x_1x_1 + 2x_1x_2 + 0 \cdot x_1x_3 + 2x_2x_1 + 2x_2x_2 - x_2x_3 + 0 \cdot x_3x_1 - x_3x_2 + x_3x_3.$$

Матрица квадратичной формы имеет вид

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

Квадратичную форму (6.5) можно записать в матричном (векторно-матричном) виде, используя произведение прямоугольных матриц.

Заметим, что матрица A является симметрической тогда и только тогда, когда она совпадает со своей транспонированной, т.е. когда

$$A' = A.$$

Обозначим через X матрицу-столбец неизвестных:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}.$$

Покажем, что в матричной записи квадратичная форма имеет следующий вид:

$$F = X'AX. \tag{6.6}$$

Действительно, произведение AX будет матрицей-столбцом:

$$AX = \begin{pmatrix} \sum a_{1j}x_j \\ \sum a_{2j}x_j \\ \dots \\ \sum a_{nj}x_j \end{pmatrix}.$$

(Здесь пишем $\sum$ вместо $\sum\limits_{j=1}^{n}$, чтобы избежать излишней громоздкости записи.)

Умножая теперь эту матрицу-столбец слева на матрицу $X' = (x_1, x_2, ..., x_n)$, получаем

$$X'AX = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j = F,$$

что и требовалось доказать.

***Пример* 6.3.** Записать квадратичную форму из примера 6.2 в матричном виде.

Решение. Используя матрицу A, найденную в примере 6.2, получаем

$$F = (x_1, x_2, x_3)\begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Рассмотрим теперь, как меняется квадратичная форма при линейном преобразовании неизвестных. Пусть задано линейное преобразование неизвестных $x_1, x_2, ..., x_n$:

$$x_i = \sum_{k=1}^{n} c_{ik} y_k, \quad i = 1, 2, \dots, n \tag{6.7}$$

с матрицей $C = (c_{ik})$, иначе говоря, задано линейное преобразование

$$X = CY, \tag{6.8}$$

где Y – столбец неизвестных $y_1, y_2, ..., y_n$.

Воспользуемся одним из свойств операции транспонирования матриц:

$$(AB)' = B'A'. \tag{6.9}$$

В соответствии с (6.9) из (6.8) получаем

$$X' = (CY)' = Y'C'. \tag{6.10}$$

Отсюда

$$F = X'AX = (Y'C')A(CY) = Y'(C'AC)Y, \quad \text{или} \quad F = Y'\widetilde{A}Y,$$

где

$$\widetilde{A} = C'AC. \tag{6.11}$$

Матрица $\widetilde{A}$ будет симметрической. Действительно, так как $A' = A$, то с учетом свойства (6.9)

$$\widetilde{A}' = (C'AC)' = C'(C'A)' = C'A'C = C'AC = \widetilde{A}.$$

Итак, мы доказали следующую теорему.

***Теорема* 6.1.** Квадратичная форма с матрицей A в результате линейного преобразования с матрицей C превращается в квадратичную форму от новых неизвестных с матрицей $C'AC$.

Предположим теперь, что матрица преобразования C – невырожденная. Тогда, очевидно, C' – также невырожденная матрица. В этом случае произведение $C'AC$ есть произведение матрицы A на невырожденные матрицы и поэтому ранг этого произведения равен рангу матрицы A.

Мы получили следующую теорему.

***Теорема* 6.2.** Ранг квадратичной формы не меняется при невырожденном линейном преобразовании.

***Пример* 6.4.** Дана квадратичная форма

$$F = F(x_1, x_2) = x_1^2 + 2x_1x_2 - 3x_2^2.$$

Найти квадратичную форму $G(y_1, y_2)$, полученную из данной линейным преобразованием

$$x_1 = 2y_1 - y_2, \quad x_2 = y_1 + y_2.$$

Р е ш е н и е. Запишем матрицу данной квадратичной формы A и матрицу преобразования C:

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

Матрица искомой квадратичной формы $\widetilde{A}$ согласно (6.11) имеет вид

$$\widetilde{A} = C'AC = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -4 \\ -4 & -4 \end{pmatrix}.$$

Следовательно, $G = G(y_1, y_2) = 5y_1^2 - 8y_1y_2 - 4y_2^2$.

***П р и м е р* 6.5.** Дана квадратичная форма

$F = F(x_1, x_2, x_3) = x_1^2 + 5x_2^2 - 4x_3^2 + 2x_1x_2 - 4x_1x_3$.

Найти квадратичную форму $G(y_1, y_2, y_3)$, полученную из данной линейным преобразованием

$$x_1 = y_1 - \frac{1}{2}y_2 + \frac{5}{6}y_3, \quad x_2 = \frac{1}{2}y_2 - \frac{1}{6}y_3, \quad x_3 = \frac{1}{3}y_3.$$

Р е ш е н и е. Запишем матрицу A данной квадратичной формы, матрицу преобразования C и вычислим $\widetilde{A} = C'AC$:

$$A = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 5 & 0 \\ -2 & 0 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1/2 & 5/6 \\ 0 & 1/2 & -1/6 \\ 0 & 0 & 1/3 \end{pmatrix},$$

$$\widetilde{A} = \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1/2 & 0 \\ 5/6 & -1/6 & 1/3 \end{pmatrix}\begin{pmatrix} 1 & 1 & -2 \\ 1 & 5 & 0 \\ -2 & 0 & -4 \end{pmatrix}\begin{pmatrix} 1 & -1/2 & 5/6 \\ 0 & 1/2 & -1/6 \\ 0 & 0 & 1/3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Следовательно, $G = G(y_1, y_2, y_3) = y_1^2 + y_2^2 - y_3^2$.

Пример 6.5 показывает, что при удачно выбранных линейных преобразованиях вид квадратичной формы можно существенно упростить.

6.2. КАНОНИЧЕСКИЙ ВИД КВАДРАТИЧНОЙ ФОРМЫ

Будем говорить, что квадратичная форма имеет **канонический вид**, если ее матрица диагональна ($a_{ij} = 0$ при $i \neq j$):

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Очевидно, в этом случае квадратичная форма представляет собой сумму квадратов неизвестных вида

$$F = a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 = \sum_{i=1}^{n} a_{ii}x_i^2.$$

В частности, квадратичная форма G, полученная в примере 6.5, имеет канонический вид.

Мы установили, что ранг квадратичной формы не меняется при невырожденных линейных преобразованиях (см. § 6.1). Пусть квадратичная форма $F(x_1, x_2, ..., x_n)$ приведена невырожденным линейным преобразованием к каноническому виду

$$b_1y_1^2 + b_2y_2^2 + \ldots + b_ny_n^2, \tag{6.12}$$

где $y_1, y_2, ..., y_n$ — новые неизвестные. Здесь некоторые из коэффициентов $b_1, b_2, ..., b_n$ могут быть нулями.

Нетрудно доказать, что *ранг квадратичной формы равен числу отличных от нуля коэффициентов в каноническом виде, к которому приводится данная квадратичная форма*.

Действительно, если квадратичная форма ранга r приведена невырожденным линейным преобразованием к виду (6.12), это означает, что матрица преобразованной квадратичной формы имеет вид

$$\begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_n \end{pmatrix}$$

и также имеет ранг r. А это равносильно тому, что на главной диагонали находится r отличных от нуля элементов.

Справедлива следующая основная теорема о квадратичных формах (приведем ее без доказательства).

***Теорема* 6.3.** Всякая квадратичная форма может быть приведена некоторым невырожденным линейным преобразованием к каноническому виду.

Следует заметить, что всякая квадратичная форма может быть приведена к каноническому виду различными способами. При этом канонический вид, к которому приводится данная квадратичная форма, не является для нее однозначно определенным.

***Пример* 6.6.** Дана квадратичная форма

$$F = 2x_1x_2 - 6x_2x_3 + 2x_3x_1.$$

Проверить, что она приводится к каноническому виду линейным преобразованием:

$$\begin{aligned} x_1 &= \frac{1}{2}y_1 + \frac{1}{2}y_2 + 3y_3, \\ x_2 &= \frac{1}{2}y_1 - \frac{1}{2}y_2 - y_3, \\ x_3 &= y_3. \end{aligned}$$

Решение:

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

Получаем

$$\tilde{A} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & 3 \\ 1/2 & -1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 6 \end{pmatrix},$$

т.е. квадратичная форма приводится к виду

$$\frac{1}{2}y_1^2 - \frac{1}{2}y_2^2 + 6y_3^2.$$

***Пример* 6.7.** Дана квадратичная форма из примера 6.6:

$$F = 2x_1x_2 - 6x_2x_3 + 2x_3x_1.$$

Проверить, что линейное преобразование:

$$\begin{aligned} x_1 &= y_1 + 3y_2 + 2y_3, \\ x_2 &= y_1 - y_2 - 2y_3, \\ x_3 &= y_2 \end{aligned}$$

также приводит эту квадратичную форму к каноническому виду.

Р е ш е н и е:

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

Получаем

$$\widetilde{A}=\begin{pmatrix}1 & 1 & 0\\ 3 & -1 & 1\\ 2 & -2 & 0\end{pmatrix}\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & -3\\ 1 & -3 & 0\end{pmatrix}\begin{pmatrix}1 & 3 & 2\\ 1 & -1 & -2\\ 0 & 1 & 0\end{pmatrix}=\begin{pmatrix}2 & 0 & 0\\ 0 & 6 & 0\\ 0 & 0 & -8\end{pmatrix},$$

т.е. квадратичная форма из примера 6.6 приводится к другому каноническому виду:

$$2y_1^2+6y_2^2-8y_3^2.$$

Из примеров 6.6 и 6.7 видим, что одна и та же квадратичная форма

$$F=2x_1x_2-6x_2x_3+2x_3x_1$$

в результате одного невырожденного линейного преобразования приобрела вид

$$G=\frac{1}{2}y_1^2-\frac{1}{2}y_2^2+6y_3^2,$$

а в результате другого –

$$G_1=2y_1^2+6y_2^2-8y_3^2.$$

Несмотря на то что G и G_1 заметно отличаются друг от друга, у них все же есть одно общее свойство: они содержат одинаковое число положительных и отрицательных коэффициентов (два и один соответственно). Это не случайно. Имеет место следующее утверждение (приведем его без доказательства).

Т е о р е м а **6.4 (закон инерции квадратичных форм).** Число положительных и число отрицательных коэффициентов в записи квадратичной формы в каноническом виде, к которому приводится данная квадратичная форма невырожденным линейным преобразованием, не зависят от выбора этого преобразования.

Наряду с каноническим видом рассматривают также **нормальный вид** квадратичной формы, т.е. сумму квадратов неизвестных с коэффициентами +1 или –1.

В частности, в примере 6.5 квадратичная форма

$$F = x_1^2 + 5x_2^2 - 4x_3^2 + 2x_1x_2 - 4x_1x_3$$

преобразована к виду

$$G = y_1^2 + y_2^2 - y_3^2,$$

который является не только каноническим, но и нормальным.

Нетрудно убедиться, что квадратичную форму, преобразованную к каноническому виду, всегда можно привести невырожденным линейным преобразованием к нормальному виду.

Действительно, пусть

$$G = c_1y_1^2 + \ldots + c_ky_k^2 - c_{k+1}y_{k+1}^2 - \ldots - c_ry_r^2,$$

где $c_1, \ldots, c_k, c_{k+1}, \ldots, c_r$ положительны.

Тогда преобразование $z_i = \sqrt{c_i}\, y_i$ $(i = 1, 2, \ldots, r)$, $z_j = y_j$ $(j = r+1, \ldots, n)$ приводит G к нормальному виду:

$$G_1 = z_1^2 + \ldots + z_k^2 - z_{k+1}^2 - \ldots - z_r^2.$$

Закон инерции квадратичных форм можно теперь сформулировать так: *число положительных и отрицательных квадратов в нормальном виде квадратичной формы не зависит от выбора линейного невырожденного преобразования, которым квадратичная форма приведена к этому виду*.

6.3. ПОЛОЖИТЕЛЬНО И ОТРИЦАТЕЛЬНО ОПРЕДЕЛЕННЫЕ КВАДРАТИЧНЫЕ ФОРМЫ

Определение. Квадратичная форма $F(x_1, x_2, \ldots, x_n)$ называется **положительно определенной**, если при всех значениях неизвестных, из которых хотя бы одно отлично от нуля, выполняется неравенство

$$F(x_1, x_2, \ldots, x_n) > 0.$$

Если же $F(x_1, x_2, \ldots, x_n) < 0$ для всех значений неизвестных, из которых хотя бы одно отлично от нуля, то квадратичная форма называется **отрицательно определенной**.

Нетрудно доказать (мы этого делать не будем), что квадратичная форма от n неизвестных является положительно определенной тогда и только тогда, когда она приводится к нормальному виду, состоящему из n положительных квадратов.

Сформулируем один из часто употребляемых критериев положительно (отрицательно) определенной квадратичной формы.

***Теорема* 6.5.** Для того чтобы квадратичная форма $F = X'AX$ была положительно (отрицательно) определенной, необходимо и достаточно, чтобы все собственные значения λ_i матрицы A были положительны (отрицательны).

***Пример* 6.8.** Выяснить, является ли квадратичная форма

$$F = 2x_1^2 - 4x_1x_2 + 5x_2^2$$

положительно определенной.

Решение. Матрица A квадратичной формы имеет вид

$$A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.$$

Составляем характеристическое уравнение:

$$\begin{vmatrix} 2-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = \lambda^2 - 7\lambda + 6 = 0.$$

Корни характеристического уравнения $\lambda_1 = 6$, $\lambda_2 = 1$ положительны, следовательно, квадратичная форма является положительно определенной.

Сформулируем еще один часто применяемый критерий положительной определенности квадратичной формы.

Пусть $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ – матрица квадратичной формы $F = X'AX$.

Главными минорами этой матрицы называются определители

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

Теорема **6.6 (критерий Сильвестра).** Для того чтобы квадратичная форма $F = X'AX$ была положительно определенной, необходимо и достаточно, чтобы все главные миноры матрицы A были положительны:

$$\Delta_1 > 0,\ \Delta_2 > 0,\ \ldots,\ \Delta_n > 0.$$

Для того чтобы квадратичная форма была отрицательно определенной, необходимо и достаточно, чтобы знаки главных миноров $\Delta_1, \Delta_2, \ldots, \Delta_n$ чередовались, причем $\Delta_1 < 0$.

(Эту теорему также примем без доказательства.)

Пример **6.9.** Убедиться, пользуясь критерием Сильвестра, что квадратичная форма

$$F = 2x_1^2 - 4x_1x_2 + 5x_2^2$$

из примера 6.8 является положительно определенной.

Решение:

$$\Delta_1 = a_{11} = 2 > 0, \quad \Delta_2 = \begin{vmatrix} 2 & -2 \\ -2 & 5 \end{vmatrix} = 6 > 0.$$

Главные миноры матрицы A положительны, следовательно, по критерию Сильвестра квадратичная форма является положительно определенной.

Пример **6.10.** Выяснить, является ли квадратичная форма

$$F = 5x_1^2 + x_2^2 + 6x_3^2 + 4x_1x_2 - 10x_1x_3 - 4x_2x_3$$

положительно определенной.

Решение. Вычисляем главные миноры:

$$\Delta_1 = 5, \quad \Delta_2 = \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1, \quad \Delta_3 = \begin{vmatrix} 5 & 2 & -5 \\ 2 & 1 & -2 \\ -5 & -2 & 6 \end{vmatrix} = 1.$$

Главные миноры положительны, следовательно, данная квадратичная форма положительно определена.

Вопросы

1. Является ли выражение $x_1x_2 + x_1x_3 + x_1x_4$ квадратичной формой?
2. Как определяется матрица квадратичной формы? Всегда ли матрица квадратичной формы является квадратной?

3. Что называется рангом квадратичной формы?
4. Как преобразуется матрица A квадратичной формы при невырожденном линейном преобразовании C?
5. Как связан ранг квадратичной формы с числом отличных от нуля коэффициентов в каноническом виде, к которому приводится эта форма?
6. В чем заключается закон инерции квадратичных форм?
7. Является ли нормальный вид квадратичной формы ее каноническим видом?
8. Каков нормальный вид положительно определенной квадратичной формы?

Раздел II
ЭЛЕМЕНТЫ АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

Глава 7
ЛИНИИ НА ПЛОСКОСТИ

7.1. ОСНОВНЫЕ ПОНЯТИЯ

Пусть на плоскости задана система координат, и пусть мы имеем на плоскости некоторую линию (прямую или кривую).

Определение. **Уравнением линии** на плоскости называется такое равенство

$$F(x, y) = 0, \tag{7.1}$$

которому удовлетворяют координаты любой точки, принадлежащей этой линии, и не удовлетворяют координаты любой точки, не принадлежащей этой линии.

Короче говоря, уравнение (7.1) есть уравнение линии, если ему удовлетворяют координаты всех тех и только тех точек, которые принадлежат этой линии.

***Пример* 7.1.** Написать уравнение множества точек, равноудаленных от оси Ox и от точки $A(0, 2)$.

Решение. Из школьного курса известно, что расстояние между точками $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$ вычисляется по формуле

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

а расстояние от точки до оси Ox есть ордината этой точки, взятая с соответствующим знаком.

Пусть $M(x, y)$ – произвольная точка искомой линии. Тогда $MM_0 = MA$ (рис. 7.1) или

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

(Здесь, очевидно, $y > 0$). Возводя в квадрат обе части уравнения, получаем

$$y^2 = x^2 + y^2 - 4y + 4,$$

или после очевидных преобразований

$$y = \frac{x^2}{4} + 1.$$

Эта линия – парабола с вершиной в точке (0, 1).

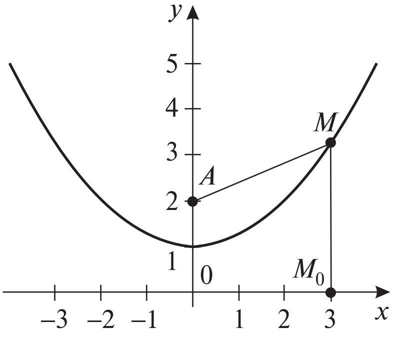


Рис. 7.1. Парабола $y = \frac{x^2}{4} + 1$

Заметим, что во многих случаях из уравнения (7.1) можно выразить y через x и записать уравнение линии в виде $y = f(x)$.

В тех случаях, когда линия задается алгебраическим уравнением n-го порядка (в частности, первого или второго), то говорим, что это линия n-го порядка (соответственно первого или второго порядка). Например, $y = 3x^2$, $x^2 + y^2 - 4 = 0$ – линии второго порядка.

7.2. ОБЩЕЕ УРАВНЕНИЕ ЛИНИИ ПЕРВОГО ПОРЯДКА. ПРЯМАЯ НА ПЛОСКОСТИ

К линиям первого порядка относятся те линии, для которых задающее их уравнение (7.1) является линейным, т.е. алгебраическим уравнением, которое содержит переменные x и y только в первой степени:

$$Ax + By + C = 0. \tag{7.2}$$

Если в этом уравнении $B \neq 0$, то можно выразить y:

$$y = -\frac{A}{B}x - \frac{C}{B},$$

или, введя новые обозначения $k = -\frac{A}{B}$, $b = -\frac{C}{B}$,

$$y = kx + b. \qquad (7.3)$$

Уравнение (7.3) называется **уравнением прямой с угловым коэффициентом** k. Здесь $k = \operatorname{tg}\varphi$, где φ — угол между прямой и положительным направлением оси Ox (рис. 7.2).

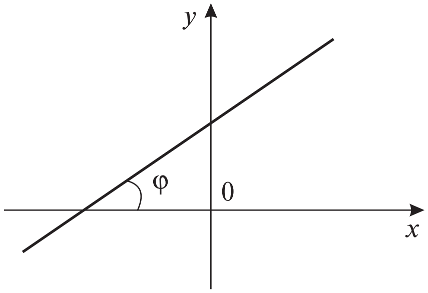


Рис. 7.2. Прямая с угловым коэффициентом $k = \operatorname{tg}\varphi$

Если же $B = 0$ в уравнении (7.2), то прямая перпендикулярна оси Ox, ее угловой коэффициент не определен и уравнение имеет вид $x = a$.

Полезно знать некоторые разновидности уравнения прямой.

1. Если известны *угловой коэффициент* k и *точка* $M(x_0, y_0)$, через которую проходит прямая, то, очевидно, выполняется тождество

$$y_0 = kx_0 + b. \qquad (*)$$

Вычитая это тождество из уравнения (7.3), получаем **уравнение прямой с заданным угловым коэффициентом и проходящей через заданную точку**:

$$y - y_0 = k(x - x_0). \qquad (7.4)$$

2. Если прямая проходит через *две заданные точки* $M_0(x_0, y_0)$ и $M_1(x_1, y_1)$, то кроме тождества (*) выполняется также тождество

$$y_1 = kx_1 + b. \qquad (**)$$

Из тождеств (*) и (**) находим

$$k = \frac{y_1 - y_0}{x_1 - x_0}$$

и с учетом (7.4) получаем **уравнение прямой, проходящей через заданные точки**:

$$y - y_0 = \frac{y_1 - y_0}{x_1 - x_0}(x - x_0). \tag{7.5}$$

***Пример* 7.2.** Составить уравнение прямой, проходящей через точки $M_0(-2, -1)$ и $M_1(1, 5)$.

Решение. Подставляем координаты этих точек в (7.5):

$$y + 1 = \frac{5+1}{1+2}(x+2), \quad \text{или} \quad y = 2x + 3.$$

3. Если известна *точка* $M_0(x_0, y_0)$, через которую проходит прямая, и *вектор* $\bar{n} = (A, B)$, перпендикулярный этой прямой, то уравнение прямой имеет вид

$$A(x - x_0) + B(y - y_0) = 0. \tag{7.6}$$

Докажем это. Пусть $M(x, y)$ – произвольная точка данной прямой. Тогда $\overline{M_0M} = (x - x_0, y - y_0)$. По условию $\overline{M_0M} \perp \bar{n}$, а это эквивалентно тому, что $\left(\overline{M_0M}, \bar{n}\right) = 0$. Записав это уравнение в координатной форме, получаем

$$A(x - x_0) + B(y - y_0) = 0,$$

что и требовалось доказать.

Раскроем скобки в последнем уравнении:

$$Ax + By - Ax_0 - By_0 = 0.$$

Обозначив $-Ax_0 - By_0 = C$, получаем

$$Ax + By + C = 0. \tag{7.7}$$

Итак, уравнение прямой есть линейное уравнение.

Докажем, что всякое уравнение первого порядка вида (7.7) есть **уравнение некоторой прямой на плоскости**.

Действительно, пусть задано уравнение первой степени (7.7). В нем хотя бы один из коэффициентов, A или B, отличен от нуля (иначе оно не было бы уравнением первой степени). Пусть, например, $A \neq 0$. Это уравнение всегда имеет решение (например, положив $y_0 = 1$, находим $x_0 = \dfrac{-B-C}{A}$). Пусть (x_0, y_0) – какое-нибудь решение уравнения (7.7), т.е.

$$Ax_0 + By_0 + C = 0. \qquad (7.7')$$

Вычитая (7.7′) из (7.7), получаем

$$A(x - x_0) + B(y - y_0) = 0.$$

Это уравнение равносильно уравнению (7.7), так как получено из (7.7) посредством тождественных преобразований. В то же время оно представляет собой уравнение прямой, как доказано выше. Следовательно, уравнение (7.7) есть уравнение прямой.

Уравнение (7.7) называется **общим уравнением прямой**, а всякий ненулевой вектор, перпендикулярный данной прямой, называется ее **нормальным вектором**. В частности, вектор $\bar{n} = (A, B)$ есть нормальный вектор прямой (7.7).

Угол между прямыми

I. Пусть заданы две прямые $y = k_1x + b_1$ и $y = k_2x + b_2$, где $k_1 = \operatorname{tg}\varphi_1$, $k_2 = \operatorname{tg}\varphi_2$. Пусть φ – угол между прямыми (рис. 7.3).

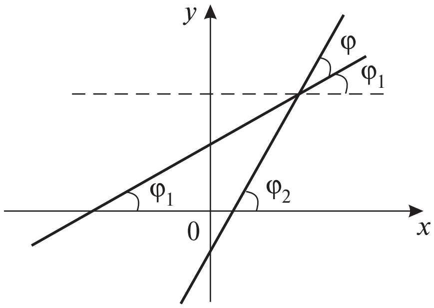


Рис. 7.3. Угол φ между прямыми

Тогда $\varphi = \varphi_2 - \varphi_1$ и по известной формуле из школьного курса тригонометрии

$$\operatorname{tg}\varphi = \operatorname{tg}(\varphi_2 - \varphi_1) = \frac{\operatorname{tg}\varphi_2 - \operatorname{tg}\varphi_1}{1 + \operatorname{tg}\varphi_1 \operatorname{tg}\varphi_2},$$

или

$$\operatorname{tg}\varphi = \frac{k_2 - k_1}{1 + k_1k_2}. \qquad (7.8)$$

Отсюда, в частности, сразу следует *условие параллельности прямых*:

$$k_1 = k_2.$$

Нетрудно также получить *условие перпендикулярности прямых*:

$$k_1 k_2 = -1.$$

***Пример* 7.3.** Найти угол между прямыми

$$y = 3x + 2, \quad y = -2x + 1.$$

Решение. Подставляя в (7.8) значения $k_1 = 3$, $k_2 = -2$, получаем

$$\operatorname{tg}\varphi = \frac{-2-3}{1-6} = 1.$$

Отсюда $\varphi = \dfrac{\pi}{4}$.

Заметим, что формула (7.8) определяет один из двух углов между пересекающимися прямыми; другой угол равен $\pi - \varphi$.

II. Пусть теперь две прямые l_1 и l_2 заданы общими уравнениями:

$$l_1\colon\ A_1x + B_1y + C_1 = 0,$$

$$l_2\colon\ A_2x + B_2y + C_2 = 0.$$

Угол между этими прямыми равен углу между их нормальными векторами (либо дополняет его до 180°). Следовательно, один из двух углов α между этими прямыми можно вычислить по формуле

$$\cos\alpha = \frac{(\bar{n}_1, \bar{n}_2)}{|\bar{n}_1||\bar{n}_2|} = \frac{A_1A_2 + B_1B_2}{\sqrt{A_1^2 + B_1^2}\sqrt{A_2^2 + B_2^2}}. \tag{7.9}$$

***Пример* 7.4.** Найти угол между прямыми, заданными общими уравнениями

$$3x - 4y + 7 = 0, \quad 8x - 6y + 15 = 0.$$

Решение. Применим формулу (7.9):

$$\cos\alpha = \frac{3\cdot 8 + 4\cdot 6}{\sqrt{3^2 + 4^2}\sqrt{8^2 + 6^2}} = \frac{24}{25}.$$

Стало быть, один из углов между данными прямыми равен $\arccos\dfrac{24}{25}$.

Условие параллельности прямых l_1 и l_2 есть условие параллельности их нормальных векторов $\bar{n}_1 = (A_1, B_1)$ и $\bar{n}_2 = (A_2, B_2)$:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2}. \tag{7.10}$$

Условие перпендикулярности прямых есть условие перпендикулярности их нормальных векторов:

$$A_1A_2 + B_1B_2 = 0. \tag{7.11}$$

Полуплоскости

Пусть прямая l задана уравнением (7.7). Ее нормальный вектор $\bar{n} = (A, B)$. Все точки плоскости, не принадлежащие l, разобьем на два множества π_1 и π_2 следующим образом:

$$M(x, y) \in \pi_1 \Leftrightarrow Ax + By + C > 0,$$

$$M(x, y) \in \pi_2 \Leftrightarrow Ax + By + C < 0.$$

Множество π_1 называется *положительной полуплоскостью* по отношению к уравнению прямой (7.7), а множество π_2 – *отрицательной полуплоскостью*. Следует подчеркнуть, что понятие положительной и отрицательной полуплоскостей определяется по отношению к уравнению прямой, а не к самой прямой. Очевидно, если умножить обе части уравнения (7.7) на –1, то получим уравнение той же самой прямой, однако при этом положительная полуплоскость станет отрицательной, а отрицательной – положительной.

Можно доказать (мы этого делать не будем), что вектор $\bar{n} = (A, B)$ направлен в ту часть плоскости, которая является положительной по отношению к уравнению прямой (7.7):

$$Ax + By + C = 0.$$

Расстояние от точки до прямой

Выведем формулу расстояния d от произвольной точки $M_0(x_0, y_0)$ до прямой (7.7).

Расстояние от точки M_0 до прямой (7.7) равно длине перпендикуляра, опущенного из M_0 на эту прямую. Обозначим через $N(x_1, y_1)$ основание этого перпендикуляра, т.е. точку пересечения перпендикуляра с прямой (7.7). Тогда по формуле расстояния между двумя точками

$$d = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}. \tag{7.12}$$

Угловой коэффициент k прямой (7.7), очевидно, равен

$$k = -\frac{A}{B}.$$

В соответствии с условием перпендикулярности угловой коэффициент перпендикуляра M_0N равен $k' = \frac{B}{A}$, а само уравнение этого перпендикуляра (с учетом того, что он проходит через точку $M_0(x_0, y_0)$) имеет вид

$$y - y_0 = \frac{B}{A}(x - x_0). \tag{7.13}$$

Далее можно было бы, решив совместно уравнения (7.7) и (7.13), найти координаты (x_1, y_1) точки N и подставить их в (7.12). Однако, несмотря на простоту этого решения, здесь возникнут громоздкие выражения. Поэтому применим другой способ. Воспользуемся тем, что точка N принадлежит M_0N. Поэтому неизвестные пока координаты x_1, y_1 точки N удовлетворяют уравнению (7.13):

$$y_1 - y_0 = \frac{B}{A}(x_1 - x_0).$$

Отсюда получаем равенство

$$\frac{y_1 - y_0}{B} = \frac{x_1 - x_0}{A}.$$

Обозначим общее значение этих дробей δ:

$$\frac{x_1 - x_0}{A} = \frac{y_1 - y_0}{B} = \delta.$$

Это значение δ неизвестно, так как неизвестны x_1 и y_1. Найдем его:

$$x_1 - x_0 = A\delta, \quad y_1 - y_0 = B\delta. \tag{*}$$

Подставляя эти разности в формулу (7.12), получим

$$d = \sqrt{(A\delta)^2 + (B\delta)^2} = \sqrt{(A^2 + B^2)\delta^2} = |\delta|\sqrt{A^2 + B^2}. \tag{7.14}$$

Следует заметить, что мы не знаем, положительно или отрицательно число δ.

Выразим x_1 и y_1 из (*):

$$x_1 = x_0 + A\delta, \quad y_1 = y_0 + B\delta$$

и подставим эти значения в уравнение (7.7). (Напомним, что точка $N(x_1, y_1)$ принадлежит и перпендикуляру к прямой (7.7), и самой прямой.) Получаем

$$A(x_0 + A\delta) + B(y_0 + B\delta) + C = 0,$$

откуда

$$\delta = -\frac{Ax_0 + By_0 + C}{A^2 + B^2}.$$

Подставим теперь найденное значение δ в формулу (7.14) и проведем необходимые сокращения:

$$d = \left| -\frac{Ax_0 + By_0 + C}{A^2 + B^2} \right| \sqrt{A^2 + B^2} = \frac{\left| Ax_0 + By_0 + C \right| \sqrt{A^2 + B^2}}{A^2 + B^2} =$$

$$= \frac{\left| Ax_0 + By_0 + C \right|}{\sqrt{A^2 + B^2}}.$$

Окончательно получаем

$$d = \frac{\left| Ax_0 + By_0 + C \right|}{\sqrt{A^2 + B^2}}. \qquad (7.15)$$

Итак, *расстояние до прямой, заданной общим уравнением*, можно найти, подставив координаты точки в левую часть этого уравнения, а затем модуль полученного числа разделив на квадратный корень из суммы квадратов коэффициентов этого уравнения прямой.

***Пример* 7.5.** Найти расстояние от точки $M_0(2, 3)$ до прямой $4x + 3y + 8 = 0$.

Решение. Применяем формулу (7.15):

$$d = \frac{\left| 4 \cdot 2 + 3 \cdot 3 + 8 \right|}{\sqrt{4^2 + 3^2}} = 5.$$

***Пример* 7.6.** Найти расстояние между параллельными прямыми l_1 и l_2:

$$l_1\colon\ 4x + 3y - 8 = 0,$$

$$l_2\colon\ 8x + 6y + 9 = 0.$$

Решение. Расстояние между двумя параллельными прямыми равно, очевидно, расстоянию от любой точки одной из этих прямых до другой прямой. Положим в уравнении первой из этих прямых $y = 0$, получим $x = 2$. Следовательно, точка $M_0(2, 0)$ принадлежит первой прямой. Находим расстояние от M_0 до прямой l_2:

$$d = \frac{|8 \cdot 2 + 6 \cdot 0 + 9|}{\sqrt{64 + 36}} = \frac{1}{4}.$$

***Пример* 7.7.** Написать уравнение биссектрисы угла между прямыми l_1 и l_2:

l_1: $3x - 4y + 7 = 0$,

l_2: $5x + 12y - 21 = 0$.

Решение. Известно, что любая точка биссектрисы находится на одинаковом расстоянии от сторон угла. Поэтому если $M(x, y)$ – точка, принадлежащая биссектрисе угла между прямыми l_1 и l_2, то

$$\frac{|3x - 4y + 7|}{\sqrt{9 + 16}} = \frac{|5x + 12y - 21|}{\sqrt{25 + 144}}.$$

Отсюда получаем

$$13|3x - 4y + 7| = 5|5x + 12y - 21|,$$

или

$$13(3x - 4y + 7) = \pm 5(5x + 12y - 21).$$

Получаем два уравнения биссектрисы:

1) $39x - 52y + 91 = 25x + 60y - 105$,
$14x - 112y + 196 = 0$,
$x - 8y + 14 = 0$;

2) $39x - 52y + 91 = -25x - 60y + 105$,
$64x + 8y - 14 = 0$,
$32x + 4y - 7 = 0$.

Итак, биссектрисы углов, образованных пересекающимися прямыми l_1 и l_2, – это прямые

$$x - 8y + 14 = 0 \quad \text{и} \quad 32x + 4y - 7 = 0.$$

***Пример* 7.8.** Написать уравнение биссектрисы внутреннего угла при вершине B треугольника ABC, где $A(1, 1)$, $B(5, -2)$, $C(2, 2)$.

Р е ш е н и е. Составим уравнения сторон AB и BC по формуле (7.5):

$$AB\colon\ y-1=\frac{-2-1}{5-1}(x-1),\quad \text{или}\quad 3x+4y-7=0;$$

$$BC\colon\ y+2=\frac{2+2}{2-5}(x-5),\quad \text{или}\quad 4x+3y-14=0.$$

Напишем уравнения обеих биссектрис угла между прямыми AB и BC:

$$\frac{|3x+4y-7|}{\sqrt{9+16}}=\frac{|4x+3y-14|}{\sqrt{16+9}},$$

$$|3x+4y-7|=|4x+3y-14|,$$

$$3x+4y-7=\pm(4x+3y-14),$$

откуда

$$l_1\colon\ x-y-7=0,$$

$$l_2\colon\ x+y-3=0.$$

Одна из этих биссектрис есть биссектриса внутреннего угла треугольника, а другая – внешнего угла.

Далее рассуждаем следующим образом:

1. Если точка M_0 принадлежит биссектрисе внутреннего угла треугольника и находится по ту сторону от точки B, что и треугольник ABC, то точка M_0 находится в той же полуплоскости относительно прямой AB, что и точка C, а также в той же полуплоскости относительно прямой BC, что и точка A.

2. Если же точка M_0 лежит по другую сторону от B, то она лежит в противоположных плоскостях как по отношению к прямой AB, так и по отношению к прямой BC (рис. 7.4).

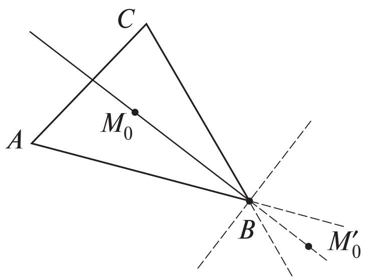


Рис. 7.4. Биссектриса внутреннего угла треугольника ABC

Возьмем на прямой l_1 произвольную точку, например (4, –3). Подставим ее координаты в уравнение прямой AB: $3\cdot 4+4\cdot(-3)-7=$ $=-7<0$. Подставим координаты точки C в это же уравнение: $3\cdot 2+$ $+4\cdot 2-7=7>0$. Подставим теперь координаты точки (4, –3) в уравнение прямой BC: $4\cdot 4+3\cdot(-3)-14=-7<0$. Подставим координаты точки A в это уравнение: $4+3-14=-7<0$.

Итак, точка (4, –3) находится по отношению к прямой BC в той же полуплоскости, что и точка A, но по отношению к прямой AB – не в той же полуплоскости, в которой находится точка C. Следовательно, точка (4, –3) принадлежит биссектрисе внешнего, а не внутреннего угла треугольника.

Итак, прямая l_1: $x-y-7=0$ – биссектриса внешнего угла при вершине B. Следовательно, искомая биссектриса есть прямая l_2: $x+y-3=0$.

Вопросы

1. Что называется уравнением линии на плоскости? Приведите примеры уравнений линий.
2. Что такое порядок алгебраической линии?
3. Что называется угловым коэффициентом прямой линии на плоскости? Определен ли угловой коэффициент прямой, параллельной оси Oy?
4. Что называется нормальным вектором прямой на плоскости? Как по общему уравнению прямой определить один из ее нормальных векторов?
5. Как определить острый угол между прямыми, заданными общими уравнениями?
6. Как вычислить расстояние между двумя параллельными прямыми на плоскости, заданными общими уравнениями?

Глава 8
ВАЖНЕЙШИЕ КРИВЫЕ ВТОРОГО ПОРЯДКА

8.1. ОКРУЖНОСТЬ. ЭЛЛИПС

Определение. **Окружностью** называется множество всех точек плоскости, находящихся на одинаковом расстоянии, называемом **радиусом**, от фиксированной точки, называемой **центром** окружности.

Пусть радиус окружности равен R и центром является точка $C(x_0, y_0)$ (рис. 8.1). Выведем уравнение этой окружности.

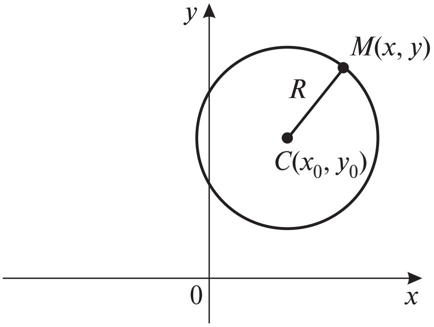


Рис. 8.1. Окружность с центром в точке C и радиусом $R = CM$

Для любой точки $M(x, y)$ окружности выполняется равенство $CM = R$, т.е. $\sqrt{(x - x_0)^2 + (y - y_0)^2} = R.$

Отсюда получаем уравнение окружности:

$$(x - x_0)^2 + (y - y_0)^2 = R^2.$$

В частности, если центр окружности совпадает с началом координат, то уравнение окружности имеет вид

$$x^2 + y^2 = R^2. \tag{8.1}$$

Уравнение (8.1) называется **каноническим уравнением окружности**.

Определение. **Эллипсом** называется линия, для всех точек которой сумма расстояний до двух фиксированных точек, называемых **фокусами**, есть величина постоянная и бóльшая, чем расстояние между фокусами.

Выведем уравнение эллипса. Выберем систему координат так, чтобы ось Ox проходила через фокусы F_1 и F_2, а ось Oy – посередине между фокусами (рис. 8.2).

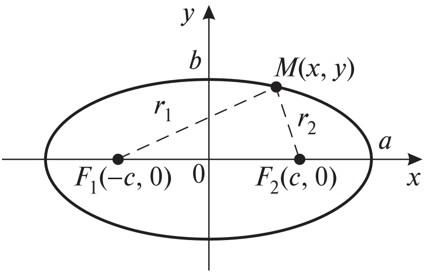


Рис. 8.2. Эллипс $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$

Пусть расстояние между фокусами равно $2c$, а сумма расстояний произвольной точки $M(x, y)$ эллипса от фокусов равна $2a$: $F_1M + F_2M = 2a$ (согласно определению $a > c$). Выразим расстояния от точки $M(x, y)$ до фокусов $F_1(-c, 0)$ и $F_2(c, 0)$ соответственно: $r_1 = F_1M = \sqrt{(x+c)^2 + y^2}$, $r_2 = F_2M = \sqrt{(x-c)^2 + y^2}$. Так как $r_1 + r_2 = 2a$, то

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a. \tag{8.2}$$

Это уравнение эллипса. Преобразуем его:

$$\sqrt{(x+c)^2 + y^2} = 2a - \sqrt{(x-c)^2 + y^2},$$

$$(x+c)^2 + y^2 = 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2,$$

$$a\sqrt{(x-c)^2 + y^2} = a^2 - cx,$$

$$a^2\left[(x-c)^2 + y^2\right] = \left(a^2 - cx\right)^2,$$

$$a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2 = a^4 - 2a^2cx + c^2x^2,$$

$$a^2x^2 - c^2x^2 + a^2y^2 = a^4 - a^2c^2.$$

Так как $a^2 - c^2 > 0$, то можно обозначить $a^2 - c^2 = b^2$. Получаем:

$$b^2x^2 + a^2y^2 = a^2b^2,$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{8.3}$$

Необходимо убедиться в том, что уравнение (8.3) действительно есть уравнение эллипса. Пока можно утверждать только то, что каждая точка $M(x, y)$, удовлетворяющая уравнению эллипса (8.2), удовлетворяет и уравнению (8.3). Однако уравнение (8.3) получено после двукратного возведения в квадрат, а мы знаем, что при возведении в квадрат обеих частей уравнения может получиться уравнение, не равносильное исходному. Убедимся, что здесь этого не произошло. Надо доказать, что каждая точка $M(x, y)$, удовлетворяющая уравнению (8.3), есть точка эллипса, т.е. что для нее выполнено условие $r_1 + r_2 = 2a$.

Итак, пусть $M(x, y)$ – произвольная точка, координаты которой удовлетворяют уравнению (8.3). Найдем расстояния r_1 и r_2 точки M от фокусов F_1 и F_2 соответственно.

Имеем

$$r_1 = \sqrt{(x+c)^2 + y^2}. \tag{*}$$

Выразим y^2 из уравнения (8.3):

$$y^2 = b^2\left(1 - \frac{x^2}{a^2}\right).$$

Однако $b^2 = a^2 - c^2$, следовательно,

$$y^2 = (a^2 - c^2)\left(1 - \frac{x^2}{a^2}\right) = a^2 - c^2 - x^2 + \frac{c^2}{a^2}x^2.$$

Подставим это значение y^2 в равенство (*):

$$r_1 = \sqrt{x^2 + 2cx + c^2 + a^2 - c^2 - x^2 + \frac{c^2}{a^2}x^2} = \sqrt{2cx + a^2 + \frac{c^2}{a^2}x^2} =$$

$$= \sqrt{\left(a + \frac{c}{a}x\right)^2}.$$

Величина e, определяемая соотношением $e = \dfrac{c}{a}$, называется **эксцентриситетом** эллипса, при этом $0 < e < 1$. Он характеризует меру вытянутости эллипса. Чем больше эксцентриситет, тем сильнее вытянут эллипс. Эксцентриситет равен нулю тогда и только тогда, когда фокусы эллипса совпадают: $F_1 = F_2$. В этом случае эллипс превращается в окружность радиуса a.

Имеем

$$r_1 = \pm\left(a + \frac{c}{a}x\right) = \pm(a + ex).$$

Слева – положительное число r_1. Следовательно, справа надо выбрать такой знак, чтобы правая часть также была положительной. Из (8.3) следует, что $|x| \le a$. Кроме того, $0 < e < 1$, следовательно, $|ex| < a$. Итак, независимо от того, $x > 0$ или $x < 0$, всегда $a + ex > 0$, поэтому справа надо взять знак «плюс»:

$$r_1 = a + ex. \tag{**}$$

Рассуждая аналогично, получим

$$r_2 = a - ex. \tag{***}$$

Из (**) и (***) имеем

$$r_1 + r_2 = 2a,$$

следовательно, точка $M(x, y)$ принадлежит эллипсу.

Мы доказали, что уравнение (8.3) есть уравнение эллипса. Оно называется **каноническим уравнением эллипса**.

Здесь a – бóльшая полуось эллипса, b – малая полуось ($b = \sqrt{a^2 - c^2}$). Из уравнения (8.3) следует, что оси Ox и Oy являются осями симметрии эллипса, а точка их пересечения – точка $O(0, 0)$ – центром симметрии.

В частном случае, когда $a = b$, фокусы эллипса сливаются, $c = 0$ и мы имеем окружность радиуса a с центром в начале координат.

8.2. ГИПЕРБОЛА

Определение. **Гиперболой** называется линия, для всех точек которой модуль разности расстояний до двух фиксированных точек, называемых **фокусами**, есть величина постоянная и меньшая, чем расстояние между фокусами.

Обозначим, как и в случае с эллипсом, расстояние между фокусами F_1 и F_2 через $2c$, выберем систему координат так, чтобы ось Ox проходила через фокусы, а ось Oy – посередине между ними (рис. 8.3). Обозначим расстояния от произвольной точки $M(x, y)$ гиперболы до фокусов F_1 и F_2 соответственно через r_1 и r_2, получаем:

$$\left|r_1 - r_2\right| = 2a.$$

Принимая во внимание равенства $r_1 = F_1M = \sqrt{(x+c)^2 + y^2}$, $r_2 = F_2M = \sqrt{(x-c)^2 + y^2}$, получаем уравнение гиперболы:

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

После преобразований, аналогичных тем, которые мы проводили с уравнением эллипса (см. § 8.1), получаем **каноническое уравнение гиперболы**:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \tag{8.4}$$

где $b^2 = c^2 - a^2$.

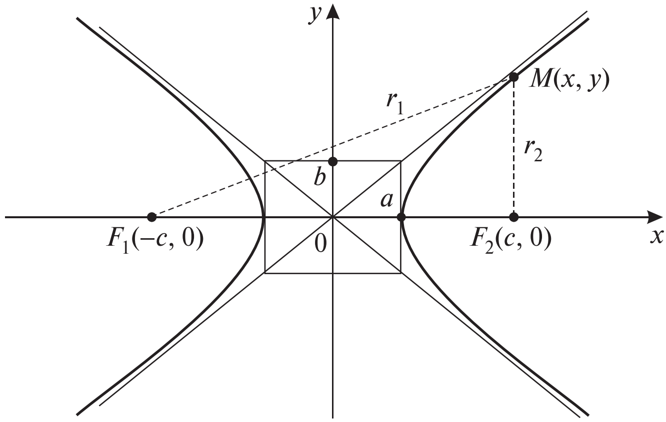


Рис. 8.3. Гипербола $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

Гипербола имеет две оси симметрии, точка пересечения которых является ее центром симметрии.

Нетрудно показать, что при $x \to \infty$ ветви гиперболы как угодно близко подходят к прямым $y = \pm\frac{b}{a}x$, называемым **асимптотами*** **гиперболы**.

---

* Подробнее об асимптотах см. в § 19.3.

8.3. ПАРАБОЛА

Определение. **Параболой** называется линия, для всех точек которой расстояние до фиксированной точки, называемой **фокусом**, равно расстоянию до фиксированной прямой, называемой **директрисой** и не проходящей через фокус.

Пусть на плоскости даны точка F и прямая d, не проходящая через эту точку. Выведем уравнение параболы с фокусом F и директрисой d. Пусть расстояние от точки F до прямой d равно p. Выберем систему координат следующим образом: ось Ox проведем через точку F перпендикулярно к прямой d, а ось Oy – посередине между точкой F и прямой d (рис. 8.4).

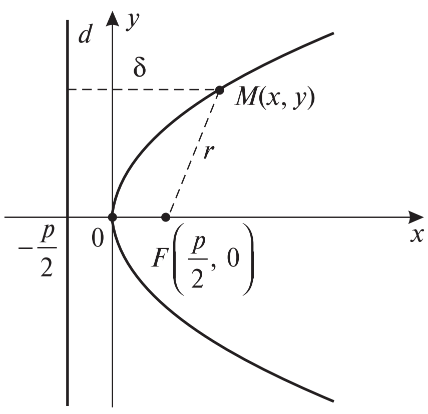


Рис. 8.4. Парабола $y^2 = 2px$

Пусть $M(x, y)$ – произвольная точка на параболе. Обозначим через δ расстояние от этой точки до директрисы, а через r – расстояние до фокуса. Согласно определению

$$r = \delta.$$

Учитывая, что $r = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2}$, $\delta = \frac{p}{2} + x$, получаем:

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \frac{p}{2} + x,$$

$$x^2 - px + \frac{p^2}{4} + y^2 = \frac{p^2}{4} + px + x^2,$$

$$y^2 = 2px. \tag{8.5}$$

Это **каноническое уравнение параболы**. Здесь число p называется **параметром** параболы.

Заметим, что каноническое уравнение параболы (8.5) отличается от знакомого по школьному курсу уравнения. Это связано с выбором системы координат. Если поменять оси координат, то вместо уравнения (8.5) получим привычное уравнение вида $y = ax^2$, где a – постоянное число. Аналогичное замечание относится и к уравнению гиперболы (8.4). Например, если для гиперболы $\frac{x^2}{2} - \frac{y^2}{2} = 1$ взять в качестве осей новой системы координат ее асимптоты, то в этой новой системе координат ее уравнение будет иметь знакомый вид $y = \frac{1}{x}$.

8.4. ОБЩЕЕ УРАВНЕНИЕ ЛИНИИ ВТОРОГО ПОРЯДКА

Общее уравнение линии второго порядка имеет вид

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0, \tag{8.6}$$

при этом $a_{11}^2 + a_{12}^2 + a_{22}^2 \neq 0$.

В аналитической геометрии доказывается, что существует девять различных кривых второго порядка:

- **эллипсы**;
- **гиперболы**;
- **параболы**;
- кривые, распадающиеся на **пару прямых**.

При этом эллипсы и пары прямых могут быть как действительными, так и мнимыми.

Таким образом, среди линий второго порядка кривыми в привычном понимании этого слова являются только эллипс, гипербола и парабола. Поэтому их называют важнейшими кривыми второго порядка.

Доказательство приведенного выше утверждения основано на преобразовании квадратичной формы. Поэтому его полезно привести здесь. Однако для этого надо предварительно изучить, как меняются координаты точек и уравнения линий при перемене системы координат.

8.5. ПРЕОБРАЗОВАНИЯ КООРДИНАТ

Прежде чем написать уравнение линии на плоскости, необходимо выбрать определенную систему координат. При этом в разных системах координат одна и та же линия будет иметь, очевидно, разные уравнения. Например, уравнение окружности радиуса R, центр которой имеет в выбранной системе Oxy координаты x_0 и y_0, выглядит так:

$$(x-x_0)^2+(y-y_0)^2=R^2.$$

В другой системе координат $O'x'y'$, у которой начало O' совпадает с центром этой окружности, уравнение этой окружности будет иметь вид

$$x'^2+y'^2=R^2.$$

Так будет, в частности, если координаты систем Oxy и $O'x'y'$ связаны соотношениями

$$x=x'+x_0,$$

$$y=y'+y_0.$$

Рассмотрим еще один пример. Пусть прямая задана в системе координат Oxy уравнением

$$x-y+2=0,$$

и пусть теперь выбрана другая система координат $O'x'y'$, связанная с прежней соотношениями

$$x=\frac{x'}{\sqrt{2}}-\frac{y'}{\sqrt{2}}-1,$$

$$y=\frac{x'}{\sqrt{2}}+\frac{y'}{\sqrt{2}}+1.$$

Тогда уравнение данной прямой в системе $O'x'y'$ будет иметь совсем простой вид:

$$y'=0.$$

Итак, мы видим, что удачный выбор координатной системы позволяет упростить уравнение рассматриваемой линии.

В аналитической геометрии переход от одной прямоугольной системы координат к другой осуществляется обычно с помощью поворота и параллельного переноса.

Параллельный перенос системы координат Oxy **в точку** $M_0(x_0, y_0)$ задается формулами

$$\begin{cases} x = x' + x_0, \\ y = y' + y_0, \end{cases} \tag{8.7}$$

выражающими старые координаты x, y через новые x', y'.

Поворот системы координат на угол α (совершаемый против часовой стрелки) задается формулами

$$\begin{cases} x = x'\cos\alpha - y'\sin\alpha, \\ y = x'\sin\alpha + y'\cos\alpha. \end{cases} \tag{8.8}$$

Выведем эти формулы.

1. Начнем с параллельного переноса. Пусть на плоскости заданы две прямоугольные системы координат: «старая» Oxy и «новая»* $O'x'y'$, причем ось $O'x'$ параллельна оси Ox, а ось $O'y'$ – оси Oy, кроме того, направления соответствующих старых и новых осей совпадают. Иначе говоря, новая система $O'x'y'$ получена из старой путем *параллельного переноса*, или *сдвига*, при котором начало координат O перемещается в точку O'.

Пусть точка O' имеет в старой системе координаты x_0 и y_0. Пусть для определенности $x_0 > 0$, $y_0 > 0$. Возьмем на плоскости произвольную точку M (рис. 8.5), и пусть (x, y) – ее координаты в старой системе, а (x', y') – в новой системе.

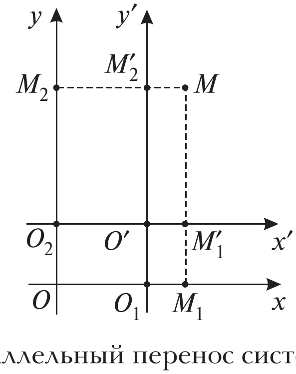


Рис. 8.5. Параллельный перенос системы координат

Рассмотрим рис. 8.5 (мы взяли также для определенности $M(x, y)$ с координатами $x > 0$, $y > 0$): здесь $OO_1 = x_0$, $OO_2 = y_0$, $OM_1 = x$,

---

* Далее слова «старая» и «новая» будем писать без кавычек.

$OM_2 = y$, $O'M_1' = x'$, $O'M_2' = y'$. Имеем $x = OM_1 = OO_1 + O_1M_1 = x_0 + x'$, $y = OM_2 = OO_2 + O_2M_2 = y_0 + y'$. Отсюда получаем (8.7).

Случай, когда точки O' и M имеют отрицательные координаты, рассматривается аналогично.

Замечание. Можно было бы рассуждать и иначе: рассмотреть векторы $\overline{OM}$, $\overline{O'M}$ и $\overline{OO'}$. Очевидно, $\overline{OM} = \overline{O'M} + \overline{OO'}$. Складывая векторы покоординатно, получим (8.7).

2. Перейдем теперь к повороту системы координат. Рассмотрим случай, когда новая система $Ox'y'$ получена из старой Oxy путем *поворота на некоторый угол* α, отсчитываемый против часовой стрелки. При этом обе системы имеют общее начало O.

Пусть, как и в предыдущем случае, точка M имеет в старой системе координаты (x, y), а в новой – соответственно (x', y').

Рассмотрим для определенности случай, когда угол $\alpha = \angle BOC$ – острый. Пусть M_1 – проекция точки M на Ox, B – проекция этой же точки M на Ox' (рис. 8.6). Стороны угла, образованного прямыми MM_1 и MB, перпендикулярны сторонам угла, образованного осями Ox, Ox' и равного α. Следовательно, $\angle AMB = \alpha$.

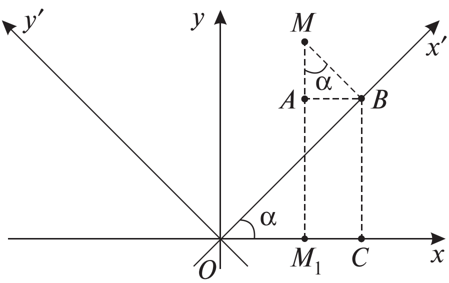


Рис. 8.6. Поворот системы координат

В соответствии с обозначениями на рис. 8.6:

$$x = OM_1,\ y = M_1M; \qquad x' = OB,\ y' = BM.$$

Однако

$$OM_1 = OC - M_1C = OC - AB.$$

Из треугольника OBC (в котором OC – катет, прилежащий к углу α) находим:

$$OC = OB\cos\alpha = x'\cos\alpha.$$

Из треугольника AMB (в котором AB – катет, противолежащий углу α) находим:

$$AB = BM \sin\alpha = y' \sin\alpha.$$

Следовательно,

$$x = OM_1 = x' \cos\alpha - y' \sin\alpha. \qquad (*)$$

Аналогично выражаем y:

$$y = M_1M = M_1A + AM = CB + AM,$$

$$CB = OB \sin\alpha = x' \sin\alpha, \qquad AM = BM \cos\alpha = y' \cos\alpha.$$

Следовательно,

$$y = M_1M = x' \sin\alpha + y' \cos\alpha. \qquad (**)$$

Из (*) и (**) получаем (8.8).

Заметим, что для случая, когда угол α не является острым, рассуждения аналогичны. Формулы (8.8) справедливы для любого угла α.

Итак, мы вывели формулы (8.7) и (8.8), выражающие старые координаты через новые. Может показаться, что полезнее были бы формулы, выражающие, наоборот, новые координаты через старые. Эти формулы получить легко. Из (8.7) сразу находим

$$\begin{cases} x' = x - x_0, \\ y' = y - y_0. \end{cases} \qquad (8.7')$$

Далее, умножив в (8.8) первое равенство на $\cos\alpha$, а второе – на $\sin\alpha$ и сложив их, получим

$$x' = x\cos\alpha + y\sin\alpha.$$

Аналогично, умножив второе равенство на $\cos\alpha$ и вычтя из него первое, умноженное на $\sin\alpha$, получим

$$y' = -x\sin\alpha + y\cos\alpha.$$

Окончательно имеем

$$\begin{cases} x' = x\cos\alpha + y\sin\alpha, \\ y' = -x\sin\alpha + y\cos\alpha. \end{cases} \qquad (8.8')$$

Однако на практике редко приходится находить новые координаты точек по их старым координатам. Значительно чаще требуется из уравнения линии в старой системе координат получать ее уравнение в новой системе. А для этого надо заменять старые координаты на новые, т.е. применять формулы (8.7) и (8.8), а не (8.7′) и (8.8′).

Рассмотрим теперь **общий случай преобразования координат**, когда надо от прямоугольной системы Oxy перейти к новой прямоугольной системе $O'x'y'$, у которой начало O' не совпадает с точкой O, а оси $O'x'$ и $O'y'$ не параллельны осям Ox и Oy соответственно. Пусть новое начало O' имеет в старой системе Oxy координаты (x_0, y_0), а ось $O'x'$ образует с осью Ox угол α. Тогда переход от системы Oxy к системе $O'x'y'$ можно осуществить в два этапа:

1) совершить параллельный перенос системы Oxy так, чтобы начало оказалось в точке $O'(x_0, y_0)$;

2) совершить поворот вокруг точки O' на угол α.

Нетрудно убедиться, что при этом старые координаты будут выражаться через новые с помощью формул

$$\begin{cases} x = x'\cos\alpha - y'\sin\alpha + x_0, \\ y = x'\sin\alpha + y'\cos\alpha + y_0. \end{cases}$$

8.6. ПРЕОБРАЗОВАНИЕ ОБЩЕГО УРАВНЕНИЯ ЛИНИИ ВТОРОГО ПОРЯДКА

Пусть дано общее уравнение линии второго порядка

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0, \tag{8.6}$$

где $a_{11}^2 + a_{12}^2 + a_{22}^2 \neq 0$.

Обозначим левую часть уравнения (8.6) через $F(x, y)$:

$$F(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33}.$$

В этом многочлене члены второго порядка образуют квадратичную форму:

$$\varphi(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2. \tag{8.9}$$

Первый шаг преобразования заключается в том, чтобы поворотом системы координат на некоторый угол α преобразовать квадратичную форму (8.9) к каноническому виду:

$$a'_{11}x'^2 + a'_{22}y'^2.$$

Заметим, что преобразование (8.8) – невырожденное линейное преобразование.

Итак, делаем преобразование (8.8):

$$F(x, y) = a_{11}(x'^2\cos^2\alpha - 2x'y'\cos\alpha\sin\alpha + y'^2\sin^2\alpha) +$$
$$+ 2a_{12}(x'^2\cos\alpha\sin\alpha - x'y'\sin^2\alpha - y'^2\sin\alpha\cos\alpha + x'y'\cos^2\alpha) +$$
$$+ a_{22}(x'^2\sin^2\alpha + 2x'y'\cos\alpha\sin\alpha + y'^2\cos^2\alpha) +$$
$$+ 2a_{13}x'\cos\alpha - 2a_{13}y'\sin\alpha + 2a_{23}x'\sin\alpha + 2a_{23}y'\cos\alpha + a_{33} =$$
$$= a'_{11}x'^2 + 2a'_{12}x'y' + a'_{22}y'^2 + 2a'_{13}x' + 2a'_{23}y' + a_{33} = F'(x', y'),$$

где новые коэффициенты имеют вид:

$$\begin{aligned}
a'_{11} &= a_{11}\cos^2\alpha + 2a_{12}\cos\alpha\sin\alpha + a_{22}\sin^2\alpha,\\
a'_{12} &= -a_{11}\cos\alpha\sin\alpha + a_{12}(\cos^2\alpha - \sin^2\alpha) + a_{22}\cos\alpha\sin\alpha,\\
a'_{22} &= a_{11}\sin^2\alpha - 2a_{12}\cos\alpha\sin\alpha + a_{22}\cos^2\alpha,\\
a'_{13} &= a_{13}\cos\alpha + a_{23}\sin\alpha,\\
a'_{23} &= -a_{13}\sin\alpha + a_{23}\cos\alpha.
\end{aligned} \qquad (8.10)$$

Угол α определяется требованием, чтобы было $a'_{12} = 0$, т.е. чтобы в преобразованном уравнении отсутствовал член, содержащий произведение неизвестных. Согласно (8.10) требование $a'_{12} = 0$ означает

$$a_{12}\cos^2\alpha + (a_{22} - a_{11})\cos\alpha\sin\alpha - a_{12}\sin^2\alpha = 0. \qquad (8.11)$$

Итак, при повороте на угол α, удовлетворяющий равенству (8.11), квадратичная форма (8.9) будет иметь канонический вид.

В уравнении (8.11) естественно предположить, что $a_{12} \neq 0$ (если бы было $a_{12} = 0$, то ничего преобразовывать не надо было бы – квадратичная форма $\varphi(x, y)$ уже имела бы вид $a_{11}x^2 + a_{22}y^2$.

Из (8.11) получаем (деля на $\cos^2\alpha$):

$$a_{12} + (a_{22} - a_{11})\operatorname{tg}\alpha - a_{12}\operatorname{tg}^2\alpha = 0,$$

или

$$a_{12}\operatorname{tg}^2\alpha - (a_{22} - a_{11})\operatorname{tg}\alpha - a_{12} = 0.$$

Решая это квадратное (относительно $\operatorname{tg}\alpha$) уравнение, получаем

$$\operatorname{tg}\alpha = \frac{a_{22} - a_{11} \pm \sqrt{(a_{22} - a_{11})^2 + 4a_{12}^2}}{2a_{12}}. \qquad (8.12)$$

Так как $(a_{22}-a_{11})^2+4a_{12}^2>0$, то по формуле (8.12) нужный нам угол α всегда можно определить.

Обозначим для более короткой записи $a'_{11}=\lambda_1$, $a'_{12}=\lambda_2$. Теперь сформулируем полученный результат.

Поворотом системы координат на угол α, определяемый по формуле (8.12), можно преобразовать квадратичную форму

$$\varphi(x, y)=a_{11}x^2+2a_{12}xy+a_{22}y^2$$

к каноническому виду

$$\varphi'(x', y')=\lambda_1 x'^2+\lambda_2 y'^2.$$

При этом весь многочлен $F(x, y)$ преобразуется к виду

$$F'(x', y')=\lambda_1 x'^2+\lambda_2 y'^2+2a'_{13}x'+2a'_{23}y'+a_{33}. \tag{8.13}$$

Заметим, что оба коэффициента λ_1 и λ_2 не могут одновременно равняться нулю: если бы было $\lambda_1=\lambda_2=0$, то квадратичная форма (8.9) в результате линейного невырожденного преобразования превратилась бы в тождественный нуль, что невозможно.

Итак, возможны **два основных случая**:

I. $\lambda_1\neq 0, \lambda_2\neq 0$.

II. Один из коэффициентов λ_1, λ_2 отличен от нуля, другой равен нулю (*параболический случай*).

Рассмотрим случай I, т.е. первый основной случай: $\lambda_1\neq 0$, $\lambda_2\neq 0$. При переносе начала координат в какую-нибудь точку $O'(x'_0, y'_0)$, т.е. при преобразовании

$$x'=x''+x'_0,$$
$$y'=y''+y'_0,$$

многочлен (8.13) принимает вид

$$F''(x'', y'')=\lambda_1 x''^2+\lambda_2 y''^2+2(\lambda_1 x'_0+a'_{13})x''+$$
$$+2(\lambda_2 y'_0+a'_{23})y''+a'_{33}, \tag{8.14}$$

где свободный член

$$a'_{33}=\lambda_1 x_0'^2+\lambda_2 y_0'^2+2a'_{13}x'_0+2a'_{23}y'_0+a_{33}=F'(x'_0, y'_0).$$

Второй шаг преобразования заключается в следующем. Подберем такие координаты (x'_0, y'_0) нового начала координат, чтобы коэффициенты при x'' и y'' в (8.14) обратились в нуль, т.е. чтобы выполнялись неравенства:

$$\lambda_1 x'_0+a'_{13}=0,\quad \lambda_2 y'_0+a'_{23}=0.$$

Так как $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, то x'_0 и y'_0 можно найти:

$$x'_0 = -\frac{a'_{13}}{\lambda_1}, \quad y'_0 = -\frac{a'_{23}}{\lambda_2}. \tag{8.15}$$

Итак, повернув оси координат на угол α, определяемый формулой (8.12), и перенеся начало системы координат в точку, координаты x'_0, y'_0 которой определяются равенствами (8.15), мы преобразуем уравнение (8.6) к виду

$$\lambda_1 x''^2 + \lambda_2 y''^2 + a'_{33} = 0. \tag{8.16}$$

Здесь также возможны **два случая**:

1) λ_1 и λ_2 разных знаков (так называемый *гиперболический случай*);

2) λ_1 и λ_2 одного и того же знака (так называемый *эллиптический случай*).

Рассмотрим первый случай (г и п е р б о л и ч е с к и й):

а) Пусть $\boldsymbol{a'_{33} \neq 0}$. Очевидно, один из коэффициентов имеет тот же знак, что и a'_{33}; пусть, например, λ_2 и a'_{33} одного знака, тогда λ_1 и a'_{33} противоположны по знаку.

Перепишем уравнение (8.16) в виде

$$\frac{x''^2}{-\frac{a'_{33}}{\lambda_1}} + \frac{y''^2}{-\frac{a'_{33}}{\lambda_2}} = 1.$$

Знаменатель $-\dfrac{a'_{33}}{\lambda_1}$ в первом члене есть положительное число – обозначим его через a^2; знаменатель $-\dfrac{a'_{33}}{\lambda_2}$ отрицателен – его обозначим через $-b^2$. Получаем уравнение

$$\frac{x''^2}{a^2} - \frac{y''^2}{b^2} = 1.$$

Это каноническое уравнение гиперболы.

б) Пусть теперь $\boldsymbol{a'_{33} = 0}$. Уравнение (8.16) приобретает вид

$$\lambda_1 x''^2 + \lambda_2 y''^2 = 0. \tag{8.17}$$

Можно считать, что $\lambda_1 > 0$, $\lambda_2 < 0$ (если это не так, то достаточно умножить обе части уравнения (8.17) на -1). Обозначим $\lambda_1 = a^2$, $\lambda_2 = -b^2$. Получаем уравнение

$$a^2 x''^2 - b^2 y''^2 = 0.$$

Его можно переписать в виде

$$(ax'' + by'')(ax'' - by'') = 0,$$

это уравнение пары прямых

$$ax'' + by'' = 0, \quad ax'' - by'' = 0,$$

пересекающихся в начале координат.

Аналогично рассматриваются эллиптический случай, когда λ_1 и λ_2 одного знака, и параболический случай, когда один из коэффициентов, λ_1 или λ_2, равен нулю.

В эллиптическом случае получим либо эллипс $\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = 1$, либо мнимый эллипс $\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = -1$, либо пару мнимых пересекающихся прямых $\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = 0$.

В параболическом случае получим либо параболу $y''^2 = 2px''$, либо пару параллельных прямых $y''^2 - a^2 = 0$, либо пару мнимых параллельных прямых $y''^2 + a^2 = 0$, либо пару совпадающих прямых $y''^2 = 0$.

Вопросы

1. Что такое полуоси эллипса?
2. Что называется эксцентриситетом эллипса? Что характеризует эксцентриситет эллипса и в каких пределах находится его значение?
3. Сколько осей симметрии имеет эллипс?
4. Какая кривая называется гиперболой?
5. Сколько осей симметрии имеет гипербола?
6. Что такое асимптоты гиперболы? Сколько асимптот имеет гипербола?
7. Каковы основные свойства гиперболы?
8. Что называется параметром параболы? Можно ли, зная параметр параболы, найти расстояние от ее фокуса до вершины?
9. Сколько осей симметрии имеет парабола?
10. Сколько существует различных видов кривых второго порядка?

Глава 9
ПРЯМЫЕ ЛИНИИ И ПЛОСКОСТИ В ПРОСТРАНСТВЕ

9.1. ПЛОСКОСТЬ В ПРОСТРАНСТВЕ

Пусть в пространстве задана система координат $Oxyz$ и пусть плоскость П проходит через точку $M_0(x_0, y_0, z_0)$ перпендикулярно вектору $\overline{N} = (A, B, C)$. Этими двумя условиями определяется единственная плоскость в пространстве $Oxyz$. Вектор $\overline{N}$ называется **нормальным вектором** плоскости П. Выведем уравнение этой плоскости.

Возьмем в плоскости П произвольную точку $M(x, y, z)$. Тогда вектор $\overline{M_0M} = (x - x_0, y - y_0, z - z_0)$ и вектор $\overline{N} = (A, B, C)$ будут взаимно перпендикулярны. Следовательно, их скалярное произведение равно нулю: $\left(\overline{N}, \overline{M_0M}\right) = 0$. Запишем последнее равенство в скалярной форме:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \tag{9.1}$$

Это **уравнение плоскости**, проходящей через точку $M_0(x_0, y_0, z_0)$ перпендикулярно данному вектору $\overline{N} = (A, B, C)$. Из (9.1) получаем

$$Ax + By + Cz - Ax_0 - By_0 - Cz_0 = 0.$$

Обозначив $-Ax_0 - By_0 - Cz_0 = D$, получим **общее уравнение плоскости**:

$$Ax + By + Cz + D = 0. \tag{9.2}$$

Итак, уравнение плоскости есть линейное уравнение, или уравнение первой степени с тремя переменными.

Нетрудно доказать, что *всякое уравнение первой степени с тремя переменными есть уравнение плоскости*.

Известно также, что плоскость однозначно определяется тремя точками, не лежащими на одной прямой. Пусть $M_0(x_0, y_0, z_0)$,

$M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$ – три точки, не лежащие на одной прямой. Тогда векторы $\overline{M_0M_1} = (x_1 - x_0,\ y_1 - y_0,\ z_1 - z_0)$ и $\overline{M_0M_2} = (x_2 - x_0, y_2 - y_0, z_2 - z_0)$ не параллельны одной прямой (не коллинеарны). Пусть $M(x, y, z)$ – произвольная точка плоскости П. Тогда вектор $\overline{M_0M}$ можно разложить по векторам $\overline{M_0M_1}$ и $\overline{M_0M_2}$. Следовательно, эти три вектора $\overline{M_0M_1}$, $\overline{M_0M_2}$ и $\overline{M_0M}$ линейно зависимы и поэтому

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{vmatrix} = 0. \tag{9.3}$$

Это **уравнение плоскости**, проходящей через три точки $M_0(x_0, y_0, z_0)$, $M_1(x_1, y_1, z_1)$ и $M_2(x_2, y_2, z_2)$, не лежащие на одной прямой.

***Пример* 9.1.** Написать уравнение плоскости, проходящей через точки $M_0(1, 2, 1)$, $M_1(3, 3, 1)$, $M_2(2, 3, 2)$.

Решение. Подставим координаты этих точек в уравнение (9.3):

$$\begin{vmatrix} x-1 & y-2 & z-1 \\ 3-1 & 3-2 & 1-1 \\ 2-1 & 3-2 & 2-1 \end{vmatrix} = 0, \quad \begin{vmatrix} x-1 & y-2 & z-1 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 0.$$

$$x - 1 - 2(y - 2) + z - 1 = 0,$$

$$x - 2y + z + 2 = 0.$$

Рассмотрим взаимное расположение двух плоскостей. Даны две плоскости:

$$A_1x + B_1y + C_1z + D_1 = 0,$$

$$A_2x + B_2y + C_2z + D_2 = 0,$$

Их нормальными векторами являются, очевидно, $\overline{N_1} = (A_1, B_1, C_1)$ и $\overline{N_2} = (A_2, B_2, C_2)$.

Угол между этими плоскостями есть угол между $\overline{N_1}$ и $\overline{N_2}$ и определяется по формуле

$$\cos\varphi = \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2}\sqrt{A_2^2 + B_2^2 + C_2^2}}. \tag{9.4}$$

Условием п а р а л л е л ь н о с т и двух плоскостей является условие пропорциональности их нормальных векторов:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}. \qquad (9.5)$$

Условие с о в п а д е н и я плоскостей выглядит так:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} = \frac{D_1}{D_2}. \qquad (9.6)$$

Условие их п е р п е н д и к у л я р н о с т и есть условие $\cos\varphi = 0$, т.е.

$$A_1A_2 + B_1B_2 + C_1C_2 = 0. \qquad (9.7)$$

9.2. ПРЯМАЯ В ПРОСТРАНСТВЕ. ПРЯМАЯ И ПЛОСКОСТЬ В ПРОСТРАНСТВЕ

Пусть прямая L проходит через точку $M_0(x_0, y_0, z_0)$ параллельно вектору $\overline{s} = (l, m, n)$. В этом случае вектор $\overline{s}$ будем называть **направляющим вектором** прямой. Пусть $M(x, y, z)$ – произвольная точка прямой L. Очевидно, векторы $\overline{M_0M} = (x - x_0, y - y_0, z - z_0)$ и $\overline{s}$ пропорциональны. Записав условие их пропорциональности в координатной форме, получим **каноническое уравнение прямой** L:

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}. \qquad (9.8)$$

Из уравнения (9.8) получаем

$$x - x_0 = lt, \quad y - y_0 = mt, \quad z - z_0 = nt,$$

где t – коэффициент пропорциональности. Отсюда

$$x = x_0 + lt, \quad y = y_0 + mt, \quad z = z_0 + nt. \qquad (9.9)$$

Это **параметрические уравнения прямой** L. (Иногда их называют в единственном числе – *параметрическим уравнением прямой*.)

Прямая в пространстве может быть также задана как линия пересечения двух плоскостей, т.е. как множество точек, координаты которых удовлетворяют системе:

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \qquad (9.10)$$

Каноническое уравнение (9.8), впрочем, тоже можно представить как пару уравнений плоскостей, рассматриваемую совместно. Нетрудно вывести каноническое или параметрическое уравнение прямой, заданной в виде (9.10). Для этого достаточно найти какую-нибудь точку $M_0(x_0, y_0, z_0)$, принадлежащую прямой, и направляющий вектор. Координаты точки M_0 найти просто: это любое решение системы (9.10). Например, положив $z_0 = 0$, из системы (9.10) найдем x_0 и y_0 и получим $M_0(x_0, y_0, 0)$. Координатами направляющего вектора $\bar{s}$, как доказывается в аналитической геометрии, могут быть числа

$$l = \begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}, \quad m = \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}, \quad n = \begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}.$$

Рассмотрим теперь взаимное расположение прямой и плоскости в пространстве. Пусть даны прямая L

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}$$

и плоскость П

$$Ax + By + Cz + D = 0.$$

Очевидно, прямая L параллельна плоскости П, когда направляющий вектор прямой $\bar{s} = (l, m, n)$ перпендикулярен нормальному вектору плоскости $\overline{N} = (A, B, C)$, т.е. *условием параллельности прямой L и плоскости* П является условие

$$Al + Bm + Cn = 0. \tag{9.11}$$

Условие пропорциональности этих векторов есть *условие перпендикулярности прямой L и плоскости* П:

$$\frac{A}{l} = \frac{B}{m} = \frac{C}{n}. \tag{9.12}$$

Угол между прямой и плоскостью есть угол между прямой и ее проекцией на плоскость, а это угол, дополнительный к углу между прямой L и нормальным вектором $\overline{N}$ плоскости П:

$$\sin\varphi = \left|\cos \widehat{\overline{N}, \bar{s}}\right| = \frac{|Al + Bm + Cn|}{\sqrt{A^2 + B^2 + C^2}\sqrt{l^2 + m^2 + n^2}}. \tag{9.13}$$

Расстояние от точки до плоскости вычисляется с помощью формулы, аналогичной формуле расстояния от точки до прямой на плоскости [см. (7.15)]. Покажем, что расстояние d от точки $M_0(x_0, y_0, z_0)$ до плоскости

$$Ax + By + Cz + D = 0 \tag{9.2}$$

вычисляется по формуле

$$d = \frac{|Ax_0 + By_0 + Cz_0 + D|}{\sqrt{A^2 + B^2 + C^2}}. \tag{9.14}$$

Напишем уравнение прямой, проходящей через точку $M_0(x_0, y_0, z_0)$ перпендикулярно плоскости (9.2). Для этого воспользуемся параметрическими уравнениями (9.9):

$$x = x_0 + lt, \quad y = y_0 + ml, \quad z = z_0 + nt.$$

Для того чтобы прямая (9.9) была перпендикулярна плоскости (9.2), надо, чтобы ее направляющий вектор $\overline{s} = (l, m, n)$ был параллелен вектору $\overline{N} = (A, B, C)$, т.е. чтобы координаты векторов $\overline{s}$ и $\overline{N}$ были пропорциональны. Проще всего, конечно, взять в качестве $\overline{s}$ вектор $\overline{N}$, т.е. взять $l = A$, $m = B$, $n = C$. Тогда параметрические уравнения (9.9) будут выглядеть так:

$$x = x_0 + At, \quad y = y_0 + Bt, \quad z = z_0 + Ct. \tag{9.9′}$$

Прямая (9.9′) перпендикулярна плоскости (9.2) и проходит через точку M_0. Следовательно, расстояние от точки M_0 до плоскости (9.2) есть расстояние между точкой M_0 и точкой M пересечения прямой (9.9′) с плоскостью (9.2). Найдем координаты этой точки M. Для этого надо решить совместно уравнения (9.2) и (9.9′). Проще всего это сделать, подставив выражения для x, y и z из (9.9′) в (9.2). Получим:

$$A(x_0 + At) + B(y_0 + Bt) + C(z_0 + Ct) + D = 0,$$

$$(A^2 + B^2 + C^2)t + (Ax_0 + By_0 + Cz_0 + D) = 0.$$

Отсюда находим t:

$$t = -\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2}.$$

Это значение t определяет координаты точки M, являющейся основанием перпендикуляра, опущенного из точки M_0 на плоскость (9.2). Подставим найденное t в (9.9′):

$$\begin{aligned} x &= x_0 + A\left(-\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2}\right), \\ y &= y_0 + B\left(-\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2}\right), \\ z &= z_0 + C\left(-\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2}\right). \end{aligned} \tag{9.9″}$$

Расстояние d от точки M_0 до плоскости (9.2) есть длина перпендикуляра M_0M, или, что то же самое, расстояние между точками $M_0(x_0, y_0, z_0)$ и $M(x, y, z)$, т.е.

$$d = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}.$$

Учитывая, что x, y и z определяются равенствами (9.9″), получаем

$$d = \sqrt{\left(A^2 + B^2 + C^2\right)\left(-\frac{Ax_0 + By_0 + Cz_0 + D}{A^2 + B^2 + C^2}\right)^2} =$$
$$= \sqrt{A^2 + B^2 + C^2}\,\frac{\left|Ax_0 + By_0 + Cz_0 + D\right|}{A^2 + B^2 + C^2},$$

или

$$d = \frac{\left|Ax_0 + By_0 + Cz_0 + D\right|}{\sqrt{A^2 + B^2 + C^2}},$$

что и требовалось доказать.

***Пример* 9.2.** Найти расстояние от точки $M_0(1,\ 0,\ 2)$ до плоскости $x + 2y - 2z + 9 = 0$.

Решение:

$$d = \frac{\left|1 + 2\cdot 0 - 2\cdot 2 + 9\right|}{\sqrt{1^2 + 2^2 + (-2)^2}} = \frac{6}{3} = 2.$$

***Пример* 9.3.** Найти расстояние от прямой

$$\frac{x+1}{2} = \frac{y-2}{2} = \frac{z}{1}$$

до плоскости $4x - 2y - 4z + 9 = 0$.

Решение. Прямая параллельна плоскости. Действительно, скалярное произведение ее направляющего вектора и нормального вектора плоскости равно нулю: $2\cdot 4 + 2\cdot(-2) + 1\cdot(-4) = 0$. Следовательно, расстояние от прямой до плоскости равно расстоянию от любой точки M_0 этой прямой до плоскости. Удобнее всего в качестве M_0 взять точку $(-1, 2, 0)$, координаты которой фигурируют в уравнении прямой. Получаем

$$d = \frac{\left|4\cdot(-1) - 2\cdot 2 - 4\cdot 0 + 9\right|}{\sqrt{4^2 + (-2)^2 + (-4)^2}} = \frac{1}{6}.$$

***Пример* 9.4.** Найти расстояние от точки $M_0(1, 2, 3)$ до прямой

$$\frac{x-6}{2}=\frac{y}{-2}=\frac{z-7}{1}.$$

Р е ш е н и е. Напишем уравнение плоскости, которая проходит через данную точку M_0 и перпендикулярна данной прямой, и найдем координаты точки M пересечения прямой и плоскости. Очевидно, M_0M – перпендикуляр, опущенный из точки M_0 на данную прямую. Его длина и есть искомое расстояние.

Уравнение плоскости, проходящей через M_0 и перпендикулярной данной прямой, есть

$$2(x-1)-2(y-2)+1(z-3)=0,$$

или

$$2x-2y+z-1=0. \tag{*}$$

Запишем уравнение данной прямой в параметрической форме:

$$x=6+2t,\quad y=-2t,\quad z=7+t. \tag{**}$$

Найдем точку пересечения прямой (**) и плоскости (*). Для этого сначала подставим x, y и z из (**) в (*) и найдем t:

$$2(6+2t)-2(-2t)+7+t-1=0,$$

$$9t+18=0,\quad t=-2.$$

Теперь, подставляя найденное значение $t=-2$ в (**), получим $x=2$, $y=4$, $z=5$. Итак, точка $M(2, 4, 5)$ есть основание перпендикуляра M_0M. Поэтому

$$d=M_0M=\sqrt{(2-1)^2+(4-2)^2+(5-3)^2}=3.$$

Заметим, что существует другой способ решения этой и подобных ей задач, основанный на понятии векторного произведения векторов, которое здесь не рассматривается.

Вопросы

1. Что называется нормальным вектором плоскости в пространстве?
2. Будет ли прямым угол между плоскостями $3x+y-z=0$ и $x-y+2z+5=0$?

3. Будут ли параллельны плоскости $3x - 2y + z = 0$ и $6x - 3y + 2z + 12 = 0$?
4. Принадлежит ли точка $M_0(1, 2, 3)$ плоскости $2x - 3y + z + 1 = 0$?
5. Чему равно расстояние от начала координат до плоскости $2x - y + 2z + 9 = 0$?
6. Какие координаты имеет точка прямой $\begin{cases} x = 2 + t, \\ y = 1 - 2t, \\ z = 3 + 2t, \end{cases}$ соответствующая значению параметра $t = -1$?
7. Принадлежит ли точка $M_0(1, 3, 2)$ прямой $\begin{cases} x = 3 - t, \\ y = 1 + t, \\ z = -4 + 3t? \end{cases}$ Если принадлежит, то чему равно соответствующее значение параметра t?
8. Будет ли вектор $\overline{a} = (2, -1, 3)$ параллелен прямой $\begin{cases} x = 3 - 4t, \\ y = 1 + 2t, \\ z = 5 - 6t? \end{cases}$
9. Какая точка прямой $\begin{cases} x = 1 + t, \\ y = 2 - 3t, \\ z = -3 + 2t \end{cases}$ соответствует значению параметра $t = 2$?
10. Пересекаются ли прямая $\begin{cases} x = 2 + t, \\ y = 3 + 2t, \\ z = -1 - 2t \end{cases}$ и плоскость $x + 2y - 5z + 2 = 0$? Если пересекаются, то каковы координаты точки пересечения?
11. Параллельна ли прямая $\dfrac{x-1}{2} = \dfrac{y+2}{-1} = \dfrac{z-3}{-2}$ плоскости $3x + 2y + 2z - 7 = 0$?
12. Как определить координаты направляющего вектора прямой, заданной парой плоскостей $\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0? \end{cases}$
13. Как найти расстояние между параллельными плоскостями?
14. Проходит ли прямая $\dfrac{x-1}{2} = \dfrac{y+1}{-1} = \dfrac{z-3}{3}$ через точки $M_1(-1, 0, 0)$ и $M_2(5, -3, 9)$?

Раздел III
ЛИНЕЙНЫЕ МОДЕЛИ В ЭКОНОМИКЕ

Глава 10
ОБЩИЕ СВЕДЕНИЯ О ЛИНЕЙНОМ ПРОГРАММИРОВАНИИ

В этой книге не предусмотрено систематическое изложение линейного программирования. Здесь мы рассмотрим лишь некоторые примеры оптимизационных задач с ограничениями, задаваемыми линейными неравенствами.

10.1. ЗАДАЧА ОБ ИСПОЛЬЗОВАНИИ РЕСУРСОВ

В экономической деятельности предприятия или отрасли постоянно возникают такие ситуации, когда надо определить, как использовать имеющиеся в распоряжении ресурсы, чтобы добиться максимального выпуска продукции. При большом количестве возможных решений этой задачи естественно выбрать наилучшее. Математически эта задача обычно сводится к нахождению наибольшего или наименьшего значения некоторой функции на множестве, определяемом системой неравенств.

Пусть предприятие из m видов ресурсов производит n видов продукции. Пусть для производства одной единицы j-го вида продукции расходуется a_{ij} единиц i-го вида ресурса. Матрица $A = (a_{ij})$ называется *технологической*.

Пусть c_j – удельная прибыль от реализации одной единицы j-й продукции. Эти удельные прибыли образуют вектор $C = (c_1, c_2, \ldots, c_n)$. Тогда произведение $CX = c_1x_1 + c_2x_2 + \ldots + c_nx_n$ есть *величина прибыли, полученной при реализации X единиц произведенной продукции*, где $X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$. Обозначим эту прибыль $f(X)$.

Пусть b_i – число единиц i-го ресурса, имеющегося в распоряжении предприятия. Тогда необходимость учитывать ограниченность запасов ресурсов при составлении планов производства выражается системой неравенств

$$a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n \le b_i, \quad i = 1, 2, ..., m. \tag{10.1}$$

Требуется при данных ресурсах выпустить такую комбинацию товаров, при которой доход предприятия оказался бы максимальным. Иначе говоря, требуется найти наибольшее значение функции $f(X) = c_1x_1 + c_2x_2 + ... + c_nx_n$ при условиях (10.1). Определенная таким образом задача называется **задачей оптимизации**. Ее записывают так:

$$f(X) = c_1x_1 + c_2x_2 + ... + c_nx_n \to \max, \tag{10.2}$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m, \end{cases} \tag{10.3}$$

$$x_1 \ge 0, x_2 \ge 0, ..., x_n \ge 0. \tag{10.4}$$

Функция $f(X)$ называется **целевой функцией**.

Определение. **Допустимым решением (планом)** данной задачи называется любой вектор X, удовлетворяющий системе ограничений (10.3) и условиям неотрицательности (10.4).

Множество допустимых решений задачи образует **область допустимых решений**.

Определение. **Оптимальным решением (планом)** задачи называется такое допустимое решение, при котором целевая функция достигает максимума (минимума).

10.2. ОБЩАЯ ЗАДАЧА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ

Сформулируем в общем виде **задачу линейного программирования**: *найти экстремум линейной функции при линейных же ограничениях на переменные*.

При этом множество значений переменных, удовлетворяющих всем линейным ограничениям задачи, называется **допустимым множеством**, а линейная функция, экстремум которой ищется, – **целевой функцией**.

На практике нередко приходится решать такие задачи линейного программирования, в которых и виды ресурсов, и виды продукции исчисляются сотнями и тысячами. Самый распространенный алгоритм решения задач линейного программирования – это так называемый **симплекс-метод**. Изложение этого метода выходит за рамки нашего курса. Приведем без доказательства лишь две теоремы, составляющие теоретическую основу симплекс-метода (и вообще линейного программирования).

***Теорема* 10.1.** Задача линейного программирования имеет оптимальное решение тогда и только тогда, когда целевая функция ограничена на допустимом множестве в направлении экстремума.

Прежде чем сформулировать вторую теорему, заметим, что допустимое множество, на котором ищется экстремум, – это многогранное тело (при $n = 2$ – многоугольник, при $n = 3$ – многогранник в трехмерном пространстве). Вершины этого многогранного тела называются *угловыми точками*.

***Теорема* 10.2.** Если экстремум целевой функции в задаче линейного программирования достигается, то он достигается в некоторой угловой точке допустимого множества.

Заметим, что угловых точек – конечное число. Симплекс-метод представляет собой направленный перебор угловых точек допустимого множества.

Рассмотрим решение задачи линейного программирования с двумя переменными **графическим методом**.

***Пример* 10.1.** Решить задачу линейного программирования:

$$f(X) = 3x_1 + 2x_2 \to \max,$$

$$\begin{cases} x_1 - x_2 + 2 \ge 0, & (1) \\ 3x_1 - 2x_2 - 6 \le 0, & (2) \\ x_2 - 3 \le 0, & (3) \end{cases}$$

$$x_1 \ge 0, \quad x_2 \ge 0.$$

Р е ш е н и е. На плоскости возьмем прямоугольную систему координат Ox_1x_2 (рис. 10.1).

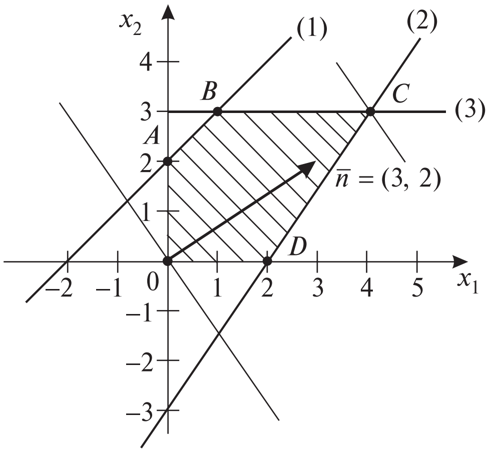


Рис. 10.1

Строим прямую $x_1 - x_2 + 2 = 0$, соответствующую первому ограничению. Чтобы выбрать нужную полуплоскость, надо подставить в неравенство (1) координаты какой-нибудь точки, не лежащей на прямой, например точки $O(0, 0)$: $0 - 0 + 2 > 0$. Получаем верное неравенство. Следовательно, точка O лежит в полуплоскости решений.

Аналогично строим прямые $3x_1 - 2x_2 - 6 = 0$ и $x_2 - 3 = 0$ и выбираем соответствующие полуплоскости.

Учитываем еще условия неотрицательности $x_1 \geq 0$, $x_2 \geq 0$.

Пересечение всех пяти полуплоскостей дает искомое допустимое множество – пятиугольник $OABCD$.

Строим линию уровня $3x_1 + 2x_2 = l$, например, при $l = 0$: $3x_1 + 2x_2 = 0$. Перемещаем линию уровня в направлении нормали. Последняя точка, по которой линия уровня еще пересекает допустимое множество, и будет точкой максимума. В нашем случае это точка C. Ее координаты найдем, решая совместно уравнения прямых, которые пересекаются в точке C:

$$\begin{cases} 3x_1 - 2x_2 - 6 = 0, \\ \qquad\quad x_2 - 3 = 0, \end{cases}$$

получаем $x_1 = 4$, $x_2 = 3$. Вычисляем $f_{\max} = f(4, 3) = 18$.

Нередко бывает, что переменные в задачах линейного программирования являются *целочисленными*. Такие задачи решать сложнее, для них разработаны специальные методы. Но если такая задача содержит две переменные, то ее можно решить графически. Надо двигать линию уровня в направлении экстремума и найти последнюю *целочисленную* точку.

10.3. ЭЛЕМЕНТЫ ТЕОРИИ ДВОЙСТВЕННОСТИ

Центральной частью линейного программирования является теория двойственности. Любой задаче линейного программирования можно поставить в соответствие другую задачу, которая называется **двойственной** (или **сопряженной**). Обе задачи (исходная и двойственная к ней) образуют пару двойственных задач линейного программирования. Каждая из задач является двойственной к другой задаче рассматриваемой пары.

Рассмотрим **задачу об использовании ресурсов**. Пусть для изготовления n видов продукции $P_1, P_2, \ldots, P_n$ используют m видов ресурсов $S_1, S_2, \ldots, S_m$ (это могут быть различные виды сырья, электроэнергия, полуфабрикаты и т.п.). Объем каждого вида ресурсов известен; иначе говоря, известен вектор запасов ресурсов $B = (b_1, b_2, \ldots, b_m)'$. Известна также норма a_{ij} расхода i-го ресурса на производство единицы j-го вида продукции, т.е. известна технологическая матрица $A = (a_{ij})$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$. Кроме того, известна прибыль, получаемая от реализации единицы каждого вида продукции, т.е. известен вектор прибылей $C = (c_1, c_2, \ldots, c_n)$. (Здесь B – вектор-столбец, а C – вектор-строка.)

Производитель составляет план выпуска продукции, обеспечивающий максимальную прибыль. Математическая модель данной задачи, как уже отмечалось, в развернутой форме имеет вид

$$f(X) = c_1x_1 + c_2x_2 + \ldots + c_nx_n \to \max, \tag{10.2}$$

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m, \end{cases} \tag{10.3}$$

$$x_1 \ge 0, x_2 \ge 0, \ldots, x_n \ge 0. \tag{10.4}$$

Здесь x_j, $j = 1, 2, \ldots, n$ – объем производства j-го вида продукции.

В более компактном виде целевая функция и система ограничений обычно записываются так:

$$f=\sum_{j=1}^{n}c_jx_j\to\max,$$

$$\sum_{j=1}^{n}a_{ij}x_j\le b_i,\quad i=1,\ 2,\ \ldots,m,$$

$$x_j\ge 0,\quad j=1,\ 2,\ \ldots,n.$$

Предположим, что имеется покупатель, желающий перекупить частично или полностью ресурсы, запасенные для выполнения рассматриваемой задачи. В рыночной экономике, как правило, категорических отказов не бывает, обычно все определяется ценой и условиями продажи.

Рассмотрим двойственную задачу, решение которой позволит определить условия продажи ресурсов. Обозначим оценку (цену) i-го ресурса через y_i, тогда вектор этих оценок будет иметь вид $Y=(y_1, y_2, ..., y_m)'$.

Затраты на приобретение i-го вида сырья в количестве b_i равны, очевидно, b_iy_i . Покупатель, естественно, желает заплатить поменьше, поэтому для него целевая функция имеет вид

$$\varphi=b_1y_1+b_2y_2+...+b_my_m\to\min. \tag{10.5}$$

Однако производителю, выступающему продавцом, выгодно так оценить свои ресурсы, чтобы суммарная их стоимость, расходуемая на каждое изделие j-й продукции, была не меньше прибыли c_j, которую он получил бы при реализации этого изделия, т.е.

$$a_{1j}y_1+a_{2j}y_2+...+a_{mj}y_m\ge c_j.$$

Следовательно, система ограничений задачи имеет вид

$$\begin{cases}a_{11}y_1+a_{21}y_2+\ldots+a_{m1}y_m\ge c_1,\\ a_{12}y_1+a_{22}y_2+\ldots+a_{m2}y_m\ge c_2,\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ a_{1n}y_1+a_{2n}y_2+\ldots+a_{mn}y_m\ge c_n.\end{cases} \tag{10.6}$$

Кроме того, очевидно, оценки всех видов ресурсов неотрицательны:

$$y_i\ge 0,\quad i=1, 2, ..., m. \tag{10.7}$$

Итак, условия (10.5)–(10.7) определяют новую задачу линейного программирования. Она называется **двойственной** к исходной задаче (10.2)–(10.4).

Рассмотрим подробнее **связь исходной и двойственной задач**:

1) коэффициенты c_j целевой функции исходной задачи являются свободными членами системы ограничений (10.6) двойственной задачи;
2) свободные члены b_i системы ограничений (10.3) исходной задачи являются коэффициентами целевой функции двойственной задачи;
3) матрица коэффициентов системы ограничений двойственной задачи является транспонированной матрицей коэффициентов системы ограничений исходной задачи.

Далее мы узнаем, что если одна из двойственных задач имеет оптимальное решение, то и другая также имеет оптимальное решение (см. теорему 10.4).

Рассмотренная пара задач относится к так называемым симметричным парам двойственных задач. В теории двойственности рассматриваются две *симметричные пары двойственных задач*. Приведем их в матрично-векторной форме (слева – *исходная* задача, справа – *двойственная* ей):

1. $$\begin{array}{ll} f = CX \to \max, & \varphi = YB \to \min, \\ AX \le B, & YA \ge C, \\ X \ge 0; & Y \ge 0. \end{array} \tag{10.8}$$

2. $$\begin{array}{ll} f = CX \to \min, & \varphi = YB \to \max, \\ AX \ge B, & YA \le C, \\ X \ge 0; & Y \ge 0. \end{array} \tag{10.9}$$

Напомним, что здесь

$$C = (c_1, c_2, \ldots, c_n), \quad Y = (y_1, y_2, \ldots, y_m),$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.$$

Заметим, что в теории двойственности используются и несимметричные пары двойственных задач (их также две), но мы их рассматривать не будем.

Сформулируем теперь более четко **правила составления двойственной задачи**:

1. Целевая функция φ двойственной задачи должна оптимизироваться противоположным по сравнению с f образом, т.е. если $f \to \max$, то $\varphi \to \min$, и наоборот.
2. В правых частях ограничений исходной задачи стоят коэффициенты при неизвестных целевой функции двойственной задачи.
3. Матрицы из коэффициентов при неизвестных в левых частях ограничений обеих задач (исходной и двойственной) являются взаимно транспонированными. При этом если исходная задача имеет размер $m \times n$ (m ограничений с n неизвестными), то двойственная – размер $n \times m$.

Пример **10.2.** Составить задачу, двойственную к данной:

$$f = x_1 + 4x_2 + x_3 \to \max,$$

$$\begin{cases} -x_1 + 2x_2 + x_3 \le 4, \\ -2x_1 - 3x_2 - x_3 \le -6, \\ 3x_1 + x_2 + 2x_3 \le 9, \end{cases}$$

$$x_j \ge 0, \quad j = 1, 2, 3.$$

Решение. Умножим правые части ограничений на соответствующие переменные двойственной задачи и составим целевую функцию (которая должна минимизироваться, так как целевая функция исходной задачи максимизируется):

$$4y_1 - 6y_2 + 9y_3 \to \min.$$

Далее транспонируем матрицу коэффициентов при неизвестных в левых частях ограничений исходной задачи, меняем все неравенства на противоположные, а в правых частях записываем соответствующие коэффициенты целевой функции исходной задачи:

$$\begin{cases} -y_1 - 2y_2 + 3y_3 \ge 1, \\ 2y_1 - 3y_2 + y_3 \ge 4, \\ y_1 - y_2 + 2y_3 \ge 1. \end{cases}$$

Окончательно получаем двойственную задачу в виде

$$\varphi = 4y_1 - 6y_2 + 9y_3 \to \min,$$

$$\begin{cases} -y_1 - 2y_2 + 3y_3 \ge 1, \\ 2y_1 - 3y_2 + y_3 \ge 4, \\ y_1 - y_2 + 2y_3 \ge 1, \end{cases}$$

$$y_i \ge 0, \quad i = 1, 2, 3.$$

Пример **10.3.** Составить задачу, двойственную к данной:

$$f = 2x_1 + 3x_3 + 2x_4 \to \min,$$

$$\begin{cases} x_1 - x_2 + 2x_3 + 3x_4 \ge 9, \\ -x_1 - 2x_2 + x_3 + 2x_4 \le -8, \end{cases}$$

$$x_j \ge 0, \quad j = \overline{1, 4}.$$

Решение. Так как целевая функция минимизируется, то все ограничения-неравенства должны иметь вид «≥». Поэтому преобразуем исходную задачу, умножив второе ограничение-неравенство на –1:

$$f = 2x_1 + 3x_3 + 2x_4 \to \min,$$

$$\begin{cases} x_1 - x_2 + 2x_3 + 3x_4 \ge 9, \\ x_1 + 2x_2 - x_3 - 2x_4 \ge 8, \end{cases}$$

$$x_j \ge 0, \quad j = \overline{1, 4}.$$

Составим теперь двойственную задачу подобно тому, как это делалось в предыдущем примере:

$$\varphi = 9y_1 + 8y_2 \to \max,$$

$$\begin{cases} y_1 + y_2 \le 2, \\ -y_1 + 2y_2 \le 0, \\ 2y_1 - y_2 \le 3, \\ 3y_1 - 2y_2 \le 2, \end{cases}$$

$$y_i \ge 0, \quad i = 1, 2.$$

Теоремы двойственности устанавливают связь между оптимальными решениями пары двойственных задач.

Рассмотрим симметричную пару двойственных задач (10.8):

I. $f = CX \to \max,$ **II.** $\varphi = YB \to \min,$

$AX \le B,$ $YA \ge C,$

$X \ge 0.$ $Y \ge 0.$

(Напомним, что если задача I имеет размер $m \times n$, то задача II – размер $n \times m$.)

Отметим **основное неравенство теории двойственности**.

Теорема **10.3.** Пусть X – любое допустимое решение исходной задачи I, а Y – любое допустимое решение двойственной задачи II. Тогда имеет место неравенство

$$f(X) \le \varphi(Y). \qquad (10.10)$$

Доказательство. Так как все переменные в обеих задачах неотрицательные, получаем (с учетом $YA \geq C$):

$$f(X) = CX \leq (YA)X. \qquad (*)$$

В силу ассоциативности умножения матрицы и с учетом $AX \leq B$

$$(YA)X = Y(AX) \leq YB = \varphi(Y). \qquad (**)$$

Соединяя (*) и (**), получаем

$$f(X) \leq \varphi(Y),$$

что и требовалось доказать.

Заметим, в частности, что применительно к задаче, рассмотренной в примере 10.3, неравенство (10.10) означает

$$2x_1 + 3x_3 + 2x_4 \leq 9y_1 + 8y_2.$$

Следствие из основного неравенства: если допустимое множество одной из задач I, II не пусто, то целевая функция другой задачи ограничена в направлении экстремума на своем допустимом множестве.

Действительно, пусть, например, множество D исходной задачи I не пусто, т.е. существует хотя бы одна точка $X^0 \in D$. Тогда, согласно неравенству (10.10), для любой точки Y из допустимого множества задачи II выполняется неравенство $\varphi(Y) \geq f(X^0)$, т.е. все значения функции φ ограничены снизу одним и тем же числом $f(X^0)$.

Теорема **10.4 (основная теорема двойственности).** Если одна из двойственных задач I или II имеет оптимальное решение, то и другая задача имеет оптимальное решение, причем экстремальные значения целевых функций равны:

$$\max f = \min \varphi. \qquad (10.11)$$

(Примем эту теорему без доказательства.)

Одним из важных следствий основной теоремы двойственности является **критерий оптимальности допустимых решений**. Рассмотрим его. Пусть X^0 и Y^0 – допустимые решения исходной и двойственной задач I и II. Для того чтобы эти решения были оптимальными, необходимо и достаточно выполнение равенства

$$f(X^0) = \varphi(Y^0). \qquad (10.12)$$

Доказательство. 1. *Необходимость*. Пусть X^0 и Y^0 – оптимальные решения. Тогда $f(X^0) = \max f$, $\varphi(Y^0) = \min \varphi$ и равенство (10.12) следует из основной теоремы двойственности.

2. *Достаточность.* Пусть выполняется равенство (10.12) и пусть X – произвольная точка, принадлежащая допустимому множеству исходной задачи. Тогда, в силу основного неравенства (10.10), имеем $f(X) \le \varphi(Y^0) = f(X^0)$. Следовательно, $f(X^0) = \max f$, т.е. X^0 – точка максимума. Аналогично доказывается, что точка Y^0, для которой выполняется равенство (10.12), есть точка минимума.

Теорема **10.5 (вторая теорема двойственности).** Для того чтобы допустимые решения $X = (x_1, x_2, ..., x_n)'$ и $Y = (y_1, y_2, ..., y_m)$ являлись оптимальными решениями пары двойственных задач I и II, необходимо и достаточно, чтобы выполнялись следующие равенства:

$$x_j\left(\sum_{i=1}^{m} a_{ij} y_i - c_j\right) = 0, \quad j = 1, 2, \ldots, n; \tag{10.13}$$

$$y_i\left(\sum_{j=1}^{n} a_{ij} x_j - b_i\right) = 0, \quad i = 1, 2, \ldots, m. \tag{10.14}$$

(Эту теорему примем без доказательства.)

Выясним смысл равенств (10.13) и (10.14). Например, второе из них означает, что если при подстановке оптимального решения в систему ограничений (10.3) i-е ограничение исходной задачи выполняется как строгое неравенство, то i-я координата оптимального решения двойственной задачи *равна нулю*. Напротив, если i-я координата оптимального решения двойственной задачи *отлична от нуля*, то i-е ограничение исходной задачи при подстановке оптимального решения обращается в равенство.

Условия (10.13) и (10.14) устанавливают в некотором смысле равновесие между задачами I и II. Поэтому теорему 10.5 называют также **теоремой равновесия**.

Пример **10.4.** Решить задачу

$$4x_1 + 3x_2 - 30x_3 \to \min,$$

$$\begin{cases} x_1 \qquad - 6x_3 \ge 1, \\ \qquad x_2 - 5x_3 \ge 2, \end{cases}$$

$$x_1, x_2, x_3 \ge 0.$$

Решение. В данной задаче три переменных. Решить ее графически, подобно тому, как это сделано в примере 10.1, нельзя. Составим для этой задачи двойственную, решим ее графическим методом, а затем, применяя вторую теорему двойственности, решим исходную задачу.

Итак, составляем двойственную задачу:

$$y_1 + 2y_2 \to \max,$$

$$\begin{cases} y_1 \phantom{{}-5y_2} \le 4, & (1) \\ \phantom{-6y_1 - {}} y_2 \le 3, & (2) \\ -6y_1 - 5y_2 \le -30, & (3) \end{cases}$$

$$y_1, y_2 \ge 0.$$

Решим ее графическим методом. На рис. 10.2 изображены область допустимых решений, нормаль $\bar{n} = (1, 2)$ и оптимальное решение – точка (4, 3).

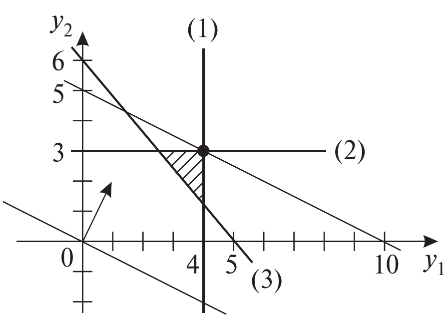


Рис. 10.2

Теперь находим решение исходной задачи с помощью второй теоремы двойственности. Так как третье ограничение двойственной задачи есть строгое неравенство при $y_1 = 4$, $y_2 = 3$, то $x_3 = 0$. Далее, так как $y_1 > 0$, $y_2 > 0$, то $x_1 - 6x_3 = 1$, $x_2 - 5x_3 = 2$, откуда $x_1 = 1$, $x_2 = 2$. Следовательно, оптимальное решение исходной задачи есть точка (1, 2, 0).

***Пример* 10.5.** С помощью второй теоремы двойственности решить задачу

$$2x_1 + 6x_2 + 3x_3 \to \min,$$

$$\begin{cases} 3 + x_1 - 3x_2 \le 0, \\ x_1 - 2x_2 + x_3 \ge 2, \end{cases}$$

$$x_1, x_2, x_3 \ge 0.$$

Р е ш е н и е. Составим для этой задачи двойственную. Сначала приведем все неравенства к виду «≥», так как целевая функция минимизируется:

$$2x_1 + 6x_2 + 3x_3 \to \min,$$

$$\begin{cases} -x_1 + 3x_2 \geq 3, \\ x_1 - 2x_2 + x_3 \geq 2, \end{cases}$$

$$x_1, x_2, x_3 \geq 0.$$

Напишем теперь двойственную задачу:

$$3y_1 + 2y_2 \to \max,$$

$$\begin{cases} -y_1 + y_2 \leq 2, \\ 3y_1 - 2y_2 \leq 6, \\ y_2 \leq 3, \end{cases}$$

$$y_1, y_2 \geq 0.$$

Эта задача совпадает с задачей, рассмотренной в примере 10.1 (только там переменные x_1, x_2, а здесь y_1, y_2). Воспользуемся ее решением: $y_1 = 4$, $y_2 = 3$.

Так как первое ограничение двойственной задачи выполняется как строгое неравенство, то $x_1 = 0$. Так как $y_1 > 0$, $y_2 > 0$, то

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

откуда $x_2 = 1$, $x_3 = 4$.

Итак, оптимальное решение исходной задачи есть (0, 1, 4).

Рассмотрим теперь задачу с четырьмя переменными, которую также можно свести к двойственной задаче, решаемой графически.

***Пример* 10.6.** Для данной задачи составить двойственную, решить ее и с помощью второй теоремы двойственности найти решение ее исходной задачи:

$$2x_1 + x_2 - 3x_3 + x_4 \to \max,$$

$$\begin{cases} x_1 + 2x_2 - x_4 \leq 4, \\ x_1 - x_2 + x_3 + 3x_4 \leq 1, \end{cases}$$

$$x_1, x_2, x_3, x_4 \geq 0.$$

Р е ш е н и е. Составим двойственную задачу:

$$4y_1 + y_2 \to \min,$$

$$\begin{cases} y_1 + y_2 \ge 2, \\ 2y_1 - y_2 \ge 1, \\ y_2 \ge -3, \\ -y_1 + 3y_2 \ge 1, \end{cases}$$

$$y_1, y_2 \ge 0.$$

Решив ее графическим методом (проделайте это самостоятельно), находим точку (1, 1).

Теперь применим вторую теорему двойственности для нахождения решения исходной задачи. Мы видим, что третье и четвертое ограничения двойственной задачи выполняются как строгие неравенства. Следовательно, $x_3 = 0$, $x_4 = 0$. Кроме того, так как $y_1 > 0$, $y_2 > 0$, то

$$\begin{cases} x_1 + 2x_2 - x_4 = 4, \\ x_1 - x_2 + x_3 + 3x_4 = 1, \end{cases}$$

откуда (с учетом $x_3 = x_4 = 0$) получаем $x_1 = 2$, $x_2 = 1$. Следовательно, оптимальное решение исходной задачи есть точка (2, 1, 0, 0). При этом, очевидно,

$$f(X)_{\max} = \varphi(Y)_{\min} = 5.$$

Вопросы

1. Что такое технологическая матрица?
2. Что называется задачей оптимизации? Как она записывается?
3. Что называется допустимым решением задачи оптимизации? Какое решение называется оптимальным?
4. Как формулируется общая задача линейного программирования?
5. Как связаны между собой исходная и двойственная задачи линейного программирования?
6. Пусть исходная задача линейного программирования имеет размер $m \times n$. Каков размер двойственной задачи?

7. В чем заключается основное неравенство теории двойственности?
8. Как формулируется основная теорема двойственности?
9. Может ли задача линейного программирования с двумя переменными быть двойственной к задаче с пятью переменными?

Глава 11
МОДЕЛЬ ЛЕОНТЬЕВА. ЛИНЕЙНАЯ МОДЕЛЬ ОБМЕНА

11.1. МОДЕЛЬ ЛЕОНТЬЕВА

Эффективное функционирование многоотраслевого хозяйства возможно лишь при наличии б а л а н с а между отраслями. Предположим, что вся производственная сфера хозяйства представляет собой n так называемых чистых отраслей.

Чистая отрасль – это условное понятие, некоторая часть хозяйства, относительно цельная, производящая свой однородный продукт и определяемая только видом выпускаемого продукта (таковы, например, добыча сырья, энергетика, сельское хозяйство и т.д.). Часть продукции идет на внутрипроизводственное потребление (как данной отраслью, так и другими отраслями), а другая часть предназначена для потребления в непроизводственной сфере.

Рассмотрим процесс производства за некоторый период времени (обычно таким промежутком служит год).

Введем следующие обозначения:

x_i – общий объем продукции i-й отрасли (валовой выпуск);

x_{ij} – объем продукции i-й отрасли, расходуемый j-й отраслью в процессе производства;

y_i – объем продукции i-й отрасли, предназначенный для потребления в непроизводственной сфере (объем конечного потребления).

Так как валовой объем продукции i-й отрасли равен сумме объемов потребления в производственной и непроизводственной сферах, то

$$x_i = \sum_{j=1}^{n} x_{ij} + y_i, \quad i = 1, 2, ..., n. \tag{11.1}$$

Уравнения (11.1) называются **соотношениями баланса**.

Поскольку продукция разных отраслей имеет разные измерения, будем рассматривать **стоимостный межотраслевой баланс**, когда все величины, входящие в (11.1), имеют стоимостное выражение.

Математическая модель, позволяющая анализировать связь между отраслями, разработана в 1936 г. американским экономистом В. Леонтьевым. Анализируя американскую экономику в период перед Второй мировой войной, он обратил внимание на следующее важное обстоятельство: в течение длительного времени величины

$$a_{ij} = \frac{x_{ij}}{x_j}, \quad i = 1,\ 2,\ \ldots, n \tag{11.2}$$

меняются очень незначительно и могут рассматриваться как постоянные числа. Это объясняется тем, что технология производства остается практически постоянной на протяжении довольно длительного времени.

Сказанное выше позволяет сделать следующее допущение: для выпуска продукции j-й отрасли объема x_j необходимо затратить продукцию i-й отрасли объема $a_{ij}x_j$, где a_{ij} – постоянный коэффициент. Это допущение называется **гипотезой линейности**. Согласно этой гипотезе

$$x_{ij} = a_{ij}x_j, \quad i = 1, 2, ..., n. \tag{11.3}$$

Согласно гипотезе линейности числа a_{ij} постоянны, они называются **коэффициентами прямых затрат**.

Теперь уравнения (11.1) с учетом (11.3) можно записать в виде системы:

$$\begin{cases} x_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n + y_1, \\ x_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n + y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n + y_n. \end{cases} \tag{11.4}$$

Введем вектор валового выпуска $\overline{x}$, матрицу прямых затрат A и вектор конечного потребления (или конечного продукта) $\overline{y}$:

$$\overline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad \overline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}. \tag{11.5}$$

Теперь система (11.4) в матричной форме имеет вид

$$\bar{x} = A\bar{x} + \bar{y}. \tag{11.6}$$

Уравнение (11.6) называется **уравнением линейного межотраслевого баланса**. Вместе с интерпретацией (11.5) вектора $\bar{x}$, матрицы A и вектора $\bar{y}$ это уравнение называют **моделью Леонтьева**.

Основная задача межотраслевого баланса: найти такой вектор валового выпуска $\bar{x}$, который при известной матрице прямых затрат A обеспечивает заданный вектор конечного потребления $\bar{y}$. Иначе говоря: сколько следует произвести продукции разных видов, чтобы обеспечить заданный уровень конечного потребления?

Очевидно, задача сводится к тому, чтобы решить уравнение (систему линейных уравнений) (11.6) с неизвестным вектором $\bar{x}$, известной матрицей A и при заданном векторе $\bar{y}$.

Условимся для краткости называть матрицу A *неотрицательной*, если все ее компоненты неотрицательны. В этом случае пишем $A \geq 0$. Аналогично определяется *неотрицательный вектор*.

В сформулированной выше задаче, очевидно, $A \geq 0$, $\bar{y} \geq 0$ (это непосредственно следует из экономического смысла A и $\bar{y}$). Искомый вектор $\bar{x}$ также должен быть неотрицательным: $\bar{x} \geq 0$.

Перепишем уравнение (11.6) в виде

$$(E - A)\bar{x} = \bar{y}. \tag{11.7}$$

Если $(E - A)$ – невырожденная матрица, то существует обратная к ней матрица $(E - A)^{-1}$ и существует (и притом единственное) решение уравнения (11.7):

$$\bar{x} = (E - A)^{-1}\bar{y}. \tag{11.8}$$

Матрица $S = (E - A)^{-1}$ называется **матрицей полных затрат**.

Выясним **экономический смысл матрицы полных затрат** $S = (s_{ij})$. Рассмотрим единичные векторы конечного продукта:

$$\bar{y}_1 = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \ \bar{y}_2 = \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \ \dots, \ \bar{y}_n = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}.$$

Для них из (11.8) получаем соответствующие векторы валового выпуска:

$$\bar{x}_1 = \begin{pmatrix} s_{11} \\ s_{21} \\ \dots \\ s_{n1} \end{pmatrix}, \quad \bar{x}_2 = \begin{pmatrix} s_{12} \\ s_{22} \\ \dots \\ s_{n2} \end{pmatrix}, \quad \dots, \quad \bar{x}_n = \begin{pmatrix} s_{1n} \\ s_{2n} \\ \dots \\ s_{nn} \end{pmatrix}.$$

Следовательно, каждый элемент s_{ij} матрицы S есть *величина валового выпуска продукции i-й отрасли, необходимого для обеспечения выпуска единицы конечного продукта j-й отрасли*.

Матрица $A \geq 0$ называется **продуктивной**, если для любого вектора $\bar{y} \geq 0$ существует решение $\bar{x} \geq 0$ уравнения (11.6). В этом случае и модель Леонтьева называется **продуктивной**.

Оказывается, нет необходимости требовать существования решения $\bar{x} \geq 0$ уравнения (11.6) *для любого* вектора $\bar{y} \geq 0$. Достаточно установить существование такого решения *хотя бы для одного* вектора $\bar{y} \geq 0$, как показывает следующая теорема (приведем ее без доказательства).

***Теорема* 11.1.** Если для $A \geq 0$ и для некоторого вектора $\bar{y} \geq 0$ уравнение (11.6) имеет решение $\bar{x} \geq 0$, то матрица A продуктивна.

Существуют различные критерии продуктивности. Приведем два из них.

Первый критерий продуктивности. Матрица $A \geq 0$ продуктивна тогда и только тогда, когда матрица $(E - A)^{-1}$ существует и неотрицательна.

Второй критерий продуктивности. Матрица $A \geq 0$ продуктивна, если сумма элементов любого ее столбца не превышает единицы:

$$\sum_{i=1}^{n} a_{ij} \leq 1.$$

11.2. ЛИНЕЙНАЯ МОДЕЛЬ ОБМЕНА

Понятия собственного вектора и собственного значения матрицы (см. § 5.3) применимы, в частности, для анализа процесса взаимных закупок товаров.

Рассмотрим следующий вопрос: какими должны быть соотношения между бюджетами стран, торгующих между собой, чтобы торговля была взаимовыгодной, т.е. практически бездефицитной для каждой из этих стран. Для ответа на этот вопрос рассмотрим **линейную модель обмена**, или **модель международной торговли**.

Пусть имеется n стран. Их национальные бюджеты обозначим соответственно через $x_1, x_2, ..., x_n$. Пусть a_{ij} – доля бюджета x_j, которую j-я страна тратит на закупку товаров у i-й страны. Будем полагать, что весь национальный бюджет каждой страны расходуется только на закупку товаров либо внутри страны, либо вне ее, т.е. справедливо равенство

$$\sum_{i=1}^{n} a_{ij} = 1, \ \ j = 1, 2, \ldots, n. \qquad (11.9)$$

Рассмотрим матрицу, составленную из коэффициентов a_{ij}:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \qquad (11.10)$$

она называется **структурной матрицей торговли**.

В соответствии с (11.9) сумма элементов любого столбца матрицы A равна единице.

Для i-й страны выручка от внутренней и внешней торговли составит

$$p_i = a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n. \qquad (11.11)$$

Условие сбалансированности торговли формулируется так: выручка от торговли каждой страны должна быть не меньше ее национального бюджета, т.е. торговля должна быть бездефицитной для каждой страны: $p_i \geq x_i$ для всех i, или

$$a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n \geq x_i, \ \ i = 1, 2, ..., n. \qquad (11.12)$$

***Теорема* 11.2.** Условием бездефицитной торговли являются равенства

$$p_i = x_i, \quad i = 1, 2, ..., n. \qquad (11.13)$$

Доказательство. Предположим противное, т.е. что $p_i > x_i$ для некоторого i. Тогда выполняется строгое неравенство:

$$\sum_{i=1}^{n} p_i > \sum_{i=1}^{n} x_i. \qquad (11.14)$$

Распишем это неравенство с учетом (11.11):

$$(a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n) + (a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n) + \ldots +$$
$$+ (a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n) > x_1 + x_2 + \ldots + x_n.$$

Группируя слагаемые, получаем

$$x_1(a_{11} + a_{21} + \ldots + a_{n1}) + x_2(a_{12} + a_{22} + \ldots + a_{n2}) + \ldots +$$
$$+ x_n(a_{1n} + a_{2n} + \ldots + a_{nn}) > x_1 + x_2 + \ldots + x_n.$$

Из (11.9) следует, что все суммы, стоящие в скобках, равны единице. Получаем противоречие:

$$x_1 + x_2 + \ldots + x_n > x_1 + x_2 + \ldots + x_n.$$

Следовательно, строгое неравенство $p_i > x_i$ невозможно ни при каком i. Поэтому все неравенства $p_i \geq x_i$ принимают вид равенств:

$$p_i = x_i, \quad i = 1, 2, \ldots, n.$$

Введем вектор бюджетов:

$$\bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}.$$

Тогда система равенств (11.13) с учетом (11.11) принимает вид

$$A\bar{x} = \bar{x}. \tag{11.15}$$

Это уравнение означает, что собственный вектор матрицы A, отвечающий собственному значению $\lambda = 1$, состоит из бюджетов стран, ведущих сбалансированную торговлю. Итак, задача свелась к нахождению собственного вектора структурной матрицы торговли, отвечающего собственному значению $\lambda = 1$.

***Пример* 11.1.** Структурная матрица торговли трех стран имеет вид

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

При каких условиях достигается сбалансированность торговли этих стран?

Р е ш е н и е. Уравнение (11.15) перепишем в виде $(A - E)\bar{x} = \bar{0}$:

$$\begin{pmatrix} -0{,}8 & 0{,}3 & 0{,}5 \\ 0{,}4 & -0{,}6 & 0{,}3 \\ 0{,}4 & 0{,}3 & -0{,}8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Ранг этой системы равен двум. Решая ее, получаем

$$\begin{cases} x_1 = \dfrac{13}{12} x_3, \\ x_2 = \dfrac{11}{9} x_3. \end{cases}$$

Положив $x_3 = 36$ (чтобы не было дробных чисел), получаем вектор

$$\bar{x} = (39,\ 44,\ 36),$$

который можно взять в качестве собственного вектора.

Итак, сбалансированность торговли этих стран достигается при условии, что их бюджеты находятся в соотношении

$$x_1 : x_2 : x_3 = 39 : 44 : 36.$$

Вопросы

1. В чем заключается гипотеза линейности?
2. Какой вид имеет уравнение линейного межотраслевого баланса?
3. Что называется моделью Леонтьева?
4. Какая матрица называется матрицей полных затрат? Каков экономический смысл этой матрицы?
5. Какая матрица называется продуктивной? В каком случае модель Леонтьева называется продуктивной?
6. Каковы критерии продуктивности матрицы?
7. Что такое структурная матрица торговли? Чем характеризуются столбцы этой матрицы?
8. В чем заключается условие сбалансированности торговли?

Раздел IV
ВВЕДЕНИЕ В АНАЛИЗ

Глава 12
ЭЛЕМЕНТЫ ТЕОРИИ МНОЖЕСТВ

12.1. ПОНЯТИЕ МНОЖЕСТВА

Множеством в математике называют совокупность объектов, объединенных по определенному признаку. Понятие множества принадлежит к числу первичных простейших математических понятий; оно не определяется, но может быть пояснено с помощью примеров.

Примерами множеств являются: множество точек данной линии, множество всех решений данного уравнения, множество предприятий некоторой отрасли. Объекты, составляющие данное множество, называются его **элементами**, или **точками**. Обычно множества обозначают прописными буквами, а входящие в них элементы – строчными. Задать множество можно перечислением его элементов или указав характеристическое свойство его элементов, т.е. такое свойство, которым обладают все элементы этого множества и только они. Например, $A = \{1, 2, 3, 4\}$ – множество, элементами которого являются числа 1, 2, 3, 4 и только они. Или $A = \{x{:}\ x > 0\}$ – множество всех положительных чисел.

Если a есть элемент множества A, то это записывают так: $a \in A$; если a не является элементом множества A, то пишут $a \notin A$.

Множество, не содержащее ни одного элемента, называется **пустым** и обозначается $\varnothing$. Например, множество действительных корней уравнения $x^2 + 1 = 0$ есть пустое множество.

Множество называется **конечным**, если число его элементов конечно, и **бесконечным**, если число его элементов бесконечно.

Множество A называется **подмножеством** множества B, если каждый элемент множества A является в то же время элементом множества B. Это записывают так: $A \subset B$. Пустое множество по определению является подмножеством любого множества.

Два множества называются **равными**, если они состоят из одних и тех же элементов: $A = B$ означает, что одновременно $A \subset B$ и $B \subset A$.

12.2. ОПЕРАЦИИ НАД МНОЖЕСТВАМИ. СЧЕТНЫЕ И НЕСЧЕТНЫЕ МНОЖЕСТВА

Определение. **Объединением** (или **суммой**) **двух множеств** A и B называется множество C, состоящее из всех элементов, принадлежащих хотя бы одному из данных множеств:

$C = A \cup B$.

Определение. **Пересечением двух множеств** A и B называется множество C, состоящее из всех элементов, которые принадлежат одновременно и множеству A, и множеству B:

$C = A \cap B$.

Определение. **Разностью между множеством** A и **множеством** B называется множество C, состоящее из всех элементов множества A, которые не принадлежат множеству B:

$C = A \setminus B$;

разность между множеством A и его подмножеством B называется **дополнением множества** B **в множестве** A.

***Пример* 12.1.** Даны множества $A = \{1, 2, 3, 9\}$ и $B = \{1, 3, 5, 7\}$. Найти объединение, пересечение и разность множеств A и B.

Решение: $A \cup B = \{1, 2, 3, 5, 7, 9\}$, $A \cap B = \{1, 3\}$, $A \setminus B = \{2, 9\}$.

Свойства операций $\cup$ и $\cap$:

1) $A \cup B = B \cup A$ и $A \cap B = B \cap A$ (коммутативность);

2) $(A \cup B) \cup C = A \cup (B \cup C)$ и $(A \cap B) \cap C = A \cap (B \cap C)$ (ассоциативность);

3) $A \cup A = A$ и $A \cap A = A$ (идемпотентность);

4) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ и $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (дистрибутивность).

Множество называется **счетным**, если все его элементы можно занумеровать натуральными числами. В противном случае множество называется **несчетным**. В более подробном курсе математики доказывается, в частности, что множество всех рациональных чисел счетно, а множество всех действительных чисел, заключенных между 0 и 1, – несчетно.

12.3. ЧИСЛОВЫЕ МНОЖЕСТВА. ЧИСЛОВАЯ ПРЯМАЯ

Действительные числа хорошо известны из школьного курса математики. Множество всех действительных чисел обычно обозначают **R**. Известны также следующие подмножества множества **R**: **N** – множество всех натуральных чисел, **Z** – целых, **Q** – рациональных, **I** – иррациональных. Очевидно, $\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R}$, $\mathbf{I} \subset \mathbf{R}$, $\mathbf{Q} \cap \mathbf{I} = \varnothing$, $\mathbf{R} = \mathbf{Q} \cup \mathbf{I}$. (Напомним, что число называется **рациональным**, если его можно представить в виде $\frac{m}{n}$, где m и n – целые числа, $n \neq 0$. Всякое действительное число, не являющееся рациональным, называется **иррациональным**.) Отметим, что множества **N**, **Z**, **Q** счетны, а множества **I** и **R** несчетны.

Числовой прямой (или **числовой осью**) называется прямая, на которой выбраны начало отсчета, положительное направление и масштаб, т.е. единица длины (рис. 12.1).

Рис. 12.1. Числовая прямая

Между множеством **R** всех действительных чисел и множеством всех точек числовой прямой существует **взаимно однозначное соответствие**: каждому действительному числу соответствует одна определенная точка числовой прямой, и наоборот, каждой точке прямой – одно определенное действительное число. Установив это взаимно однозначное соответствие, мы отождествляем точки числовой прямой и соответствующие действительные числа. Понятия «число x» и «точка x» становятся неразличимыми. Поэтому часто вместо «точка x» говорят «число x» и наоборот.

Отметим наиболее употребительные числовые множества. Пусть a и b – два числа, причем $a < b$, тогда:

- **отрезок** $[a, b]$ – это множество всех чисел x, удовлетворяющих неравенству $a \leq x \leq b$;
- **интервал** (a, b) – множество всех чисел, удовлетворяющих строгому неравенству $a < x < b$;
- **полуинтервалы** $(a, b]$ и $[a, b)$ – числовые множества, характеризующиеся неравенствами соответственно $a < x \leq b$ и $a \leq x < b$.

Интервалы и полуинтервалы могут быть, в частности, бесконечными: $(-\infty, a)$, $(b, +\infty)$, $(-\infty, +\infty)$, $(-\infty, a]$, $[b, +\infty)$. (Очевидно, интервал $(-\infty, +\infty)$ есть вся числовая прямая.)

Все перечисленные множества принято объединять единым термином **промежуток**. Говоря «промежуток», мы имеем в виду либо отрезок, либо интервал, либо полуинтервал.

Окрестностью точки a называется всякий интервал, содержащий эту точку a. Интервал $(a-\varepsilon, a+\varepsilon)$ называется ε-**окрестностью** точки a.

12.4. МОДУЛЬ ДЕЙСТВИТЕЛЬНОГО ЧИСЛА

Определение. **Модулем** (или **абсолютной величиной**) действительного числа x называется само это число x, если оно неотрицательно, и число, противоположное числу x, если x отрицательно:

$$|x| = \begin{cases} x, & \text{если } x \geq 0, \\ -x, & \text{если } x < 0. \end{cases}$$

Очевидно, $|x| \geq 0$.

Известны, в частности, следующие **свойства модулей**:

$$|x+y| \leq |x| + |y|; \quad |x-y| \geq |x| - |y|;$$

$$|xy| = |x| \cdot |y|; \quad \left|\frac{x}{y}\right| = \frac{|x|}{|y|}.$$

Модуль разности двух чисел $|x-a|$ есть расстояние между точками x и a числовой прямой, в частности, $|x|$ есть расстояние от точки 0 до точки x. Множество точек x, удовлетворяющих условию $|x-a| < \varepsilon$, есть, очевидно, ε-окрестность точки a.

12.5. МЕТОД МАТЕМАТИЧЕСКОЙ ИНДУКЦИИ

Метод математической индукции – один из самых важных методов математических доказательств. Он применяется для доказательства утверждений, зависящих от натурального числа n.

Метод математической индукции заключается в следующем: чтобы доказать некоторое утверждение, зависящее от натурального числа n, надо:

1) проверить его справедливость при $n = 1$ (или при наименьшем n, при котором утверждение имеет смысл);

2) предполагая справедливость утверждения для $n = k$, доказать его справедливость для $n = k + 1$.

Затем делается вывод о справедливости данного утверждения для любого n.

***Пример* 12.2.** Доказать, что $1 + 3 + 5 + ... + (2n - 1) = n^2$.

Решение. Обозначим $1 + 3 + 5 + ... + (2n - 1) = S_n$.

1. Очевидно, при $n = 1$ формула верна: $1 = 1^2$.

2. Предполагая $S_k = k^2$, докажем, что $S_{k+1} = (k + 1)^2$. Действительно,

$$S_{k+1} = S_k + [2(k + 1) - 1] = k^2 + (2k + 1) = (k + 1)^2,$$

что и требовалось доказать. На основании метода математической индукции делаем вывод, что $S_n = n^2$.

***Пример* 12.3.** Доказать справедливость формулы сложных процентов $S_n = S_0\left(1+\frac{i}{100}\right)^n$, где S_0 – первоначальный капитал, i – процентная ставка, n – число периодов начисления.

Решение. 1. При $n = 1$ имеем $S_1 = S_0 + S_0\frac{i}{100} = S_0\left(1+\frac{i}{100}\right)$, т.е. формула верна.

2. Предполагаем, что $S_k = S_0\left(1+\frac{i}{100}\right)^k$. Докажем, что $S_{k+1} = S_0\left(1+\frac{i}{100}\right)^{k+1}$. Имеем

$$S_{k+1} = S_0\left(1+\frac{i}{100}\right)^k + S_0\left(1+\frac{i}{100}\right)^k\frac{i}{100} =$$

$$= S_0\left(1+\frac{i}{100}\right)^k\left(1+\frac{i}{100}\right) = S_0\left(1+\frac{i}{100}\right)^{k+1},$$

что и требовалось доказать. Итак, формула сложных процентов верна.

12.6. СОЕДИНЕНИЯ И БИНОМ НЬЮТОНА

Соединения

Пусть X – множество, состоящее из n элементов: $X = \{x_1, x_2, \ldots, x_n\}$. Из элементов множества X будем образовывать различные подмножества, которые называются **соединениями**. В зависимости от того, содержит ли соединение все элементы множества X или часть их, и от того, играет ли роль порядок расположения элементов, различают **три вида соединений**:

- размещения;
- перестановки;
- сочетания.

Определение. Соединения, содержащие каждое m элементов из данных n элементов множества X, которые отличаются друг от друга либо самими элементами, либо их порядком, называют **размещениями** из n элементов по m.

Например, при составлении расписания на определенный день в классе, где изучаются 10 предметов и каждый день по 5 уроков, рассматривают размещения по 5 элементов из 10.

Число размещений из n элементов по m обозначают A_n^m. Справедлива формула

$$A_n^m = n(n-1)(n-2)\cdots[n-(m-1)], \quad 1 \le m \le n. \tag{12.1}$$

Определение. Соединения, каждое из которых содержит n элементов множества X и которые отличаются только порядком элементов, называются **перестановками** из n элементов.

Число перестановок из n элементов обозначают символом P_n.

Перестановки являются частным случаем размещений, когда $m = n$. По формуле (12.1)

$$P_n = A_n^n = n(n-1)(n-2)\cdots 1, \quad \text{или} \quad P_n = 1\cdot 2\cdots(n-1)n. \tag{12.2}$$

Произведение $1\cdot 2\cdots(n-1)n$ называют «n факториал» и обозначают $n!$. При $n = 0$ по определению полагают $0! = 1$. Формулу (12.2) можно записать теперь так:

$$P_n = n!. \tag{12.3}$$

Пользуясь символом $n!$, можно формулу (12.1) представить в виде

$$A_n^m = \frac{n!}{(n-m)!}. \tag{12.1$'$}$$

Определение. Соединения, содержащие каждое m элементов из данных n элементов множества X, которые отличаются друг от друга хотя бы одним элементом, называются **сочетаниями** из n элементов по m.

Порядок расположения элементов внутри сочетания во внимание не принимается. Число сочетаний из n элементов по m обозначается C_n^m. Из определения следует, что

$$A_n^m = C_n^m P_m.$$

Поэтому

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n(n-1)(n-2)\cdots[n-(m-1)]}{m!}, \tag{12.4}$$

или

$$C_n^m = \frac{n!}{m!(n-m)!}. \tag{12.4′}$$

Можно доказать, что

$$C_n^{m+1} + C_n^m = C_{n+1}^{m+1}. \tag{12.5}$$

Формула бинома Ньютона

Для любого натурального числа n справедлива формула

$$(a+b)^n = C_n^0 a^n + C_n^1 a^{n-1} b + \ldots + C_n^m a^{n-m} b^m + \ldots + C_n^n b^n, \tag{12.6}$$

которая называется **формулой бинома Ньютона**.

Для доказательства справедливости формулы (12.6) применим метод математической индукции.

1. Пусть $n = 1$, тогда

$$(a+b)^1 = C_1^0 a + C_1^1 b = a + b$$

(мы воспользовались тем, что $C_1^0 = \dfrac{1}{0!} = 1$, $C_1^1 = \dfrac{1}{1} = 1$).

2. Предполагая, что формула (12.6) верна при $n = k$, докажем, что она верна для $n = k + 1$, т.е. докажем, что

$$(a+b)^{k+1} = C_{k+1}^0 a^{k+1} + C_{k+1}^1 a^k b + \ldots + C_{k+1}^{m+1} a^{k-m} b^{m+1} + \ldots +$$
$$+ C_{k+1}^k a b^k + C_{k+1}^{k+1} b^{k+1}. \tag{12.7}$$

Имеем

$$(a+b)^{k+1} = (a+b)^k (a+b) =$$
$$= (C_k^0 a^k + C_k^1 a^{k-1} b + \ldots + C_k^m a^{k-m} b^m + \ldots + C_k^k b^k)(a+b) =$$
$$= C_k^0 a^{k+1} + C_k^1 a^k b + \ldots + C_k^{m+1} a^{k-m} b^{m+1} + \ldots + C_k^k a b^k +$$
$$+ C_k^0 a^k b + \ldots + C_k^m a^{k-m} b^{m+1} + \ldots + C_k^{k-1} a b^k + C_k^k b^{k+1}.$$

Приводя подобные члены, получим

$$(a+b)^{k+1} = C_k^0 a^{k+1} + (C_k^0 + C_k^1) a^k b + \ldots + (C_k^m + C_k^{m+1}) a^{k-m} b^{m+1} +$$
$$+ \ldots + (C_k^{k-1} + C_k^k) a b^k + C_k^k b^{k+1}.$$

В силу того что $C_k^0 = 1 = C_{k+1}^0$, $C_k^0 + C_k^1 = C_{k+1}^1$, $C_k^m + C_k^{m+1} = C_{k+1}^{m+1}$, $C_k^{k-1} + C_k^k = C_{k+1}^k, C_k^k = 1 = C_{k+1}^{k+1}$ [см. формулы (12.4′), (12.5)], получаем формулу (12.7). На основании метода математической индукции заключаем, что формула (12.6) верна для любого n.

Коэффициенты C_n^0, C_n^1, …, C_n^m, …, C_n^n в формуле (12.6) называются **биномиальными коэффициентами**.

Вопросы

1. Всякое ли множество содержит бесконечное число элементов?
2. Могут ли для множеств A и B одновременно быть верными утверждения: «A есть подмножество множества B» и «B есть подмножество множества A»?
3. В каком случае объединение двух множеств совпадает с их пересечением?
4. Что такое разность между множеством A и множеством B?
5. Что такое дополнение множества A до множества B?
6. Конечным или бесконечным является счетное множество?
7. Какие числа называются рациональными? Счетным или несчетным является множество всех рациональных чисел?
8. В чем заключается взаимно однозначное соответствие между множеством всех действительных чисел и множеством всех точек числовой прямой?
9. Какой общий термин употребляют для названия числового множества, являющегося либо отрезком, либо интервалом, либо полуинтервалом?

10. Всегда ли верно равенство $\sqrt{a^2} = a$? Если нет, то чему равен этот корень $\sqrt{a^2}$?
11. В чем заключается геометрический смысл модуля действительного числа?
12. Можно ли утверждать, что модуль суммы двух действительных чисел равен сумме их модулей? Верно ли аналогичное утверждение для произведения и частного двух действительных чисел?
13. Какому двойному неравенству эквивалентно неравенство $\left|a\right| < b$?

Глава 13
ФУНКЦИЯ

13.1. ПОНЯТИЕ ФУНКЦИИ

Одним из наиболее важных понятий в математике и ее приложениях является понятие функции.

Определение. Пусть даны числовые множества X и Y. Пусть каждому элементу $x \in X$ по какому-либо закону f поставлен в соответствие некоторый (единственный) элемент $y \in Y$. Тогда говорят, что на множестве X задана **функция*** $y = f(x)$.

При этом x называется **независимой переменной** (или **аргументом**), y – **зависимой переменной**, множество $D(f) = X$ – **областью определения** функции. Множество $R(f)$ всех значений функции называется **областью изменения** функции. Очевидно, $R(f) \subseteq Y$.

Если множество X специально не оговорено, то областью определения функции считают множество всех таких значений x, при которых функция $y = f(x)$ вообще имеет смысл (это так называемая естественная область определения).

Заметим, что для обозначения функции и ее аргумента применяются различные буквы, например:

$$y = y(x), \quad y = F(x), \quad s = s(t), \quad y = \varphi(x).$$

Наиболее употребительными являются следующие **способы задания функции**:

1) ***аналитический*** – связь между аргументом и функцией задается в виде формулы (или формул). Так, функции $y = 2x + 3$, $y = \dfrac{1}{x^2}$, $y = x^3 + \dfrac{2x}{\sqrt{x^2+1}}$ заданы аналитически.

Следует заметить, что одна функция может определяться и набором формул: разным участкам области определения функции

---

* Точнее говоря, **числовая функция**.

соответствуют разные формулы (разные аналитические выражения). Например:

$$y = \begin{cases} x^2 - 1, & \text{если } x \le 1, \\ 1 - x, & \text{если } x > 1; \end{cases}$$

2) ***табличный*** – функция задается таблицей, содержащей значения аргумента x и соответствующие значения функции $f(x)$. Примерами могут служить таблицы бухгалтерской отчетности, таблица логарифмов. На табличном способе задания, хранения и обработки информации основаны по существу и базы данных, следовательно, в их основе – также табличная форма функциональной зависимости;

3) ***графический*** – функция задана графически, если начерчен ее график, т.е. соответствие между аргументом и функцией задано посредством графика. К достоинствам этого способа можно отнести его наглядность, к недостаткам – невысокую точность.

Существуют и другие, менее распространенные способы задания функций, например ***словесный***, заключающийся в том, что функция описывается правилом ее составления.

Перейдем к рассмотрению **основных свойств функций**:

1) ***четность*** и ***нечетность***. Функция $y = f(x)$, заданная на симметричном относительно начала координат интервале, называется **четной**, если для любых значений x из ее области определения выполняется равенство $f(-x) = f(x)$. Если же $f(-x) = -f(x)$, то функция называется **нечетной**. Функция, которая не является четной или нечетной, называется **функцией общего вида**.

Например: 1) $y = x^4$ – четная функция, так как $f(-x) = (-x)^4 = x^4 = f(x)$; 2) $y = \sin x$ – нечетная функция, так как $f(-x) = \sin(-x) = -\sin x = -f(x)$; 3) $y = x^2 + \sin x$ – функция общего вида, так как $f(-x) = (-x)^2 + \sin(-x) = x^2 - \sin x$, $f(-x) \ne f(x)$, $f(-x) \ne -f(x)$.

График четной функции симметричен относительно оси Ox, а график нечетной функции – относительно начала координат;

2) ***монотонность***. Функция $y = f(x)$ называется **возрастающей** на промежутке X, если для любых $x_1, x_2 \in X$ из неравенства $x_2 > x_1$ следует $f(x_2) > f(x_1)$; функция называется **убывающей**, если из $x_2 > x_1$ следует $f(x_2) < f(x_1)$.

Функция называется **монотонной** на промежутке X, если она или возрастает на всем этом промежутке, или убывает на нем.

Заметим, что мы дали определение функции монотонной в строгом смысле. Вообще к монотонным функциям относятся

неубывающие функции, т.е. такие, для которых из $x_2 > x_1$ следует $f(x_2) \geq f(x_1)$, и **невозрастающие**, т.е. такие, для которых из $x_2 > x_1$ следует $f(x_2) \leq f(x_1)$;

3) ***ограниченность.*** Функция $y = f(x)$ называется **ограниченной** в данной области, если существует такое число $M > 0$, что $|f(x)| \leq M$ для всех x из этой области.

Например, функция $\frac{1}{x^2+1}$ ограничена на всей числовой прямой, так как $\left|\frac{1}{x^2+1}\right| \leq 1$ для любого $x \in \mathbf{R}$;

4) ***периодичность.*** Функция $y = f(x)$ называется **периодической**, если существует такое число $T \neq 0$, что $f(x + T) = f(x)$ для всех x из области определения функции.

В этом случае T называется **периодом** функции. Очевидно, если T – период функции $y = f(x)$, то периодами этой функции являются также $2T$, $3T$ и т.д. Поэтому обычно периодом функции называют наименьший положительный период (если он существует). Например, функция $y = \sin x$ имеет период $T = 2\pi$, а функция $y = \operatorname{tg} 3x$ – период $T = \frac{\pi}{3}$.

13.2. ОСНОВНЫЕ ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ

Перечислим основные элементарные функции и кратко напомним их основные свойства, известные из школьного курса математики.

1. **Степенная функция** $y = x^a$, a – любое действительное число:

1) a – натуральное число. Область определения функции – вся числовая прямая. Функция является *нечетной* при a нечетном и *четной* при a четном.

Если a – нечетное, то функция возрастает на $(-\infty, +\infty)$; если a – четное, то убывает на $(-\infty, 0]$ и возрастает на $[0, +\infty)$;

2) a – целое отрицательное число. В этом случае функция определена при всех значениях x, кроме $x = 0$.

Функция является *нечетной* при a нечетном и *четной*, если a – четное. Если a – нечетное, то функция убывает на $(-\infty, 0)$ и $(0, +\infty)$;

3) $a = \frac{1}{n}$, $n \neq 0$. Если n – четное, то функция определена на $[0, +\infty)$; если n – нечетное, то на $(-\infty, +\infty)$. Функция возрастает на всей области определения.

На рис. 13.1—13.4 изображены графики степенной функции при различных значениях a.

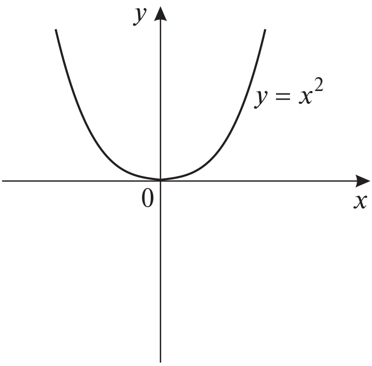


Рис. 13.1. График степенной функции $y = x^a$ при $a = 2$

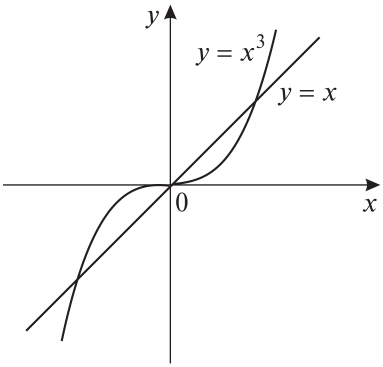


Рис. 13.2. График степенной функции $y = x^a$ при $a = 1$ и $a = 3$

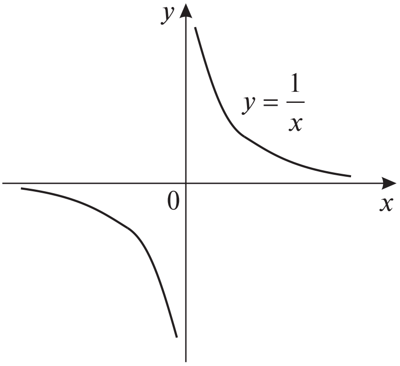


Рис. 13.3. График степенной функции $y = x^a$ при $a = -1$

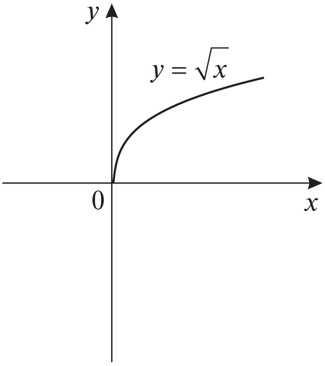


Рис. 13.4. График степенной функции $y = x^a$ при $a = \frac{1}{2}$

Степенная функция не является периодической ни при каком a.

2. **Показательная функция** $y = a^x$, $a > 0$, $a \neq 1$. Эта функция определена на всей числовой прямой; является функцией общего вида; возрастает на $(-\infty, +\infty)$ при $a > 1$ (рис. 13.5, *а*), убывает на $(-\infty, +\infty)$ при $0 < a < 1$ (рис. 13.5, *б*). Непериодическая.

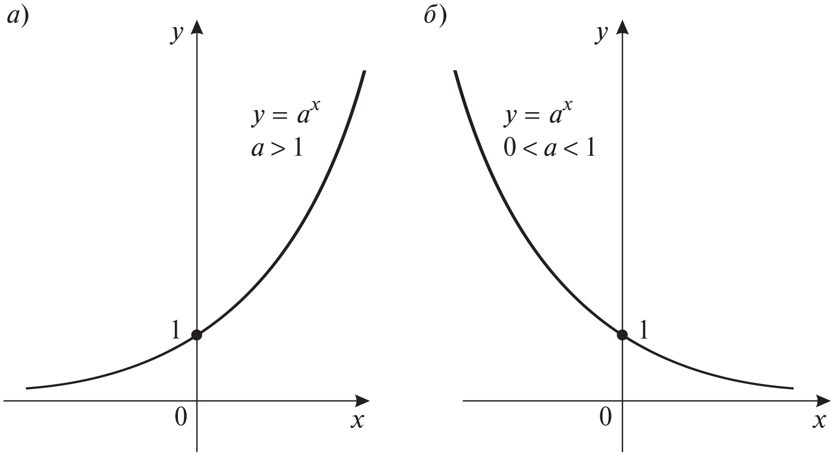


Рис. 13.5. График показательной функции $y = a^x$ при $a > 1$ (*а*) и при $0 < a < 1$ (*б*)

3. **Логарифмическая функция** $y = \log_a x$, $a > 0$, $a \neq 1$. Определена на $(0, +\infty)$, является функцией общего вида; возрастает на $(0, +\infty)$ при $a > 1$ (рис. 13.6, *а*); убывает на $(0, +\infty)$ при $0 < a < 1$ (рис. 13.6, *б*); непериодическая.

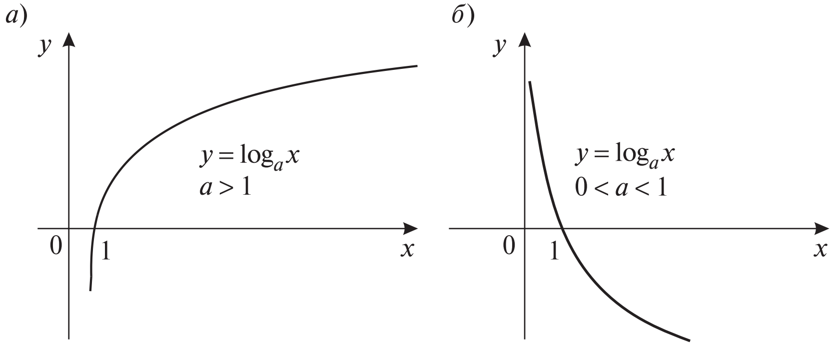


Рис. 13.6. График логарифмической функции $y = \log_a x$ при $a > 1$ (*а*) и при $0 < a < 1$ (*б*)

Напомним, что показательная функция $y = a^x$ и логарифмическая функция $y = \log_a x$ являются взаимно обратными.

4. **Тригонометрические функции**:

1) $y = \sin x$ (рис. 13.7). Нечетная периодическая функция с периодом $T = 2\pi$, определенная на $(-\infty, +\infty)$;

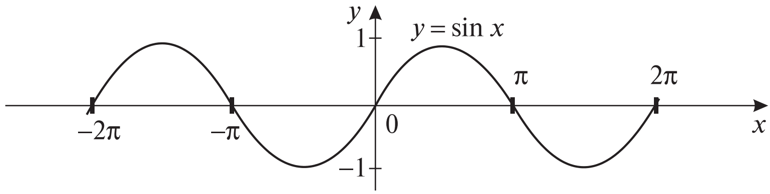


Рис. 13.7. График функции $y = \sin x$

2) $y = \cos x$ (рис. 13.8). Четная периодическая функция с периодом $T = 2\pi$, определенная на $(-\infty, +\infty)$;

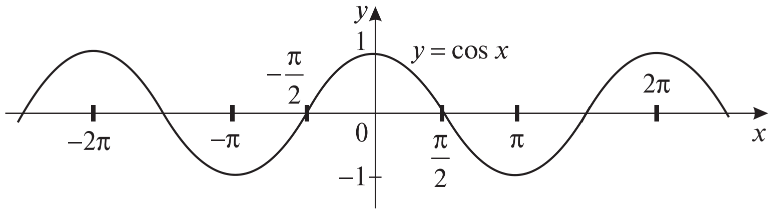


Рис. 13.8. График функции $y = \cos x$

3) $y = \operatorname{tg} x$ (рис. 13.9). Область определения имеет вид $\left(-\frac{\pi}{2} + \pi n, \frac{\pi}{2} + \pi n\right), n \in \mathbf{Z}$. Функция нечетная, возрастает на $\left(-\frac{\pi}{2} + \pi n, \frac{\pi}{2} + \pi n\right)$, $n \in \mathbf{Z}$. Является периодической функцией с периодом $T = \pi$;

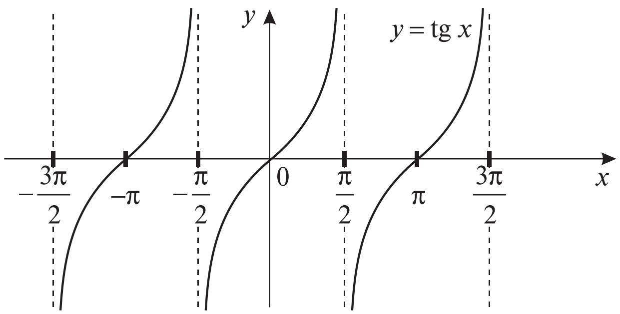


Рис. 13.9. График функции $y = \operatorname{tg} x$

4) $y = \operatorname{ctg} x$ (рис. 13.10). Область определения: $(\pi n, \pi + \pi n)$, $n \in \mathbf{Z}$. Четная периодическая функция, убывает на $(\pi n, \pi + \pi n)$, $n \in \mathbf{Z}$; период $T = \pi$.

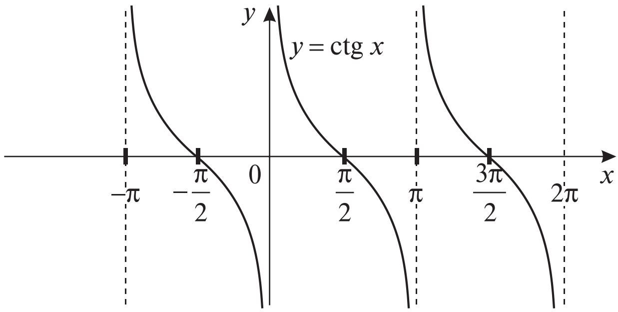


Рис. 13.10. График функции $y = \operatorname{ctg} x$

5. **Обратные тригонометрические функции**:

1) $y = \arcsin x$; 2) $y = \arccos x$;

3) $y = \operatorname{arctg} x$; 4) $y = \operatorname{arcctg} x$.

Функции $\arcsin x$ и $\arccos x$ определены на $[-1, 1]$, а функции $\operatorname{arctg} x$ и $\operatorname{arcctg} x$ – на всей числовой прямой.

13.3. ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ

Пусть $y = f(x)$ определена на промежутке X, ее область изменения есть Y, и пусть разным значениям x соответствуют разные значения y. Тогда для каждого значения $y \in Y$ существует единственное число $x \in X$, при котором $f(x) = y$. Полученная в этом случае функция $x = \varphi(y)$, определенная на X с областью изменения Y, называется **обратной**.

Так как обычно независимую переменную обозначают через x, а функцию – через y, то функцию, обратную к функции $y = f(x)$, обозначают также в виде $y = f^{-1}(x)$.

Взаимно обратными функциями являются, в частности, $y = a^x$ и $y = \log_a x$ (рис. 13.11), $y = \sin x$ и $y = \arcsin x$ и т.д. Графики взаимно обратных функций симметричны относительно прямой $y = x$.

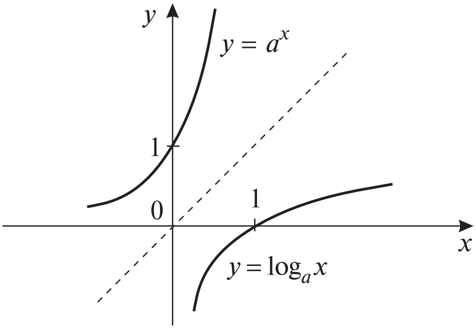


Рис. 13.11. Графики взаимно обратных функций
$y = a^x$ и $y = \log_a x$

Хорошо известно, что над функциями можно производить арифметические действия: сложение, вычитание, умножение, деление.

Рассмотрим еще одно действие над функциями, называемое *взятием функции от функции*, или *построением сложной функции*. Пусть функция $y = f(u)$ определена на множестве U, ее область изменения есть Y, а ее аргумент u есть функция от x: $u = \varphi(x)$, определенная на множестве X с областью изменения U. Тогда функция $y = f(\varphi(x))$, определенная на X, называется **сложной функцией**, или **функцией от функции** (**суперпозицией функций**).

Например, две функции $y = \lg u$ и $u = 1 - x^2$ определяют сложную функцию $y = \lg(1 - x^2)$ с областью определения $(-1, 1)$.

Заметим, что операция взятия функции от функции может производиться любое число раз. Например, функция $y = \sqrt{\lg \sin x^2}$ получается в результате следующих операций:

$$y = \sqrt{u}, \quad u = \lg v, \quad v = \sin w, \quad w = x^2.$$

Определение. **Элементарной** называется функция, которая получена из основных элементарных функций и констант с помощью конечного числа операций сложения, вычитания, умножения, деления и взятия функции от функции.

Элементарные функции делятся на *алгебраические* и *трансцендентные*.

К алгебраическим функциям относятся:

а) многочлены

$$y = a_0 x^n + a_1 x^{n-1} + \ldots + a_n;$$

б) дробно-рациональные функции

$$y = \frac{a_0 x^n + a_1 x^{n-1} + \ldots + a_n}{b_0 x^m + b_1 x^{m-1} + \ldots + b_m},$$

т.е. функции, определяемые как отношение двух многочленов;

в) иррациональные функции, т.е. функции, полученные путем конечного числа суперпозиций и арифметических действий над степенными функциями с целыми и дробными показателями и не являющиеся рациональными.

Примеры иррациональных функций: $y = x^2 + \sqrt[3]{x}$, $y = \dfrac{\sqrt{x^2+1}}{x^3+\sqrt{x}}$ и т.п.

Вообще, **алгебраической** называется функция $y = f(x)$, которая удовлетворяет уравнению вида

$$P_0 y^n + P_1 y^{n-1} + \ldots + P_n = 0,$$

где $P_0, P_1, \ldots, P_n$ – многочлены, зависящие от x.

Функция, которая не является алгебраической, называется **трансцендентной**.

К трансцендентным относятся показательная, логарифмическая, тригонометрические функции.

13.4. ПРИМЕНЕНИЕ ФУНКЦИЙ В ЭКОНОМИКЕ

Функции широко применяются в экономической теории. Приведем наиболее часто применяемые функции одного аргумента.

Функция полезности (рис. 13.12) – субъективная числовая оценка данным индивидом полезности u количества x некоторого товара. В широком смысле функция полезности – зависимость полезности (эффекта) некоторого действия от его уровня (интенсивности).

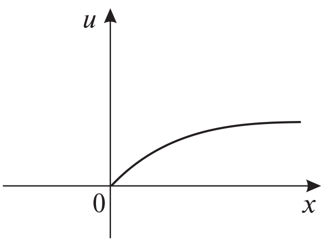


Рис. 13.12. График функции полезности

Функция выпуска, однофакторная производственная функция (рис. 13.13) – зависимость объема y выпускаемой продукции от объема x перерабатываемого ресурса. Функция выпуска является частным видом **производственной функции**, которая выражает зависимость результата производственной деятельности от обусловивших его факторов.

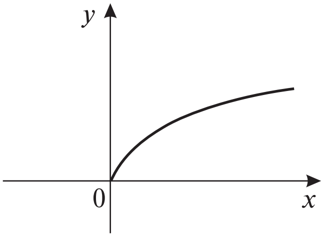


Рис. 13.13. График функции выпуска

Функция издержек – зависимость издержек производства от объема продукции. Функция издержек также есть частный вид производственной функции.

Функция спроса и предложения – зависимость объема спроса D и предложения S от цены на товар p.

Рассмотрим какой-нибудь товар. Пусть $D(p)$ – количество (число единиц) товара, которое покупатель на рынке желает купить при данной цене p за единицу. $D = D(p)$ называется *функцией спроса* на товар. Эта функция – убывающая. Обычно она имеет вид

$$D = kp^a + c, \qquad (13.1)$$

где $a < 0$ (рис. 13.14).

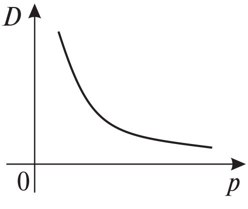


Рис. 13.14. График функции спроса

С другой стороны, пусть $S(p)$ – число единиц товара, предлагаемого продавцами на рынке по цене p. Очевидно, предложение растет с ростом цены. Поэтому *функция предложения* $S = S(p)$ – возрастающая функция. Она обычно имеет вид

$$S = p^b + d, \tag{13.2}$$

где $b \geq 1$ (рис. 13.15).

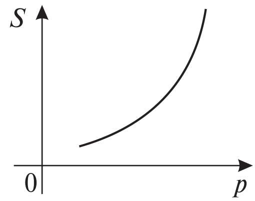


Рис. 13.15. График функции предложения

Для экономики представляет интерес условие, когда спрос р а в е н предложению:

$$D(p) = S(p). \tag{13.3}$$

Цена $p = p_0$, при которой выполняется равенство (13.3), называется **равновесной**. Точка пересечения кривых D и S (графиков функций $D = D(p)$ и $S = S(p)$) называется **точкой равновесия**.

При увеличении благосостояния населения константа c в формуле (13.1) увеличивается, кривая D поднимается вверх, точка равновесия смещается вправо (рис. 13.16).

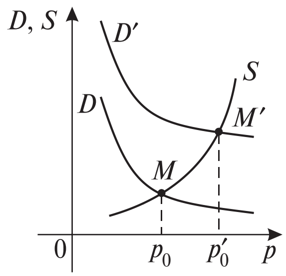


Рис. 13.16. Положение точки равновесия в зависимости от благосостояния населения

Вопросы

1. Что такое естественная область определения функции?
2. Какие существуют способы задания функций?
3. Каким свойством обладает график нечетной функции?
4. Какой общий термин употребляют, говоря о возрастающей и убывающей функциях?
5. Сколько разных периодов имеет периодическая функция?
6. Пусть функция $y = f(x)$ имеет наименьший период T. Является ли периодической функция $y = f(3x)$ и если является, то каков ее наименьший период?
7. Как получить график обратной функции из графика самой функции?
8. Какая функция является обратной для функции $y = x^3$?
9. Относится ли функция $y = 2x + 3$ к основным элементарным функциям?
10. Как можно получать элементарные функции из основных элементарных?
11. Будет ли элементарной функцией сумма элементарных функций? А квадратный корень из элементарной функции?
12. Какие функции одной переменной наиболее часто применяются в экономике?

Глава 14
ПРЕДЕЛЫ

14.1. ПОСЛЕДОВАТЕЛЬНОСТЬ. ПРЕДЕЛ ПОСЛЕДОВАТЕЛЬНОСТИ

Если по некоторому закону каждому натуральному числу n поставлено в соответствие одно определенное действительное число x_n, то говорят, что задана **числовая последовательность**:

$$x_1, x_2, \ldots, x_n, \ldots . \tag{14.1}$$

Иначе говоря, числовая последовательность — это *функция натурального аргумента*: $x_n = f(n)$, т.е. функция, определенная на множестве натуральных чисел.

Последовательность (14.1) записывается коротко в виде $\{x_n\}$. Числа $x_1, x_2, \ldots, x_n, \ldots$ называют **членами** последовательности, а n-й член x_n — **общим членом** последовательности.

Последовательность считается заданной, если задан ее общий член или указан способ получения любого ее элемента. Например, формула $x_n = \dfrac{2n+1}{5n+2}$ определяет последовательность

$$\frac{3}{7}, \frac{5}{12}, \frac{7}{17}, \frac{9}{22}, \ldots, \frac{2n+1}{5n+2}, \ldots . \tag{14.2}$$

Последовательность может быть *монотонной* или *немонотонной*, *ограниченной* или *неограниченной*. (Нет необходимости определять эти понятия, так как в § 13.1 уже были даны определения монотонной и ограниченной функции.)

В частности, последовательность (14.2) является монотонно убывающей. Действительно, рассмотрим разность $x_n - x_{n+1}$:

$$x_n - x_{n+1} = \frac{2n+1}{5n+2} - \frac{2(n+1)+1}{5(n+1)+2} = \frac{2n+1}{5n+2} - \frac{2n+3}{5n+7} =$$

$$= \frac{(2n+1)(5n+7)-(2n+3)(5n+2)}{(5n+2)(5n+7)} =$$

$$= \frac{10n^2+19n+7-10n^2-19n-6}{(5n+2)(5n+7)} = \frac{1}{(5n+2)(5n+7)}.$$

Итак, $x_n - x_{n+1} > 0$, т.е. $x_{n+1} < x_n$ для любого n.

Рассмотрим эту же последовательность (14.2). Если, например, $n = 100$, то $x_n = \dfrac{201}{502}$; если $n = 100\,000$, то $x_n = \dfrac{200\,001}{500\,002}$. Мы видим, что с ростом n члены последовательности x_n все меньше отличаются от $\dfrac{2}{5}$, и это различие может стать как угодно малым.

В частности, при $n = 100\,000$

$$x_n - \frac{2}{5} = \frac{200\,001}{500\,002} - \frac{2}{5} = 0{,}0000003999\ldots < 0{,}000001.$$

Определение. Число a называется **пределом последовательности** $\{x_n\}$, если для любого (сколь угодно малого) числа $\varepsilon > 0$ существует такой номер N, что при всех $n > N$ выполняется неравенство

$$\left| x_n - a \right| < \varepsilon. \qquad (14.3)$$

Если последовательность $\{x_n\}$ имеет предел a, то она называется **сходящейся** (к числу a). В этом случае пишем:

$$\lim_{n\to\infty} x_n = a. \qquad (14.4)$$

Иногда вместо (14.4) пишут: $x_n \to a$ при $n \to \infty$.

Если вернуться к последовательности (14.2), то мы видим, что для $\varepsilon = 0{,}000001$ неравенство (14.3) выполняется при $n > 100\,000$.

Выясним **геометрический смысл предела числовой последовательности**. Неравенство (14.3) равносильно двойному неравенству

$$-\varepsilon < x_n - a < \varepsilon, \quad \text{или} \quad a - \varepsilon < x_n < a + \varepsilon.$$

Это означает, что при $n > N$ все члены последовательности $\{x_n\}$ находятся в ε-окрестности точки a (рис. 14.1).

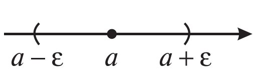


Рис. 14.1. ε-окрестность точки a

Следовательно, вне этой окрестности может находиться лишь конечное число членов этой последовательности.

14.2. ПРЕДЕЛ ФУНКЦИИ

Предел функции в бесконечности

Рассматривая предел последовательности $x_n = f(n)$, мы имели дело с функцией, аргумент которой, n, возрастая, принимал лишь натуральные значения. Теперь рассмотрим функцию $y = f(x)$. Ее аргумент x в процессе изменения может принимать любые (не только натуральные и не только целые) значения.

Определение. Число b называется **пределом функции** $y = f(x)$ **при** x, **стремящемся к бесконечности**, если для любого (сколь угодно малого) $\varepsilon > 0$ существует такое число $M > 0$, что для всех x, удовлетворяющих условию $|x| > M$, выполняется неравенство

$$|f(x) - b| < \varepsilon.$$

В этом случае пишем:

$$\lim_{x\to\infty} f(x) = b. \tag{14.5}$$

Иногда вместо (14.5) пишут: $f(x) \to b$ при $x \to \infty$.

Смысл определения предела функции в бесконечности в основном тот же, что и для предела последовательности:

- $\lim\limits_{n\to\infty} x_n = a$ означает, что члены последовательности сколь угодно мало отличаются от a, если n достаточно велико;
- $\lim\limits_{x\to\infty} f(x) = b$ означает, что значения функции сколь угодно мало отличаются от b, если x достаточно велико по абсолютной величине.

Замечание. Если в сформулированном выше определении условие $|x| > M$ заменить на условие $x > M$, то получим определение предела функции при $x \to +\infty$. Если же его заменить на условие $x < -M$, то получим определение предела при $x \to -\infty$.

Предел функции в точке

Пусть функция $y = f(x)$ определена в некоторой окрестности точки a, кроме, может быть, самой точки a.

Определение. Число b называется **пределом функции** $y = f(x)$ **при** x, **стремящемся к** a, если для любого (сколь угодно малого) числа $\varepsilon > 0$ существует такое число $\delta > 0$, что для всех $x \neq x_0$, удовлетворяющих условию $|x - a| < \delta$, выполняется неравенство

$$|f(x) - b| < \varepsilon.$$

В этом случае пишем:

$$\lim_{x \to a} f(x) = b. \tag{14.6}$$

Геометрический смысл этого определения таков: для любой ε-окрестности точки b (на оси Oy) существует такая δ-окрестность точки a (на оси Ox), что как только x попадает в эту δ-окрестность точки a, соответствующее значение функции $f(x)$ будет принадлежать ε-окрестности точки b, т.е. $x \in (a - \delta, a + \delta) \Rightarrow$ $\Rightarrow f(x) \in (b - \varepsilon, b + \varepsilon)$.

14.3. БЕСКОНЕЧНО МАЛЫЕ ВЕЛИЧИНЫ. БЕСКОНЕЧНО БОЛЬШИЕ ВЕЛИЧИНЫ

Определение. Функция $\alpha = \alpha(x)$ называется **бесконечно малой величиной** (или просто **бесконечно малой**) при $x \to x_0$ (при $x \to \infty$), если ее предел равен нулю: $\lim\limits_{x \to x_0} \alpha(x) = 0$ ($\lim\limits_{x \to \infty} \alpha(x) = 0$).

Аналогично определяется бесконечно малая последовательность.

Рассмотрим теперь **связь переменной величины с ее пределом**.

Теорема **14.1.** Число b является пределом функции $f(x)$ при $x \to x_0$ (при $x \to \infty$) тогда и только тогда, когда

$$f(x) = b + \alpha(x), \tag{14.7}$$

где $\alpha(x)$ – бесконечно малая при $x \to x_0$ (при $x \to \infty$).

Доказательство. 1. *Необходимость.* Пусть $\lim\limits_{x \to x_0} f(x) = b$. Обозначим $\alpha(x) = f(x) - b$. Пусть задано $\varepsilon > 0$. Тогда существует такое $\delta > 0$, что для всех $x \neq x_0$, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство $|f(x) - b| < \varepsilon$, т.е. $|\alpha(x)| < \varepsilon$, а это означает, что $\alpha(x)$ – бесконечно малая.

2. *Достаточность.* Пусть $f(x) = b + \alpha(x)$, где $\alpha(x)$ – бесконечно малая. Тогда разность $f(x) - b$ есть бесконечно малая, т.е. для всякого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех $x \neq x_0$, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство $|f(x) - b| < \varepsilon$, а это означает, что $\lim\limits_{x \to x_0} f(x) = b$.

Мы доказали теорему для случая $x \to x_0$. Доказательство для $x \to \infty$ аналогично.

Определение. Функция $y = f(x)$ называется **бесконечно большой величиной** (или просто **бесконечно большой**) при $x \to x_0$, если для всякого (сколь угодно большого) $M > 0$ существует такое $\delta > 0$, что для всех x, не равных x_0 и удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство

$$|f(x)| > M.$$

В этом случае также говорят, что $f(x)$ имеет бесконечный предел при $x \to x_0$, или $f(x)$ стремится к бесконечности при $x \to x_0$, и пишут: $\lim\limits_{x \to x_0} f(x) = \infty$, или $f(x) \to \infty$ при $x \to x_0$.

Аналогично определяется бесконечно большая при $x \to \infty$.

Между понятиями бесконечно малой и бесконечно большой существует очевидная связь: если $\alpha(x)$ – бесконечно малая при $x \to x_0$ $(x \to \infty)$, то $f(x) = \dfrac{1}{\alpha(x)}$ есть бесконечно большая при $x \to x_0$ $(x \to \infty)$; если $f(x)$ – бесконечно большая, то $\alpha(x) = \dfrac{1}{f(x)}$ есть бесконечно малая.

Свойства бесконечно малых:

1. *Алгебраическая сумма конечного числа бесконечно малых есть бесконечно малая.*

2. *Произведение бесконечно малой величины на величину ограниченную есть бесконечно малая.*

Следствием этого утверждения являются следующие утверждения:

1) *произведение бесконечно малой на константу есть бесконечно малая*;

2) *произведение двух бесконечно малых есть бесконечно малая.*

Докажем в качестве примера свойство 2. Пусть $\alpha = \alpha(x)$ есть бесконечно малая при $x \to x_0$, а y – величина ограниченная, $|y| < M$. Пусть задано $\varepsilon > 0$. Тогда для $\varepsilon' = \dfrac{\varepsilon}{M}$ существует такое $\delta > 0$, что для всех $x \neq x_0$, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство

$$|\alpha| < \varepsilon' = \frac{\varepsilon}{M}.$$

Но тогда

$$|\alpha y| < \frac{\varepsilon}{M} M = \varepsilon, \text{ т.е. } |\alpha y| < \varepsilon.$$

А это означает, что $|\alpha y|$ есть бесконечно малая.

Бесконечно малые можно сравнивать. В частности, если $\lim\limits_{x\to x_0}\dfrac{\alpha(x)}{\beta(x)}=0$, то говорят, что $\alpha(x)$ – **бесконечно малая более высокого порядка**, чем $\beta(x)$, и пишут:

$$\alpha(x)=o(\beta(x)).$$

14.4. ОСНОВНЫЕ ТЕОРЕМЫ О ПРЕДЕЛАХ

Единственность предела

***Теорема* 14.2.** Если функция имеет предел, то этот предел – единственный.

Доказательство. Предположим противное: $\lim\limits_{x\to a}f(x)=b_1$, $\lim\limits_{x\to a}f(x)=b_2$ и $b_1\neq b_2$. Тогда на основании теоремы 14.1

$$f(x)=b_1+\alpha_1(x) \quad \text{и} \quad f(x)=b_2+\alpha_2(x),$$

где $\alpha_1(x)$ и $\alpha_2(x)$ – бесконечно малые. Вычитая почленно эти равенства, получаем

$$0=b_1-b_2+[\alpha_1(x)-\alpha_2(x)].$$

Отсюда

$$\alpha_1(x)-\alpha_2(x)=b_2-b_1=c=\text{const}\neq 0.$$

Это равенство невозможно, так как разность $\alpha_1(x)-\alpha_2(x)$ есть бесконечно малая. Следовательно, предположение о существовании двух различных пределов неверно.

Предел суммы, произведения, частного

Будем рассматривать пределы функций $u=u(x)$ и $v=v(x)$ при $x\to a$ или при $x\to\infty$. Краткая запись $\lim u$ будет означать или $\lim\limits_{x\to a}u(x)$, или $\lim\limits_{x\to\infty}u(x)$. Аналогично для $\lim v$.

1. *Предел алгебраической суммы равен алгебраической сумме пределов*:

$$\lim(u+v)=\lim u+\lim v.$$

2. *Предел произведения равен произведению пределов*:

$$\lim(uv)=\lim u\cdot\lim v.$$

3. *Предел частного равен частному пределов* (при условии, что предел делителя отличен от нуля):

$$\lim \frac{u}{v} = \frac{\lim u}{\lim v}.$$

Сформулируем эти утверждения более четко для случая $x \to a$ и докажем. (Для $x \to \infty$ доказательства аналогичны.)

1. Если функции $u = u(x)$ и $v = v(x)$ имеют при $x \to a$ пределы $\lim\limits_{x\to a} u(x) = b_1$, $\lim\limits_{x\to a} v(x) = b_2$, то их сумма $u(x) + v(x)$ также имеет предел $\lim\limits_{x\to a} [u(x) + v(x)] = b_1 + b_2$.

Доказательство. Так как $\lim\limits_{x\to a} u(x) = b_1$, $\lim\limits_{x\to a} v(x) = b_2$, то по теореме 14.1 функции $u(x)$ и $v(x)$ можно представить в виде $u(x) = b_1 + \alpha_1(x)$, $v(x) = b_2 + \alpha_2(x)$, где $\alpha_1(x)$ и $\alpha_2(x)$ – бесконечно малые, $\lim\limits_{x\to a} \alpha_1(x) = 0$, $\lim\limits_{x\to a} \alpha_2(x) = 0$. Следовательно,

$$u(x) + v(x) = [b_1 + \alpha_1(x)] + [b_2 + \alpha_2(x)] = (b_1 + b_2) + [\alpha_1(x) + \alpha_2(x)].$$

Здесь $b_1 + b_2$ есть величина постоянная, а величина $\alpha_1(x) + \alpha_2(x)$ – бесконечно малая. Снова применяя теорему 14.1, получаем

$$\lim_{x\to a} [u(x) + v(x)] = b_1 + b_2, \text{ т.е. } \lim_{x\to a} [u(x) + v(x)] = \lim_{x\to a} u(x) + \lim_{x\to a} v(x).$$

2. Если функции $u = u(x)$ и $v = v(x)$ имеют при $x \to a$ пределы $\lim\limits_{x\to a} u(x) = b_1$, $\lim\limits_{x\to a} v(x) = b_2$, то произведение $u(x)v(x)$ также имеет предел и $\lim\limits_{x\to a} [u(x)v(x)] = b_1 b_2$.

Доказательство. Так как $\lim\limits_{x\to a} u(x) = b_1$, $\lim\limits_{x\to a} v(x) = b_2$, то в силу теоремы 14.1 $u(x) = b_1 + \alpha_1(x)$, $v(x) = b_2 + \alpha_2(x)$, где $\alpha_1(x)$ и $\alpha_2(x)$ – бесконечно малые при $x \to a$. Имеем:

$$u(x)v(x) = [b_1 + \alpha_1(x)][b_2 + \alpha_2(x)] =$$
$$= b_1 b_2 + b_1\alpha_2(x) + b_2\alpha_1(x) + \alpha_1(x)\alpha_2(x).$$

Обозначим $b_1\alpha_2(x) + b_2\alpha_1(x) + \alpha_1(x)\alpha_2(x) = \alpha(x)$. Как было установлено ранее, $b_1\alpha_2(x)$, $b_2\alpha_1(x)$ и $\alpha_1(x)\alpha_2(x)$ – бесконечно малые, следовательно, их сумма $\alpha(x)$ также есть бесконечно малая. Итак,

$$u(x)v(x) = b_1 b_2 + \alpha(x),$$

где $\alpha(x)$ – бесконечно малая. Это и означает, что $u(x)v(x)$ имеет предел, равный $b_1 b_2$, что и требовалось доказать.

Следствие. Постоянный множитель можно выносить за знак предела:

$$\lim_{x\to a} cu(x) = c\lim_{x\to a} u(x).$$

Действительно, если $\lim\limits_{x\to a} u(x) = b$, c = const, то $\lim\limits_{x\to a} c = c$. Тогда $\lim\limits_{x\to a}[cu(x)] = \lim\limits_{x\to a} c \lim\limits_{x\to a} u(x) = c\lim\limits_{x\to a} u(x)$.

3. Если функции $u = u(x)$ и $v = v(x)$ имеют при $x \to a$ пределы $\lim\limits_{x\to a} u(x) = b_1$, $\lim\limits_{x\to a} v(x) = b_2$, причем $b_2 \neq 0$, то функция $\dfrac{u(x)}{v(x)}$ также имеет предел при $x \to a$ и $\lim\limits_{x\to a}\dfrac{u(x)}{v(x)} = \dfrac{b_1}{b_2}$.

Доказательство. Пусть $\lim\limits_{x\to a} u(x) = b_1$, $\lim\limits_{x\to a} v(x) = b_2 \neq 0$. По теореме 14.1 $u(x) = b_1 + \alpha_1(x)$, $v(x) = b_2 + \alpha_2(x)$, где $\alpha_1(x)$, $\alpha_2(x)$ – бесконечно малые. Выполним несложные тождественные преобразования:

$$\frac{u(x)}{v(x)} = \frac{b_1 + \alpha_1(x)}{b_2 + \alpha_2(x)} = \frac{b_1}{b_2} + \left(\frac{b_1 + \alpha_1(x)}{b_2 + \alpha_2(x)} - \frac{b_1}{b_2}\right) = \frac{b_1}{b_2} + \frac{b_2\alpha_1(x) - b_1\alpha_2(x)}{b_2[b_2 + \alpha_2(x)]}.$$

Итак,

$$\frac{u(x)}{v(x)} = \frac{b_1}{b_2} + \frac{b_2\alpha_1(x) - b_1\alpha_2(x)}{b_2[b_2 + \alpha_2(x)]}.$$

Здесь $\dfrac{b_1}{b_2}$ – постоянное число, а дробь $\dfrac{b_2\alpha_1(x) - b_1\alpha_2(x)}{b_2[b_2 + \alpha_2(x)]}$ есть бесконечно малая – это следует из свойств бесконечно малых: $b_2\alpha_1(x) - b_1\alpha_2(x)$ есть бесконечно малая, а $\dfrac{1}{b_2[b_2 + \alpha_2(x)]}$ ограничена.

Предельный переход в неравенствах

Предположим, что неравенства, о которых речь пойдет ниже, выполняются в некоторой окрестности точки a (исключая, может быть, эту точку) или для достаточно больших x.

4. Если функция $u = u(x)$ неотрицательна: $u \geq 0$, то $\lim u \geq 0$.

5. Если для функций $u = u(x)$ и $v = v(x)$ выполняется неравенство $u \geq v$, то $\lim u \geq \lim v$.

6. Если для функций $u = u(x)$, $v = v(x)$, $w = w(x)$ выполняется неравенство $u \le v \le w$ и $\lim u = \lim w = b$, то $\lim v = b$.

Для примера сформулируем более подробно и докажем последнее утверждение.

***Теорема* 14.3.** Если в некоторой окрестности точки a выполняются условия $u(x) \le v(x) \le w(x)$ и при $x \to a$ функции $u = u(x)$ и $w = w(x)$ имеют один и тот же предел: $\lim\limits_{x\to a} u(x) = \lim\limits_{x\to a} w(x) = b$, то функция $v(x)$ имеет тот же предел: $\lim\limits_{x\to a} v(x) = b$.

Доказательство. Пусть задано произвольное $\varepsilon > 0$. Тогда, так как $\lim\limits_{x\to a} u(x) = b$, существует такое $\delta_1 > 0$, что для всех $x \ne a$, удовлетворяющих условию $|x - a| < \delta_1$, выполняется неравенство

$$|u(x) - b| < \varepsilon. \qquad (*)$$

Так как $\lim\limits_{x\to a} w(x) = b$, то существует такое $\delta_2 > 0$, что для всех $x \ne a$, удовлетворяющих условию $|x - a| < \delta_2$, выполняется неравенство

$$|w(x) - b| < \varepsilon. \qquad (**)$$

Если δ – наименьшее из δ_1 и δ_2, то для всех $x \ne a$, удовлетворяющих условию $|x - a| < \delta$, оба неравенства (*) и (**) выполняются одновременно, т.е. одновременно

$$b - \varepsilon < u(x) < b + \varepsilon, \quad b - \varepsilon < w(x) < b + \varepsilon.$$

Из последних неравенств и условий теоремы получаем

$$b - \varepsilon < u(x) \le v(x) \le w(x) < b + \varepsilon,$$

следовательно, $b - \varepsilon < v(x) < b + \varepsilon$, т.е.

$$|v(x) - b| < \varepsilon.$$

А это означает, что $\lim\limits_{x\to a} v(x) = b$. (Доказательство для $x \to \infty$ аналогично.)

Односторонние пределы

Если $f(x)$ стремится к пределу b при x, стремящемся к a так, что $x < a$, то b называется **пределом функции** $f(x)$ **при** x, **стремящемся к** a **слева**, или **левосторонним пределом**. В этом случае пишут $\lim\limits_{x\to a-0} f(x) = b$ (или $\lim\limits_{x\to a-} f(x) = b$). Аналогично предел $f(x)$ при $x \to a$,

$x > a$, называется **правосторонним пределом** и записывается в виде $\lim\limits_{x\to a+0} f(x) = b$ (или $\lim\limits_{x\to a+} f(x) = b$).

Нетрудно доказать, что функция $f(x)$ имеет предел при $x \to a$ тогда и только тогда, когда существуют одновременно левосторонний и правосторонний пределы и они равны между собой:

$$\lim_{x\to a-} f(x) = \lim_{x\to a+} f(x) = b.$$

В этом случае и предел в обычном смысле также равен b:

$$\lim_{x\to a} f(x) = b.$$

Достаточный признак существования предела

Теорема **14.4.** Монотонная ограниченная последовательность имеет конечный предел.

В частности, если последовательность $\{x_n\}$ возрастает и ограничена сверху (т.е. существует такое M, что $x_n < M$ для всех n), то она имеет предел. Аналогично убывающая и ограниченная снизу последовательность также имеет предел. При этом возрастание и убывание можно понимать в широком смысле (т.е. $x_{n+1} \ge x_n$ и $x_{n+1} \le x_n$ соответственно для всех n).

Интуитивно справедливость этой теоремы представляется почти очевидной, однако строгое ее доказательство мы не приводим, так как оно основано на сведениях из теории действительного числа, которые в настоящей книге не рассматриваются.

14.5. ДВА ЗАМЕЧАТЕЛЬНЫХ ПРЕДЕЛА

1. Докажем, что существует предел функции $\dfrac{\sin x}{x}$ при $x \to 0$ и этот предел равен единице:

$$\lim_{x\to 0} \frac{\sin x}{x} = 1. \tag{14.8}$$

Этот предел обычно называют **первым замечательным пределом**.

Доказательство. Рассмотрим окружность радиуса R: $OA = OM = R$ (на рис. 14.2 изображен ее сектор). Пусть x – радианная мера острого угла AOM: $0 < x < \dfrac{\pi}{2}$.

Тогда $MB = OM \cdot \sin x = R\sin x$, $NA = OA \cdot \operatorname{tg} x = R\operatorname{tg} x$.

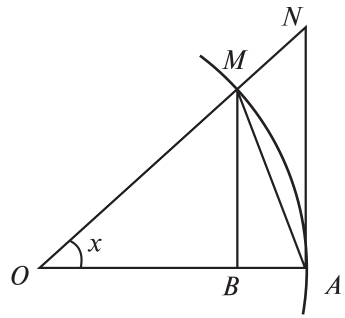


Рис. 14.2

Пусть S_1 – площадь треугольника OMA. Тогда $S_1 = \frac{1}{2} OA \cdot MB = \frac{1}{2} R^2 \sin x$. Обозначим через S_2 площадь сектора OMA. Тогда $S_2 = \frac{1}{2} R^2 x$. Обозначим через S_3 площадь треугольника AON. Тогда $S_3 = \frac{1}{2} OA \cdot NA = \frac{1}{2} R^2 \operatorname{tg} x$. Очевидно,

$$S_1 < S_2 < S_3,$$

следовательно,

$$\frac{1}{2} R^2 \sin x < \frac{1}{2} R^2 x < \frac{1}{2} R^2 \operatorname{tg} x.$$

Отсюда

$$\sin x < x < \operatorname{tg} x.$$

Разделив последнее неравенство на $\sin x$, получим

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}, \quad \text{или} \quad \cos x < \frac{\sin x}{x} < 1.$$

Итак, функция $\frac{\sin x}{x}$ заключена между функциями $u(x) = \cos x$ и $w(x) = 1$, имеющими при $x \to 0$ один и тот же предел 1*. Отсюда на основании теоремы 14.3 получаем равенство

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

Для $x < 0$ доказательство аналогично.

---

* Стремление $\cos x$ к 1 при $x \to 0$ следует из того, что $1 - \cos x = 2\sin^2 \frac{x}{2}$.

Пример **14.1.** Найти $\lim\limits_{x\to 0}\dfrac{\sin ax}{ax}$.

Решение. Сделаем замену $\alpha = ax$. Очевидно, $x\to 0$ равносильно тому, что $\alpha\to 0$. Получаем

$$\lim_{x\to 0}\frac{\sin ax}{ax} = \lim_{\alpha\to 0}\frac{\sin\alpha}{\alpha} = 1.$$

Пример **14.2.** Найти $\lim\limits_{x\to 0}\dfrac{\sin ax}{bx}$.

Решение: $\lim\limits_{x\to 0}\dfrac{\sin ax}{bx} = \lim\limits_{x\to 0}\dfrac{\sin ax}{ax}\dfrac{a}{b} = \dfrac{a}{b}\lim\limits_{x\to 0}\dfrac{\sin ax}{ax} = \dfrac{a}{b}\cdot 1 = \dfrac{a}{b}$.

Пример **14.3.** Найти $\lim\limits_{x\to 0}\dfrac{\operatorname{tg} x}{x}$.

Решение: $\lim\limits_{x\to 0}\dfrac{\operatorname{tg} x}{x} = \lim\limits_{x\to 0}\dfrac{\sin x}{x}\dfrac{1}{\cos x} = \lim\limits_{x\to 0}\dfrac{\sin x}{x}\lim\limits_{x\to 0}\dfrac{1}{\cos x} = 1\cdot 1 = 1.$

Итак, запомним:

$$\lim_{x\to 0}\frac{\operatorname{tg} x}{x} = 1.$$

Пример **14.4.** Найти $\lim\limits_{x\to 0}\dfrac{1-\cos x}{x^2}$.

Решение: $\lim\limits_{x\to 0}\dfrac{1-\cos x}{x^2} = \lim\limits_{x\to 0}\dfrac{(1-\cos x)(1+\cos x)}{x^2(1+\cos x)} =$

$$= \lim_{x\to 0}\frac{1-\cos^2 x}{x^2(1+\cos x)} = \lim_{x\to 0}\frac{\sin^2 x}{x^2(1+\cos x)} = \lim_{x\to 0}\left(\frac{\sin x}{x}\right)^2\frac{1}{1+\cos x} = 1\cdot\frac{1}{2} = \frac{1}{2}.$$

(Можно вычислить этот предел иначе, воспользовавшись формулой $1-\cos x = 2\sin^2\dfrac{x}{2}$.)

2. Рассмотрим последовательность $\{a_n\}$, общий член которой имеет вид

$$a_n = \left(1+\frac{1}{n}\right)^n.$$

Докажем, что эта последовательность монотонно возрастает и ограничена, следовательно, в силу теоремы 14.4 имеет предел.

Применим формулу бинома Ньютона:

$$\left(1+\frac{1}{n}\right)^n = 1+n\frac{1}{n}+\frac{n(n-1)}{1\cdot 2}\left(\frac{1}{n}\right)^2+\frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}\left(\frac{1}{n}\right)^3+\ldots+$$

$$+\frac{n(n-1)(n-2)\cdots[n-(n-1)]}{1\cdot 2\cdot 3\cdots n}\left(\frac{1}{n}\right)^n.$$

Преобразуем полученное выражение:

$$\left(1+\frac{1}{n}\right)^n = 1+1+\frac{1}{1\cdot 2}\frac{n(n-1)}{n^2}+\frac{1}{1\cdot 2\cdot 3}\frac{n(n-1)(n-2)}{n^3}+\ldots+$$

$$+\frac{1}{1\cdot 2\cdot 3\cdots n}\frac{n(n-1)(n-2)\cdots[n-(n-1)]}{n^n}=$$

$$=1+1+\frac{1}{1\cdot 2}\left(1-\frac{1}{n}\right)+\frac{1}{1\cdot 2\cdot 3}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\ldots+$$

$$+\frac{1}{1\cdot 2\cdot 3\cdots n}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-1}{n}\right). \qquad (*)$$

Из последнего равенства следует, что $\left(1+\frac{1}{n}\right)^n \geq 2$, а также следует, что рассматриваемая последовательность возрастает с возрастанием n.

Действительно, при переходе от n к $n+1$ каждое слагаемое последней суммы (начиная с третьего) возрастает:

$$\frac{1}{1\cdot 2}\left(1-\frac{1}{n}\right)<\frac{1}{1\cdot 2}\left(1-\frac{1}{n+1}\right) \text{ и т.д.}$$

и добавляется еще один (положительный) член.

Итак, последовательность $\{a_n\}$ монотонно возрастает.

Докажем теперь, что $\{a_n\}$ ограничена. Очевидно, $1-\frac{1}{n}<1$, $\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)<1$ и т.д. Поэтому из (*) получаем неравенство

$$\left(1+\frac{1}{n}\right)^n<1+1+\frac{1}{1\cdot 2}+\frac{1}{1\cdot 2\cdot 3}+\ldots+\frac{1}{1\cdot 2\cdot 3\cdots n}.$$

Замечая, что

$$\frac{1}{1\cdot 2\cdot 3}<\frac{1}{2^2},\quad \frac{1}{1\cdot 2\cdot 3\cdot 4}<\frac{1}{2^3},\ \ldots,\ \frac{1}{1\cdot 2\cdot 3\cdots n}<\frac{1}{2^{n-1}},$$

получаем

$$\left(1+\frac{1}{n}\right)^n < 1+1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^{n-1}}.$$

Члены правой части, начиная со второго, образуют геометрическую прогрессию:

$$1+\frac{1}{2}+\frac{1}{2^2}+\ldots+\frac{1}{2^{n-1}}=\frac{1-\left(\frac{1}{2}\right)^n}{1-\frac{1}{2}}=2-\left(\frac{1}{2}\right)^{n-1}<2.$$

Следовательно,

$$\left(1+\frac{1}{n}\right)^n < 3.$$

Итак,

$$2 \le \left(1+\frac{1}{n}\right)^n < 3.$$

Мы доказали, что последовательность $\{a_n\}$, где $a_n=\left(1+\frac{1}{n}\right)^n$, является возрастающей и ограниченной. Поэтому на основании теоремы 14.4 она имеет предел.

Определение. Предел $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ называется **числом e**:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \mathrm{e}. \tag{14.9}$$

Предел (14.9) называется **вторым замечательным пределом**.

Число e – иррациональное число. Более того, оно является числом трансцендентным, т.е. не является корнем никакого алгебраического уравнения с целыми коэффициентами.

Известно, что

$$\mathrm{e} = 2{,}7182818284\ldots$$

В большинстве случаев на практике полагают $\mathrm{e} \approx 2{,}72$.

Рассмотрим пример. Пусть первоначальный вклад в банк составил S_0 денежных единиц. Банк выплачивает ежегодно $p\%$. Тогда по истечении года сумма вклада составит $S_0+\frac{p}{100}S_0=S_0\left(1+\frac{p}{100}\right)$, т.е. умножится на $\left(1+\frac{p}{100}\right)$. Через два года она снова умножится на $\left(1+\frac{p}{100}\right)$ и составит $S_0\left(1+\frac{p}{100}\right)^2$ и т.д. Таким образом, при $p\%$ годовых по истечении n лет сумма вклада будет равна

$$S_n=S_0\left(1+\frac{p}{100}\right)^n. \qquad (14.10)$$

Это **формула сложных процентов** (см. пример 12.3).

Следует заметить, что проценты по вкладу могут начисляться не обязательно раз в году, а, например, ежеквартально, ежемесячно, каждый день. Формула (14.10) позволяет вычислить сумму вклада S_n по истечении n периодов при процентной ставке $p\%$ за период (независимо от того, какова длина этих периодов).

Представим себе, что банк, находящийся в Москве, закончив рабочий день, перечисляет (учитывая разницу во времени) некоторую сумму S_0 банку, находящемуся во Владивостоке, на 12 часов – с 20 часов текущего дня до 8 часов следующего дня по московскому времени. Владивостокский банк возвращает деньги к началу работы московского банка, выплачивая 1% за пользование этим краткосрочным кредитом. Затем, на следующий день, московский банк повторяет эту операцию, но уже с полученной суммой в 101% от S_0 и т.д. (Такой договор между банками вряд ли возможен, однако ставка 1% в день в начале 1990-х годов была реальна.)

По истечении 300 дней московский банк получит сумму

$$S_{300}=S_0\left(1+\frac{1}{100}\right)^{300}=S_0\left[\left(1+\frac{1}{100}\right)^{100}\right]^3\approx S_0\mathrm{e}^3\approx S_0\cdot 2{,}7^3\approx 19{,}68S_0,$$

т.е. первоначальная сумма увеличится за год почти в 20 раз.

В общем случае пусть сумма S_0 помещена в банк на t лет под $p\%$ годовых и проценты начисляются n раз в году. Тогда процентная ставка за $\frac{1}{n}$ часть года составит $\frac{p}{n}\%$, а размер вклада за t лет (при nt начислениях) составит

$$S_n=S_0\left(1+\frac{p}{100n}\right)^{nt},$$

или, если обозначить $\frac{p}{100}=r$:

$$S_n = S_0\left(1+\frac{r}{n}\right)^{nt}.$$

Преобразуем последнее выражение:

$$S_n = S_0\left[\left(1+\frac{r}{n}\right)^{\frac{n}{r}}\right]^{rt}.$$

Введем обозначение: $\frac{n}{r}=m$. Пусть $n\to\infty$. Тогда $m\to\infty$. Получаем

$$S=\lim_{n\to\infty} S_n = \lim_{m\to\infty} S_0\left[\left(1+\frac{1}{m}\right)^{m}\right]^{rt} = S_0\mathrm{e}^{rt}.$$

Эта формула

$$S=S_0\mathrm{e}^{rt}, \quad \text{или} \quad S=S_0\mathrm{e}^{\frac{p}{100}t},$$

называется **формулой непрерывных процентов**.

Рассмотрим теперь функцию

$$f(x)=\left(1+\frac{1}{x}\right)^{x}.$$

Здесь x изменяется непрерывно, принимая любые (а не только натуральные) значения.

Теорема **14.5.** Предел функции $f(x)=\left(1+\frac{1}{x}\right)^{x}$ при $x\to\infty$ существует и равен e.

Доказательство. Пусть $x\to+\infty$. Для каждого значения x существует такое натуральное число n, что

$$n\le x\le n+1.$$

Из этих неравенств получаем:

$$\frac{1}{n}\ge\frac{1}{x}\ge\frac{1}{n+1},$$

$$1+\frac{1}{n}\ge1+\frac{1}{x}\ge1+\frac{1}{n+1},$$

$$\left(1+\frac{1}{n}\right)^{n+1}\ge\left(1+\frac{1}{x}\right)^{x}\ge\left(1+\frac{1}{n+1}\right)^{n}.$$

Очевидно, если $x \to +\infty$, то $n \to \infty$. Найдем пределы переменных, между которыми заключена функция $\left(1+\frac{1}{x}\right)^x$:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n\left(1+\frac{1}{n}\right) = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \lim_{n\to\infty}\left(1+\frac{1}{n}\right) =$$

$$= \mathrm{e}\cdot 1 = \mathrm{e};$$

$$\lim_{n\to\infty}\left(1+\frac{1}{n+1}\right)^n = \lim_{n\to\infty}\frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{1+\frac{1}{n+1}} = \ldots = \frac{\mathrm{e}}{1} = \mathrm{e}.$$

Итак, обе переменные величины, между которыми заключена функция $\left(1+\frac{1}{x}\right)^x$, имеют один и тот же предел e. Следовательно, по теореме 14.3

$$\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = \mathrm{e}.$$

Для случая $x \to -\infty$ с помощью замены $x = -y$ легко доказать, что также

$$\lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = \mathrm{e}.$$

Итак,

$$\lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^x = \mathrm{e}. \qquad (14.11)$$

Сделаем в (14.11) замену $\frac{1}{x}=\alpha$. Тогда $x\to\pm\infty$ равносильно $\alpha\to 0$. Получаем

$$\lim_{\alpha\to 0}(1+\alpha)^{\frac{1}{\alpha}} = \mathrm{e}. \qquad (14.12)$$

Мы получили для числа e три формулы: (14.9), (14.11) и (14.12). Число e – одна из фундаментальных величин в математике.

Показательная функция с основанием e (рис. 14.3)

$y = \mathrm{e}^x$

(и вообще функция вида $y = \mathrm{e}^{ax}$) играет важную роль в математике и ее приложениях. Она применяется в статистике, физике, химии, в исследовании демографических процессов и т.п. Эту функцию (и ее график) называют **экспонентой**.

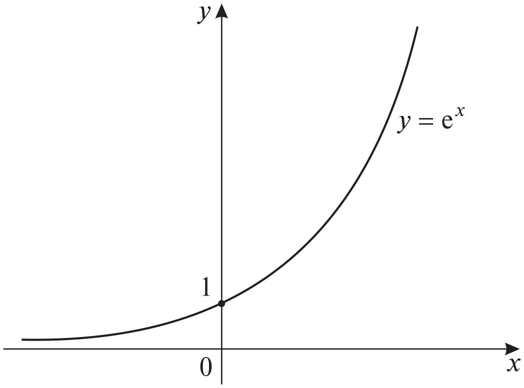


Рис. 14.3. Экспонента

Логарифм с основанием e называется **натуральным логарифмом** и обозначается символом ln: $\log_{\mathrm{e}} x = \ln x$.

Вопросы

1. Что такое общий член последовательности?
2. Могут ли в числовой последовательности различным номерам соответствовать одинаковые числа?
3. Пусть все члены монотонной последовательности умножили на –1. Будет ли полученная последовательность монотонной?
4. Пусть число 4 является пределом числовой последовательности. Можно ли утверждать, что вне интервала (3, 5) содержится лишь конечное число членов последовательности?
5. Имеет ли предел последовательность 1, 0, 1, 0, 1, 0, ...?
6. Пусть число 5 является пределом числовой последовательности. Может ли эта последовательность иметь отрицательные члены?

7. Может ли число -1 быть пределом числовой последовательности, все члены которой положительны?
8. Может ли последовательность иметь два различных предела?
9. Пусть для всех x выполняется неравенство $f(x) \geq 5{,}001$. Может ли в этом случае быть $\lim\limits_{x \to 1} f(x) = 5$?
10. Что такое число e?
11. Какая функция называется экспонентой? Какая кривая называется экспонентой?

Глава 15
НЕПРЕРЫВНОСТЬ ФУНКЦИИ

15.1. ОСНОВНЫЕ ПОНЯТИЯ

Понятие непрерывности функции является одним из основных понятий математического анализа.

***Определение* 1.** Функция $f(x)$ называется **непрерывной** в точке x_0, если выполняются следующие три условия: 1) $f(x)$ определена в точке x_0; 2) существует конечный предел $f(x)$ при $x \to x_0$; 3) этот предел равен значению функции в точке x_0, т.е.

$$\lim_{x \to x_0} f(x) = f(x_0). \tag{15.1}$$

Заметим, что равенство (15.1) и условие непрерывности $f(x)$ в точке x_0 можно записать в виде

$$\lim_{x \to x_0} f(x) = f\left(\lim_{x \to x_0} x\right).$$

С геометрической точки зрения непрерывная функция – это функция, график которой есть непрерывная кривая.

Существует несколько эквивалентных определений непрерывности.

Обозначим разность $x - x_0$ через Δx. Будем говорить, что при переходе от значения x_0 к значению x аргумент получает *приращение* $\Delta x = x - x_0$. При этом функция $y = f(x)$ получает соответствующее приращение $\Delta y = f(x_0 + \Delta x) - f(x_0)$. С учетом сказанного равенство (15.1) равносильно равенству

$$\lim_{\Delta x \to 0} \Delta y = 0.$$

***Определение* 2.** Функция $y = f(x)$ называется **непрерывной** в точке x_0, если она определена в этой точке и некоторой ее окрестности и выполняется равенство

$$\lim_{\Delta x \to 0} \Delta y = 0. \tag{15.2}$$

(Это определение легко запоминается в следующей форме: *функция непрерывна, если бесконечно малому приращению аргумента соответствует бесконечно малое приращение функции*.)

***Пример* 15.1.** Покажем, что функция $y = \sin x$ непрерывна в произвольной точке x. Придадим аргументу приращение Δx. Тогда функция получит приращение

$$\Delta y = \sin(x + \Delta x) - \sin x = 2\cos\frac{(x + \Delta x) + x}{2}\sin\frac{(x + \Delta x) - x}{2} =$$

$$= 2\cos\left(x + \frac{\Delta x}{2}\right)\sin\left(\frac{\Delta x}{2}\right).$$

Если $\Delta x \to 0$, то $\sin\frac{\Delta x}{2} \to 0$ (так как $\left|\sin\frac{\Delta x}{2}\right| < \left|\frac{\Delta x}{2}\right|$); при этом $\cos\left(x + \frac{\Delta x}{2}\right)$ ограничена. Поэтому

$$\lim_{\Delta x \to 0} \Delta y = \lim_{\Delta x \to 0} 2\cos\left(x + \frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2} = 0.$$

Следовательно, функция $y = \sin x$ непрерывна.

Аналогично можно доказать, что *любая основная элементарная функция непрерывна в каждой точке, в которой она определена*.

Справедливы следующие ***утверждения***:

1. Если функции $f(x)$ и $g(x)$ непрерывны в точке x_0, то их сумма $\varphi(x) = f(x) + g(x)$ также непрерывна в этой точке.

2. Произведение двух непрерывных функций есть непрерывная функция.

3. Частное двух непрерывных функций есть функция непрерывная, если знаменатель в рассматриваемой точке не обращается в нуль (т.е. если $f(x)$ и $g(x)$ непрерывны в точке x_0 и $g(x_0) \neq 0$, то $\varphi(x) = \frac{f(x)}{g(x)}$ непрерывна в точке x_0).

4. Если $u = \varphi(x)$ непрерывна при $x = x_0$ и $f(u)$ непрерывна в точке $u_0 = \varphi(x_0)$, то сложная функция $y = f(\varphi(x))$ непрерывна в точке x_0.

Доказательства этих утверждений просты и основаны на свойствах пределов.

Д о к а ж е м, например, утверждение 2. Пусть $f(x)$ и $g(x)$ непрерывны в точке x_0: $\lim\limits_{x\to x_0} f(x)=f(x_0)$, $\lim\limits_{x\to x_0} g(x)=g(x_0)$, и пусть $\varphi(x)=$ $=f(x)g(x)$. Так как предел произведения равен произведению пределов, то $\lim\limits_{x\to x_0}\varphi(x)=\lim\limits_{x\to x_0} f(x)\lim\limits_{x\to x_0} g(x)=f(x_0)g(x_0)=\varphi(x_0)$. Итак, $\lim\limits_{x\to x_0}\varphi(x)=\varphi(x_0)$, т.е. $\varphi(x)$ непрерывна в точке x_0.

***Т е о р е м а* 15.1.** Всякая элементарная функция непрерывна в каждой точке, в которой она определена.

Доказательство следует из утверждений 1–4, сформулированных выше*.

Если функция $f(x)$ не является непрерывной в точке x_0, то точка x_0 называется **точкой разрыва** функции $f(x)$. Различают **точки разрыва первого рода**, когда существуют конечные пределы $\lim\limits_{x\to x_0-} f(x)$ и $\lim\limits_{x\to x_0+} f(x)$, и **точки разрыва второго рода**, когда хотя бы один из этих односторонних пределов бесконечен или не существует. Среди точек разрыва первого рода следует отметить также **точки устранимого разрыва**, когда предел функции $f(x)$ при $x\to x_0$ существует, но или он не равен $f(x_0)$, или функция не определена при $x=x_0$.

***П р и м е р* 15.2.**

1. $f(x)=\operatorname{arctg}\dfrac{1}{x}$. Здесь $x_0=0$ – точка разрыва первого рода, так как $\lim\limits_{x\to 0-}\operatorname{arctg}\dfrac{1}{x}=-\dfrac{\pi}{2}$, $\lim\limits_{x\to 0+}\operatorname{arctg}\dfrac{1}{x}=\dfrac{\pi}{2}$.

2. $f(x)=\dfrac{1}{x}$. Здесь $x_0=0$ – точка разрыва второго рода, поскольку $\lim\limits_{x\to 0+}\dfrac{1}{x}=+\infty$, $\lim\limits_{x\to 0-}\dfrac{1}{x}=-\infty$.

3. $f(x)=\begin{cases}\dfrac{\sin x}{x}, & x\neq 0,\\ 0, & x=0.\end{cases}$ Здесь $x_0=0$ – точка устранимого разрыва, так как существует $\lim\limits_{x\to 0}\dfrac{\sin x}{x}=1$. Этот разрыв можно устранить, если изменить значение функции в точке $x=0$, положив $f(0)=1$.

* С учетом того, что основные элементарные функции непрерывны.

Если функция $y = f(x)$ непрерывна в каждой точке некоторого промежутка, то говорят, что эта функция **непрерывна на этом промежутке**.

15.2. СВОЙСТВА ФУНКЦИЙ, НЕПРЕРЫВНЫХ НА ОТРЕЗКЕ

***Теорема* 15.2 (первая теорема Вейерштрасса).** Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она ограничена на этом отрезке.

***Теорема* 15.3 (вторая теорема Вейерштрасса).** Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она достигает на этом отрезке наименьшего значения m и наибольшего значения M (т.е. существуют на этом отрезке точка c_1, в которой $f(c_1) = m$, и точка c_2, в которой $f(c_2) = M$).

***Теорема* 15.4 (первая теорема Больцано – Коши).** Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и на его концах принимает значения разных знаков*. Тогда внутри $[a, b]$ существует такая точка c, что $f(c) = 0$.

***Теорема* 15.5 (вторая теорема Больцано – Коши).** Пусть функция $f(x)$ непрерывна на $[a, b]$ и пусть m – наименьшее, а M – наибольшее значения $f(x)$ на $[a, b]$. Тогда для любого C, удовлетворяющего условию $m < C < M$, существует внутри $[a, b]$ такая точка c, что $f(c) = C$.

Доказать эти теоремы непросто, и мы это делать не будем. Однако все они являются частными случаями следующего ***утверждения*** (которое интуитивно кажется очевидным): *если функция $f(x)$ непрерывна на отрезке $[a, b]$, то область ее изменения есть отрезок.*

15.3. ЭКОНОМИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ НЕПРЕРЫВНОСТИ

Большинство функций, применяемых в экономике, являются непрерывными. Таковы, в частности, функции спроса и предложения, функция полезности, функция выпуска (см. § 13.4). Среди функций, используемых в экономике, есть и разрывные функции.

---

* Например, $f(a) < 0$, $f(b) > 0$.

1. *Налоговая ставка* (рис. 15.1) – функция, выражающая зависимость налоговой ставки N в процентах от величины годового дохода q. Эта функция разрывна на концах промежутков, и разрывы эти первого рода.

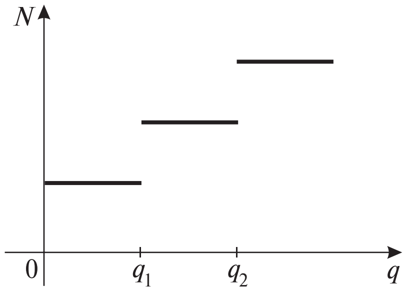


Рис. 15.1. Налоговая ставка

Однако сама величина подоходного налога P является непрерывной функцией годового дохода q (рис. 15.2).

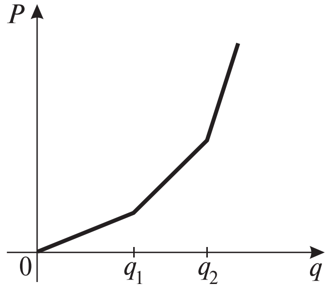


Рис. 15.2. Подоходный налог

Из непрерывности функции $P = P(q)$ следует, в частности, что если доходы налогоплательщиков отличаются незначительно, то и различие в величинах их подоходного налога также невелико.

2. Как известно, две основные категории рыночных отношений – *спрос* и *предложение*. И то и другое зависит от многих факторов, среди которых главный – это *цена* товара. Обозначим цену товара p (price), объем спроса – D (demand), а величину предложения – S (supply). По своему смыслу функции $D = D(p)$ и $S = S(p)$

непрерывно зависят от p. Значит, при малых колебаниях цен спрос и предложение меняются незначительно. Однако иногда спрос меняется скачкообразно. Происходит это обычно по причинам психологического характера, в частности, при «пробитии» круглой цены. Бывает, что при росте цены на некоторый товар спрос некоторое время уменьшается незначительно, но, как только цена превысит определенную сумму (например, 100 денежных единиц), спрос резко падает. В этом случае функция $D = D(p)$ при указанном значении p имеет разрыв.

При анализе функций $D = D(p)$ и $S = S(p)$ (если рассматривать их на таком ценовом отрезке, где они не имеют разрывов) можно воспользоваться свойствами непрерывных функций (см. § 15.2). Рассмотрим разность $D(p) - S(p)$.

При малых p, очевидно, $D(p) - S(p) > 0$ (спрос превышает предложение), а при больших p, наоборот, $D(p) - S(p) < 0$. Применяя к разности $D(p) - S(p)$ первую теорему Больцано – Коши (см. § 15.2), приходим к заключению, что существует такая цена p_0, для которой $D(p_0) - S(p_0) = 0$, т.е. $D(p_0) = S(p_0)$. Эта цена p_0 называется **равновесной** (мы упоминали о ней в § 13.4).

15.4. СРАВНЕНИЕ БЕСКОНЕЧНО МАЛЫХ

Пусть одновременно рассматриваются несколько бесконечно малых величин α, β, γ, ..., которые являются функциями одного и того же аргумента x, и величины эти являются бесконечно малыми при стремлении x к некоторому конечному пределу a или к бесконечности.

Во многих случаях представляет интерес сравнение этих бесконечно малых между собой по характеру их стремления к нулю. Речь идет о сравнительной «скорости» их стремления к нулю: какая из бесконечно малых стремится к нулю «быстрее», а какая — «медленнее».

Для сравнения двух бесконечно малых обычно изучают поведение их о т н о ш е н и я. При этом, рассматривая отношение $\frac{\alpha}{\beta}$ (или $\frac{\beta}{\alpha}$), предполагают, что переменная, стоящая в знаменателе, не обращается в нуль, по крайней мере для значений x, достаточно близких к a (или для достаточно больших по модулю при $x \to \infty$).

I. Если отношение $\frac{\beta}{\alpha}$ имеет конечный предел, отличный от нуля, то бесконечно малые α и β называются **бесконечно малыми одного порядка**.

В этом случае, очевидно, и отношение $\frac{\alpha}{\beta}$ имеет конечный предел.

Пример **15.3.** Бесконечно малые $\alpha = 3x$ и $\beta = \sin 2x$ являются бесконечно малыми одного порядка при $x \to 0$, так как (см. пример 14.2)

$$\lim_{x\to 0}\frac{\beta}{\alpha} = \lim_{x\to 0}\frac{\sin 2x}{3x} = \frac{2}{3}.$$

Пример **15.4.** Бесконечно малые $\alpha = x$ и $\beta = \sqrt{1+x}-1$ также являются бесконечно малыми одного порядка, так как

$$\lim_{x\to 0}\frac{\beta}{\alpha} = \lim_{x\to 0}\frac{\sqrt{1+x}-1}{x} = \lim_{x\to 0}\frac{x}{x\left(\sqrt{1+x}+1\right)} = \lim_{x\to 0}\frac{1}{\sqrt{1+x}+1} = \frac{1}{2}.$$

II. Если отношение $\frac{\beta}{\alpha}$ само оказывается бесконечно малым, т.е. $\lim\limits_{x\to a}\frac{\beta(x)}{\alpha(x)} = 0$ $\left(\lim\limits_{x\to a}\frac{\alpha(x)}{\beta(x)} = \infty\right)$, то говорят, что бесконечно малая β является **бесконечно малой высшего порядка** относительно бесконечно малой α (а бесконечно малая α – **бесконечно малой низшего порядка** относительно бесконечно малой β).

Пример **15.5.** Бесконечно малая $\beta = 1-\cos 2x$ есть бесконечно малая высшего порядка относительно $\alpha = x$. Действительно,

$$\lim_{x\to 0}\frac{1-\cos 2x}{x} = \lim_{x\to 0}\frac{2\sin^2 x}{x} = 0.$$

Заметим, что если β есть бесконечно малая высшего порядка относительно бесконечно малой α, то это обстоятельство записывают так:

$$\beta = o(\alpha).$$

В частности, пример 15.5 показывает, что

$$1-\cos 2x = o(x).$$

III. Бесконечно малая β называется **бесконечно малой *k*-го порядка** относительно бесконечно малой α, если β и α^k – бесконечно малые одного порядка, т.е. если существует конечный предел отношения $\frac{\beta}{\alpha^k}$, отличный от нуля.

Пример **15.6.** Если $\alpha = x$, а $\beta = 1 - \cos x$, то при $x \to 0$ бесконечно малая β есть бесконечно малая второго порядка относительно бесконечно малой α. Действительно,

$$\lim_{x\to 0}\frac{\beta}{\alpha^2} = \lim_{x\to 0}\frac{1-\cos x}{x^2} = \lim_{x\to 0}\frac{2\sin^2\frac{x}{2}}{x^2} = \frac{1}{2}.$$

Пример **15.7.** Если $\alpha = x$, а $\beta = \sqrt{1+x^3} - 1$, то при $x \to 0$ бесконечно малая β есть бесконечно малая третьего порядка относительно бесконечно малой α. Убедимся в этом:

$$\lim_{x\to 0}\frac{\beta}{\alpha^3} = \lim_{x\to 0}\frac{\sqrt{1+x^3}-1}{x^3} = \lim_{x\to 0}\frac{x^3}{x^3\left(\sqrt{1+x^3}+1\right)} = \frac{1}{2}.$$

IV. Бесконечно малые α и β называются **эквивалентными**, если предел их отношения равен единице:

$$\lim\frac{\beta}{\alpha} = 1.$$

Если α и β эквивалентны, то пишут $\alpha \sim \beta$.

Теорема **15.6.** Для того чтобы бесконечно малые α и β были эквивалентны, необходимо и достаточно, чтобы их разность $\gamma = \beta - \alpha$ была бесконечно малой высшего порядка относительно α и β.

Доказательство. 1. *Необходимость*. Пусть $\gamma = o(\alpha)$, $\gamma = o(\beta)$. Тогда $\frac{\gamma}{\alpha} = \frac{\beta}{\alpha} - 1$, откуда $\lim\frac{\beta}{\alpha} = 1$ (так как $\lim\frac{\gamma}{\alpha} = 0$).

2. *Достаточность*. Пусть $\alpha \sim \beta$, т.е. $\lim\frac{\beta}{\alpha} = 1$. Тогда из равенства $\frac{\gamma}{\alpha} = \frac{\beta}{\alpha} - 1$ получаем $\lim\frac{\gamma}{\alpha} = 1 - 1 = 0$.

Заметим, что предел отношения бесконечно малых может вообще не существовать. В этом случае говорят, что бесконечно малые несравнимы между собой. Рассмотрим традиционный пример.

Пример **15.8.** Бесконечно малые $\alpha = x$ и $\beta = x\sin\frac{1}{x}$ $(x \to 0)$ несравнимы между собой. Действительно, отношение этих бесконечно малых $\frac{\beta}{\alpha} = \sin\frac{1}{x}$ не имеет предела при $x \to 0$.

Вопросы

1. Какие из основных элементарных функций непрерывны?
2. Как классифицируются точки разрыва функции?
3. Является ли налоговая ставка непрерывной функцией от величины дохода?
4. Является ли величина подоходного налога непрерывной функцией годового дохода?
5. Пусть $D = D(p)$ – функция, выражающая зависимость спроса D от цены p. Разрыв какого рода имеет функция $D = D(p)$ при скачкообразном изменении спроса?
6. На чем основано сравнение бесконечно малых?
7. Являются ли эквивалентные бесконечно малые бесконечно малыми одного порядка?
8. Пусть α и β – две бесконечно малые разных порядков. Какая из них быстрее стремится к нулю – та, что более высокого порядка, или та, что более низкого порядка?
9. Любые ли две бесконечно малые величины сравнимы между собой?

Раздел V
ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ

Глава 16
ПРОИЗВОДНАЯ ФУНКЦИИ. ДИФФЕРЕНЦИАЛ

16.1. ПРОИЗВОДНАЯ

Пусть функция $y = f(x)$ определена на некотором промежутке X. Придадим значению аргумента $x_0 \in X$ произвольное приращение Δx так, чтобы точка $x_0 + \Delta x$ также принадлежала X. Тогда функция $f(x)$ получит соответствующее приращение $\Delta y = f(x_0 + \Delta x) - f(x_0)$.

Определение. **Производной функции** $y = f(x)$ **в точке** x_0 называется предел отношения приращения функции в этой точке к приращению аргумента при $\Delta x \to 0$:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

(при условии, что этот предел существует).

Производная обозначается $f'(x_0)$, или $y'(x_0)$, или просто y', или $\frac{dy}{dx}$. (Экономисты применяют также обозначение $Mf(x)$ для производной $f'(x)$ и термин «маржинальное значение функции f в точке x»).

Итак, по определению

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \qquad (16.1)$$

Если функция $f(x)$ имеет производную в каждой точке множества X, то производная $f'(x)$ также является функцией от аргумента x, определенной на X.

Геометрический смысл производной

Для выяснения геометрического смысла производной необходимо сформулировать определение касательной к графику функции в данной точке.

Определение. **Касательной** к графику функции $y = f(x)$ в точке M_0 называется предельное положение секущей M_0M при стремлении точки M к точке M_0 по кривой $y = f(x)$.

Пусть на кривой $y = f(x)$ зафиксирована точка $M_0(x_0, y_0)$, где $y_0 = f(x_0)$ (рис. 16.1). Придадим аргументу приращение Δx, т.е. перейдем от значения $x = x_0$ к значению $x_0 + \Delta x$. Получим на кривой точку $M(x_0 + \Delta x, y_0 + \Delta y)$. Из треугольника M_0MA имеем

$$\operatorname{tg}\varphi = \frac{MA}{M_0A} = \frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

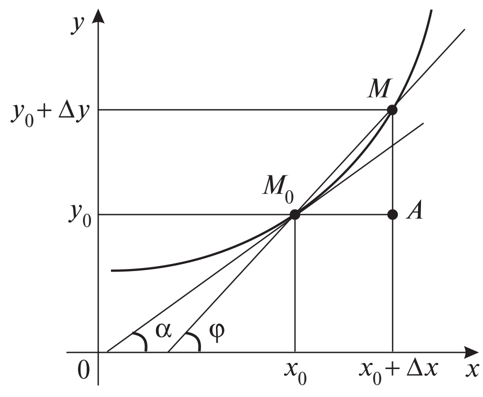


Рис. 16.1. Геометрический смысл производной

Пусть $\Delta x \to 0$. Тогда точка M будет перемещаться вдоль кривой и в пределе совпадет с точкой M_0. При этом

$$\operatorname{tg}\alpha = \lim_{\Delta x \to 0} \operatorname{tg}\varphi = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Если производная $f'(x)$ в точке x_0 существует, то, согласно определению производной, получаем

$$\operatorname{tg}\alpha = f'(x_0).$$

Итак, производная $f'(x_0)$ равна *угловому коэффициенту** *касательной* к графику функции $y = f(x)$ в точке $M_0(x_0, f(x_0))$.

Физический смысл производной

Предположим, что функция $s = f(t)$ описывает закон движения точки по прямой как зависимость пройденного пути s от времени t. К моменту времени t_0 пройденный путь равен $s_0 = f(t_0)$, а к моменту $t_0 + \Delta t$ он равен $s = f(t_0 + \Delta t)$. Тогда за промежуток Δt пройден путь $\Delta s = s - s_0$, а средняя скорость за время Δt есть отношение $\dfrac{\Delta s}{\Delta t}$. Предел этого отношения при $\Delta t \to 0$, т.е.

$$\lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = f'(t),$$

определяет *мгновенную скорость* точки в момент времени t_0 как производную пути по времени.

16.2. ПРИМЕНЕНИЕ ПРОИЗВОДНОЙ В ЭКОНОМИКЕ

1. Производительность труда. Пусть функция $q = q(t)$ выражает объем q произведенной за время t продукции и пусть требуется найти производительность труда в момент t_0. Рассмотрим период времени от момента t_0 до $t_0 + \Delta t$. За этот период объем произведенной продукции составит $\Delta q = q(t_0 + \Delta t) - q(t_0)$. *Средняя производительность* за этот период равна $\dfrac{\Delta q}{\Delta t}$. Тогда производительность труда в момент t_0 можно определить как предельное значение средней производительности при $\Delta t \to 0$:

$$q'(t) = \lim_{\Delta t \to 0} \frac{\Delta q}{\Delta t}.$$

Как видим, математически задача о производительности в момент t_0 не отличается от задачи нахождения мгновенной скорости движения (см. § 16.1).

---

* Угловой коэффициент есть тангенс угла наклона к положительному направлению оси Ox.

Возможна и другая постановка задачи о производительности. Пусть количество продукции q зависит только от приложенного труда x (для фирмы это просто численность персонала): $q = q(x)$.

Для оценки эффективности производства часто применяется *средняя производительность*, которая в данном случае определяется в виде отношения $\frac{q}{x}$.

Однако возникает вопрос: как изменится объем продукции при изменении численности персонала? Ответ на этот вопрос можно получить, введя понятие **предельной производительности**. Это — производная от продукции q по величине приложенного труда x:

$$q' = \frac{dq}{dx}.$$

Предельная производительность при такой постановке задачи приближенно равна *изменению объема выпускаемой продукции при изменении численности персонала на единицу*.

Если число работников a велико, то приращение $\Delta a = 1$ можно считать достаточно малым, чтобы воспользоваться приближенным равенством $q'(a) \approx \frac{\Delta q}{\Delta a} = \frac{q(a+1)-q(a)}{1} = q(a+1)-q(a)$, откуда $q(a + 1) = q(a) + q'(a)$. В данном случае $q'(a)$ есть дополнительная продукция, произведенная новым сотрудником за единицу времени.

Пусть v — цена продукции, а p — зарплата работника за единицу времени. Тогда если $vq'(a) > p$, то надо нанять еще одного работника, так как он приносит фирме больше, чем она ему платит. Это правило называется **«золотым» правилом экономики**.

2. Себестоимость продукции. Рассмотрим зависимость себестоимости C произведенной продукции от ее объема q: $C = C(q)$.

Предельной себестоимостью называют величину

$$MC \approx \lim_{\Delta q \to 0} \frac{\Delta C}{\Delta q} = C'(q).$$

Наряду с себестоимостью в микроэкономике важную роль играет еще один предельный показатель — *эластичность*. Рассмотрим его позднее — при изучении так называемой логарифмической производной (см. § 20.2).

16.3. ДИФФЕРЕНЦИРУЕМОСТЬ ФУНКЦИИ. СВЯЗЬ МЕЖДУ ДИФФЕРЕНЦИРУЕМОСТЬЮ И НЕПРЕРЫВНОСТЬЮ

Определение. Функция $y = f(x)$ называется **дифференцируемой в точке** x_0, если ее приращение Δy в этой точке можно представить в виде

$$\Delta y = A\Delta x + \alpha\Delta x, \tag{16.2}$$

где A – некоторое число (не зависящее от Δx), а $\alpha = \alpha(\Delta x)$ – бесконечно малая при $\Delta x \to 0$.

***Теорема* 16.1.** Для того чтобы функция $y = f(x)$ была дифференцируемой в точке x_0, необходимо и достаточно, чтобы она имела в этой точке конечную производную.

Доказательство. 1. *Необходимость*. Пусть дано, что функция $f(x)$ дифференцируема в точке x_0, т.е. ее приращение можно представить в виде (16.2). Поделим это равенство на $\Delta x \neq 0$:

$$\frac{\Delta y}{\Delta x} = A + \alpha.$$

Переходя к пределу при $\Delta x \to 0$ (и учитывая, что $\alpha \to 0$ при $\Delta x \to 0$), получаем

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0) = \lim_{\Delta x \to 0} (A + \alpha) = A,$$

т.е. производная функции $f(x)$ в точке x_0 существует и равна A.

2. *Достаточность*. Пусть теперь дано, что функция $f(x)$ имеет производную в точке x_0, т.е. существует предел

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = A.$$

Тогда в соответствии с теоремой 14.1 о связи переменной величины с ее пределом

$$\frac{\Delta y}{\Delta x} = A + \alpha,$$

где α – бесконечно малая при $\Delta x \to 0$. Отсюда

$$\Delta y = A\Delta x + \alpha\Delta x.$$

А это и означает, что $f(x)$ дифференцируема.

Теорема 16.1 позволяет называть функцию одного аргумента дифференцируемой, если она имеет производную. Операция нахождения производной называется **дифференцированием**.

Непрерывность дифференцируемой функции

Теорема **16.2.** Если функция дифференцируема в точке x_0, то она непрерывна в этой точке.

Доказательство. Так как $f(x)$ дифференцируема в точке x_0, то ее приращение в этой точке имеет вид (16.2). Переходя к пределу в этом равенстве при $\Delta x \to 0$, получаем

$$\lim_{\Delta x \to 0} \Delta y = \lim_{\Delta x \to 0} (A\Delta x + \alpha\Delta x) = 0,$$

т.е. $\lim\limits_{\Delta x \to 0} \Delta y = 0$, а это и означает, что функция непрерывна.

16.4. ВЫЧИСЛЕНИЕ ПРОИЗВОДНОЙ

Схема вычисления производной функции $f(x)$:

1. Придадим аргументу x приращение Δx и найдем соответствующее значение функции $f(x + \Delta x)$.
2. Найдем приращение функции $\Delta y = f(x + \Delta x) - f(x)$.
3. Составим отношение $\dfrac{\Delta y}{\Delta x}$.
4. Найдем предел этого отношения при $\Delta x \to 0$:

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Пример **16.1.** Если $y = c = \text{const}$, то $y' = 0$. Действительно, $\Delta y = 0$ при любом Δx, так что $y' = 0$. Итак, если $c = \text{const}$, то

$$c' = 0.$$

Пример **16.2.** Найти производную функции $y = x^2$.

Решение. 1. Придадим аргументу приращение Δx и найдем $f(x + \Delta x) = (x + \Delta x)^2$.

2. Найдем приращение функции:

$$\Delta y = (x + \Delta x)^2 - x^2 = x^2 + 2x\Delta x + \Delta x^2 - x^2 = 2x\Delta x + \Delta x^2.$$

3. Составим отношение $\dfrac{\Delta y}{\Delta x}$:

$$\frac{\Delta y}{\Delta x} = 2x + \Delta x.$$

4. Найдем предел:

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} (2x + \Delta x) = 2x.$$

Итак, $(x^2)' = 2x$.

***Пример* 16.3.** Найти производную функции $y = \sin x$.

Решение:

1. $f(x+\Delta x) = \sin(x+\Delta x)$.

2. $\Delta y = \sin(x+\Delta x) - \sin x = 2\cos\dfrac{(x+\Delta x)+x}{2}\sin\dfrac{(x+\Delta x)-x}{2} =$

$$= 2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right).$$

3. $\dfrac{\Delta y}{\Delta x} = \dfrac{2\sin\dfrac{\Delta x}{2}\cos\left(x+\dfrac{\Delta x}{2}\right)}{\Delta x} = \dfrac{2\sin\dfrac{\Delta x}{2}}{\Delta x}\cos\left(x+\dfrac{\Delta x}{2}\right) =$

$$= \frac{\sin\dfrac{\Delta x}{2}}{\dfrac{\Delta x}{2}}\cos\left(x+\frac{\Delta x}{2}\right).$$

4. $y' = \lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x} = \lim\limits_{\Delta x\to 0}\dfrac{\sin\dfrac{\Delta x}{2}}{\dfrac{\Delta x}{2}}\lim\limits_{\Delta x\to 0}\cos\left(x+\dfrac{\Delta x}{2}\right) = 1\cdot\cos x = \cos x$

(с учетом первого замечательного предела (см. § 14.5) и непрерывности функции $\cos x$).

Итак, $(\sin x)' = \cos x$.

Правила дифференцирования

Предполагаем, что $u = u(x)$ и $v = v(x)$ – дифференцируемые функции, $c = \text{const}$.

I. $(cu)' = cu'$.

II. $(u \pm v)' = u' \pm v'$.

III. $(uv)' = u'v + uv'$.

IV. $\left(\dfrac{u}{v}\right)' = \dfrac{u'v - uv'}{v^2}$.

Сформулируем эти правила более подробно и докажем их.

I. Если $y = cu(x)$, где $c = \text{const}$, то $y' = cu'(x)$, т.е. *постоянный множитель можно выносить за знак производной*.

Доказательство. 1. Придадим x приращение Δx. Тогда

$y + \Delta y = cu(x+\Delta x)$.

2. Найдем приращение функции:

$\Delta y = cu(x+\Delta x) - cu(x) = c[u(x+\Delta x) - u(x)]$.

3. Составим отношение $\frac{\Delta y}{\Delta x}$:

$$\frac{\Delta y}{\Delta x} = c\frac{u(x+\Delta x)-u(x)}{\Delta x}.$$

4. Вычислим предел этого отношения при $\Delta x \to 0$, т.е. вычислим y':

$$y' = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = c\lim_{\Delta x\to 0}\frac{u(x+\Delta x)-u(x)}{\Delta x} = cu'(x).$$

(Мы здесь воспользовались тем, что постоянный множитель можно выносить за знак предела (см. § 14.4).)

II. Если $y = u(x) \pm v(x)$, то $y' = u'(x) \pm v'(x)$, т.е. *производная алгебраической суммы дифференцируемых функций равна алгебраической сумме производных этих функций*.

Доказательство. 1. Придадим аргументу x приращение Δx. Тогда функции $u = u(x)$, $v = v(x)$ получат значения $u + \Delta u = u(x + \Delta x)$, $v + \Delta v = v(x + \Delta x)$, а тогда

$$y + \Delta y = (u + \Delta u) \pm (v + \Delta v).$$

2. Найдем приращение функции y:

$$\Delta y = (u + \Delta u) \pm (v + \Delta v) - (u \pm v) = \Delta u \pm \Delta v.$$

3. Составим отношение $\frac{\Delta y}{\Delta x}$:

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} \pm \frac{\Delta v}{\Delta x}.$$

4. Вычислим предел этого отношения при $\Delta x \to 0$:

$$y' = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x} \pm \lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x} = u' \pm v'.$$

Итак, $(u \pm v)' = u' \pm v'$.

(Здесь мы воспользовались тем, что предел алгебраической суммы равен алгебраической сумме пределов (см. § 14.4).)

Доказанное утверждение можно распространить на любое число слагаемых. В частности,

$$(u + v + w)' = u' + v' + w'.$$

III. Если $u = u(x)$, $v = v(x)$, то $(uv)' = u'v + uv'$, т.е. *производная произведения двух дифференцируемых функций равна произведению производной первой из этих функций на вторую плюс произведение первой функции на производную второй функции*.

Доказательство. 1. Придадим x приращение Δx. Тогда функции u и v получат значения $u + \Delta u$, $v + \Delta v$, а их произведение $y = uv$ — значение $(u + \Delta u)(v + \Delta v)$.

2. Найдем приращение функции y:

$$\Delta y = (u + \Delta u)(v + \Delta v) - uv = uv + \Delta uv + u\Delta v + \Delta u\Delta v - uv =$$

$$= \Delta uv + u\Delta v + \Delta u\Delta v.$$

3. Составим отношение $\dfrac{\Delta y}{\Delta x}$:

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x}v + u\frac{\Delta v}{\Delta x} + \Delta u\frac{\Delta v}{\Delta x}.$$

4. Вычислим предел при $\Delta x \to 0$:

$$y' = (uv)' = \lim_{\Delta x \to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0}\frac{\Delta u}{\Delta x}v + u\lim_{\Delta x \to 0}\frac{\Delta v}{\Delta x} + \lim_{\Delta x \to 0}\Delta u\lim_{\Delta x \to 0}\frac{\Delta v}{\Delta x} =$$

$$= u'v + uv' + 0 \cdot v' = u'v + uv'.$$

(Здесь $\lim\limits_{\Delta x \to 0}\Delta u = 0$, так как u, будучи дифференцируемой функцией, непрерывна.)

Итак, $(uv)' = u'v + uv'$.

На основании доказанного утверждения легко получить правило дифференцирования произведения трех и вообще любого конечного числа функций.

Пусть дано произведение трех функций $y = uvw$. Представим это произведение в виде $u(vw)$:

$$y' = u'(vw) + u(vw)' = u'vw + u(v'w + vw') = u'vw + uv'w + uvw'.$$

Легко понять, что для случая n сомножителей тем же приемом можно получить аналогичную формулу производной произведения:

$$(u_1u_2u_3\cdots u_n)' = u_1'u_2u_3\cdots u_n + u_1u_2'u_3\cdots u_n + u_1u_2u_3'\cdots u_n + \ldots + u_1u_2u_3\cdots u_n'.$$

IV. Если $y = \dfrac{u}{v}$, то $y' = \dfrac{u'v - uv'}{v^2}$,

т.е. *производная дроби (отношения двух функций) равна дроби, знаменатель которой есть квадрат знаменателя данной дроби, а числитель — разность между произведением знаменателя на производную числителя и произведением числителя на производную знаменателя.*

Доказательство. 1. Придадим аргументу x приращение Δx. Тогда функции u и v получат значения $u + \Delta u$, $v + \Delta v$, а их отношение $y = \dfrac{u}{v}$ — значение $y + \Delta y = \dfrac{u + \Delta u}{v + \Delta v}$.

2. $\Delta y = \dfrac{u+\Delta u}{v+\Delta v} - \dfrac{u}{v} = \dfrac{v\Delta u - u\Delta v}{v(v+\Delta v)}.$

3. $\dfrac{\Delta y}{\Delta x} = \dfrac{\dfrac{v\Delta u - u\Delta v}{\Delta x}}{v(v+\Delta v)} = \dfrac{\dfrac{\Delta u}{\Delta x}v - u\dfrac{\Delta v}{\Delta x}}{v(v+\Delta v)}.$

4. $y' = \lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x} = \dfrac{v\lim\limits_{\Delta x\to 0}\dfrac{\Delta u}{\Delta x} - u\lim\limits_{\Delta x\to 0}\dfrac{\Delta v}{\Delta x}}{v\lim\limits_{\Delta x\to 0}(v+\Delta v)}.$

Отсюда, учитывая, что $\lim\limits_{\Delta x\to 0}\Delta v = 0$ (так как функция v, будучи дифференцируемой, непрерывна), получаем

$$y' = \frac{u'v - uv'}{v^2}, \text{ или } \left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}.$$

Производная сложной функции

Теперь сформулируем и докажем правило V.

V. Пусть функция $u = \varphi(x)$ имеет в точке x_0 производную $u'_x = \varphi'(x_0)$, а функция $y = f(u)$ имеет в соответствующей точке $u_0 = \varphi(x_0)$ производную $y'_u = f'(u_0)$. Тогда сложная функция $y = f(\varphi(x))$ в упомянутой точке x_0 также имеет производную и справедлива следующая формула:

$$y'_x = y'_u u'_x.$$

Доказательство. Придадим x приращение Δx. Пусть Δu – соответствующее приращение функции $u = \varphi(x)$, Δy – приращение функции $y = f(u)$, вызванное приращением Δu. Воспользуемся равенством (16.2), которое, заменяя x на u, перепишем в виде

$$\Delta y = y'_u\,\Delta u + \alpha\Delta u$$

(здесь α зависит от Δu; $\alpha \to 0$ при $\Delta u \to 0$). Разделив это равенство почленно на Δx, получим

$$\frac{\Delta y}{\Delta x} = y'_u\frac{\Delta u}{\Delta x} + \alpha\frac{\Delta u}{\Delta x}.$$

Если $\Delta x \to 0$, то $\Delta u \to 0$ (так как u непрерывна), а тогда $\alpha \to 0$. Следовательно, существует предел

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = y'_u\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x} = y'_u u'_x,$$

т.е.

$$y'_x = y'_u u'_x.$$

Пример **16.4.** Найти производную функции $y = \sin^2 x$.

Решение: $y = u^2$, где $u = \sin x$. В соответствии с правилом V и с учетом примеров 16.2 и 16.3 получаем

$$y' = 2u \cdot u'_x = 2\sin x\,(\sin x)' = 2\sin x\cos x = \sin 2x.$$

Производная обратной функции

Пусть $y = f(x)$ – дифференцируемая и строго монотонная функция на некотором промежутке X, а функция $x = \varphi(y)$ является для нее обратной. Можно показать, что $\varphi(y)$ непрерывна на соответствующем промежутке Y.

Теорема **16.3.** Пусть функция $f(x)$ строго монотонна и непрерывна в промежутке X и в точке x_0 имеет конечную и отличную от нуля производную $f'(x_0)$. Тогда для обратной функции $x = \varphi(y)$ в соответствующей точке $y_0 = f(x_0)$ также существует производная, причем

$$\varphi'(y_0) = \frac{1}{f'(x_0)}. \tag{16.3}$$

Доказательство. Придадим значению $y = y_0$ произвольное приращение Δy. Тогда функция $x = \varphi(y)$ получит соответствующее приращение Δx. Заметим, что при $\Delta y \neq 0$ также и $\Delta x \neq 0$ – в силу однозначности функции $y = f(x)$. Имеем

$$\frac{\Delta x}{\Delta y} = \frac{1}{\dfrac{\Delta y}{\Delta x}}.$$

Пусть теперь $\Delta y \to 0$. Тогда $\Delta x \to 0$, так как $\varphi(y)$ непрерывна. Но тогда знаменатель правой части написанного равенства стремится к пределу $f'(x_0) \neq 0$, следовательно, существует предел для левой части равенства, равный $\dfrac{1}{f'(x_0)}$; он и есть производная $\varphi'(y)$. Итак,

$$x'_y = \frac{1}{y'_x}. \tag{16.4}$$

Последнее равенство можно переписать и в таком виде:

$$y'_x = \frac{1}{x'_y}. \tag{16.5}$$

16.5. ПРОИЗВОДНЫЕ ОСНОВНЫХ ЭЛЕМЕНТАРНЫХ ФУНКЦИЙ

Производная логарифмической функции

Выведем сначала формулу для производной функции $\boldsymbol{y = \ln x}$:

$$\Delta y = \ln(x + \Delta x) - \ln x = \ln \frac{x + \Delta x}{x} = \ln\left(1 + \frac{\Delta x}{x}\right),$$

$$\frac{\Delta y}{\Delta x} = \frac{1}{\Delta x} \ln\left(1 + \frac{\Delta x}{x}\right),$$

$$y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{\Delta x} \ln\left(1 + \frac{\Delta x}{x}\right).$$

Обозначим $\dfrac{\Delta x}{x} = t$; отсюда $\Delta x = tx$. Очевидно, $\Delta x \to 0$ тогда и только тогда, когда $t \to 0$. Получаем

$$y' = \lim_{t \to 0} \frac{1}{tx} \ln(1 + t) = \frac{1}{x} \lim_{t \to 0} \ln(1 + t)^{1/t}.$$

Отсюда с учетом второго замечательного предела (см. § 14.5) и непрерывности логарифмической функции получаем

$$y' = \frac{1}{x} \lim_{t \to 0} \ln(1 + t)^{1/t} = \frac{1}{x} \ln \mathrm{e} = \frac{1}{x},$$

т.е.

$$(\ln x)' = \frac{1}{x}.$$

Пусть теперь $\boldsymbol{y = \log_a x}$. Очевидно, $\log_a x = \dfrac{\ln x}{\ln a}$. Получаем

$$y' = (\log_a x)' = \left(\frac{\ln x}{\ln a}\right)' = \frac{1}{\ln a} (\ln x)' = \frac{1}{x \ln a},$$

т.е.

$$(\log_a x)' = \frac{1}{x \ln a}.$$

Производная показательной функции

Пусть $\boldsymbol{y = a^x}$. Логарифмируя это равенство, получаем

$$\ln y = x \ln a. \qquad (*)$$

В соответствии с правилом дифференцирования сложной функции

$$(\ln y)' = \frac{1}{y} y' = \frac{y'}{y}.$$

Дифференцируя равенство (*), получаем

$$\frac{y'}{y} = \ln a, \quad y' = y \ln a,$$

или

$$y' = a^x \ln a,$$

т.е.

$$(a^x)' = a^x \ln a.$$

При $a = \mathrm{e}$, очевидно,

$$(\mathrm{e}^x)' = \mathrm{e}^x.$$

Производная степенной функции

Пусть $\boldsymbol{y = x^n}$, где n – любое действительное число. Логарифмируем эту функцию:

$$\ln y = n \ln x.$$

Дифференцируя обе части равенства, получаем

$$\frac{y'}{y} = n\frac{1}{x}, \quad y' = \frac{ny}{x} = \frac{nx^n}{x} = nx^{n-1},$$

т.е.

$$(x^n)' = nx^{n-1}.$$

Производные тригонометрических функций

Выведем теперь формулы для производных тригонометрических функций:

1. $\boldsymbol{y = \sin x}$. Производную этой функции мы уже нашли (см. пример 16.3):

$$(\sin x)' = \cos x.$$

2. $\boldsymbol{y = \cos x}$:

$$\Delta y = \cos(x + \Delta x) - \cos x = -2\sin\frac{x + \Delta x + x}{2}\sin\frac{x + \Delta x - x}{2} =$$

$$= -2\sin\frac{\Delta x}{2}\sin\left(x + \frac{\Delta x}{2}\right),$$

$$\frac{\Delta y}{\Delta x} = -\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\sin\left(x + \frac{\Delta x}{2}\right).$$

Переходя к пределу при $\Delta x \to 0$ и учитывая, что $\sin x$ есть непрерывная функция, получаем

$$y' = -\sin x,$$

т.е.

$$(\cos x)' = -\sin x.$$

3. $\boldsymbol{y = \mathrm{tg}\, x}$:

$$y' = \left(\frac{\sin x}{\cos x}\right)' = \frac{(\sin x)'\cos x - \sin x(\cos x)'}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x},$$

т.е.

$$(\mathrm{tg}\, x)' = \frac{1}{\cos^2 x}.$$

4. $\boldsymbol{y = \mathrm{ctg}\, x}$. Действуя аналогично, получаем

$$(\mathrm{ctg}\, x)' = -\frac{1}{\sin^2 x}.$$

Производные обратных тригонометрических функций

И наконец, выведем формулы для производных обратных тригонометрических функций:

1. $\boldsymbol{y = \arcsin x}$. Эта функция является обратной для функции $x = \sin y$. По теореме о производной обратной функции (см. § 16.4)

$$y'_x = \frac{1}{x'_y} = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - \sin^2 y}}$$

(корень взят со знаком «плюс», так как $\cos y > 0$ при $-\dfrac{\pi}{2} < y < \dfrac{\pi}{2}$). Так как $\sin y = x$, то окончательно получаем

$$(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}.$$

2. $\boldsymbol{y = \arccos x}$:

$$(\arccos x)' = -\frac{1}{\sqrt{1 - x^2}}.$$

Вывод аналогичен предыдущему.

3. $\boldsymbol{y = \operatorname{arctg} x}$. Эта функция является обратной для $x = \operatorname{tg} y$. Так как $x'_y = \dfrac{1}{\cos^2 y}$, то

$$y'_x = \frac{1}{x'_y} = \cos^2 y = \frac{\cos^2 y}{\cos^2 y + \sin^2 y} = \frac{1}{1 + \operatorname{tg}^2 y} = \frac{1}{1 + x^2},$$

т.е.

$$(\operatorname{arctg} x)' = \frac{1}{1 + x^2}.$$

4. $\boldsymbol{y = \operatorname{arcctg} x}$:

$$(\operatorname{arcctg} x)' = -\frac{1}{1 + x^2}.$$

Вывод аналогичен предыдущему.

Мы вывели формулы производных всех основных элементарных функций. Сведем их теперь в таблицу и напомним еще раз правила дифференцирования.

Таблица производных

1. $c' = 0$.

2. $(x^n)' = nx^{n-1}$ (n — любое действительное число).

3. $(\log_a x)' = \dfrac{1}{x \ln a}$; $(\ln x)' = \dfrac{1}{x}$.

4. $(a^x)' = a^x \ln a$; $(\mathrm{e}^x)' = \mathrm{e}^x$.

5. $(\sin x)' = \cos x$.

6. $(\cos x)' = -\sin x$.

7. $(\operatorname{tg} x)' = \dfrac{1}{\cos^2 x}$.

8. $(\operatorname{ctg} x)' = -\dfrac{1}{\sin^2 x}$.

9. $(\arcsin x)' = \dfrac{1}{\sqrt{1 - x^2}}$.

10. $(\arccos x)' = -\dfrac{1}{\sqrt{1 - x^2}}$.

11. $(\operatorname{arctg} x)' = \dfrac{1}{1 + x^2}$.

12. $(\operatorname{arcctg} x)' = -\dfrac{1}{1 + x^2}$.

Правила дифференцирования

I. $(cu)' = cu'$.

II. $(u \pm v)' = u' \pm v'$.

III. $(uv)' = u'v + uv'$.

IV. $\left(\dfrac{u}{v}\right)' = \dfrac{u'v - uv'}{v^2}$.

V. Если $y = f(u)$, $u = \varphi(x)$, то $y'_x = y'_u u'_x$.

Формулы 1–12 и правила I–V составляют основу для практического дифференцирования.

16.6. ДИФФЕРЕНЦИАЛ

Напомним, что функция $y = f(x)$ называется *дифференцируемой* в точке x_0, если ее приращение Δy можно представить в виде (16.2):

$$\Delta y = A\Delta x + \alpha\Delta x,$$

где $\alpha \to 0$ при $\Delta x \to 0$.

Здесь величина $A\Delta x$, если $A \neq 0$, является главным членом разложения Δy, т.е. главной линейной относительно Δx частью приращения функции.

Определение. **Дифференциалом** dy **функции** $y = f(x)$ в точке x_0 называется главная линейная относительно Δx часть приращения функции в этой точке:

$$dy = A\Delta x. \tag{*}$$

При $A = 0$ дифференциал также определяют формулой (*), т.е. в этом случае $dy = 0$.

Так как из теоремы 16.1 следует $A = f'(x_0)$, то

$$dy = f'(x_0)\Delta x. \tag{16.6}$$

Дифференциалом dx независимой переменной x будем называть ее приращение Δx, и равенство (16.6) можем записать в виде

$$dy = f'(x_0)dx, \tag{16.6$'$}$$

откуда $f'(x_0) = \dfrac{dy}{dx}$. Теперь мы видим, что $\dfrac{dy}{dx}$ не просто символическое обозначение производной, а обычное отношение дифференциала функции dy к дифференциалу ее аргумента dx. Теперь формулу (16.2) с учетом (16.6) можно переписать в виде

$$\Delta y = f'(x_0)\Delta x + o(\Delta x)$$

или

$$\Delta f(x_0) = f'(x_0)\Delta x + o(\Delta x). \tag{16.2$'$}$$

Геометрический смысл дифференциала

Пусть точка M на графике функции $y = f(x)$ соответствует значению аргумента $x = x_0$, а точка N – значению аргумента $x = x_0 + \Delta x$ (рис. 16.2). Тогда $MA = \Delta x$, $AN = \Delta y$. Проведем касательную к кривой $y = f(x)$ в точке M. Пусть α – угол между этой касательной и осью Ox. Мы знаем, что $\operatorname{tg}\alpha = f'(x_0)$.

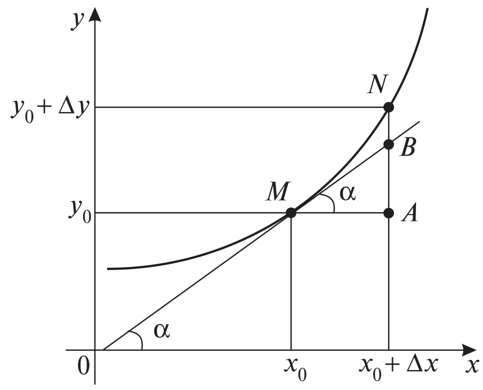


Рис. 16.2. Геометрический смысл дифференциала

Рассмотрим прямоугольный треугольник MAB. Очевидно, $MA = \Delta x$, $AB = MA \cdot \operatorname{tg}\alpha = \Delta x \operatorname{tg}\alpha = f'(x_0)\Delta x = dy$.

Итак, в то время как Δy есть приращение ординаты кривой, dy является соответствующим *приращением ординаты касательной*.

Применение дифференциала в приближенных вычислениях

Приближенные вычисления с помощью дифференциала основаны на приближенной замене приращения функции на ее дифференциал. Так как дифференциал есть главная часть приращения функции, то

$$\Delta y \approx dy,$$

или, подробнее,

$$\Delta f(x_0) = f(x_0 + \Delta x) - f(x_0) \approx f'(x_0)\Delta x.$$

Отсюда

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x.$$

Пример **16.5.** Вычислить приближенно $\sqrt[3]{8{,}24}$.

Решение. Пусть $f(x) = \sqrt[3]{x}$, $x_0 = 8$, тогда $f(x_0) = 2$, $\Delta x = 0{,}24$. Воспользуемся тем, что $\Delta y \approx dy = f'(x_0)\Delta x$: $f'(x) = (x^{1/3})' = \dfrac{1}{3}x^{-2/3} =$

$= \dfrac{1}{3\sqrt[3]{x^2}}$, $\quad f'(x_0) = \dfrac{1}{3\sqrt[3]{8^2}} = \dfrac{1}{12}$, $\quad \Delta y \approx dy = \dfrac{1}{12} \cdot 0{,}24 = 0{,}02$. Откуда

$$f(x_0 + \Delta x) = \sqrt[3]{8{,}24} \approx 2 + 0{,}02 = 2{,}02.$$

Задача нахождения дифференциала функции, очевидно, сводится к нахождению производной и умножению ее на дифференциал аргумента. Поэтому большинство теорем и формул, относящихся к производным, остаются верными и для дифференциалов. В частности:

I. $d(cu) = cdu \quad (c = \text{const}).$ **III.** $d(uv) = vdu + udv.$

II. $d(u \pm v) = du \pm dv.$ **IV.** $d\left(\dfrac{u}{v}\right) = \dfrac{vdu - udv}{v^2}.$

Найдем выражение для дифференциала сложной функции. Пусть $y = f(u)$, $u = \varphi(x)$, или $y = f(\varphi(x))$. Если $y = f(u)$ и $u = \varphi(x)$ – дифференцируемые функции своих аргументов, то производная сложной функции равна, как известно, $y' = f'(u)u'$.

Тогда дифференциал функции

$$dy = f'(x)dx = f'(u)u'dx = f'(u)du,$$

так как $u'dx = du$. Таким образом, одновременно

$$dy = f'(x)dx \quad \text{и} \quad dy = f'(u)du,$$

т.е. форма дифференциала не зависит от того, чем является аргумент функции – независимой переменной или функцией другого аргумента. Это свойство дифференциала получило название **инвариантности** (т.е. неизменности) **формы дифференциала**.

Производные и дифференциалы высших порядков

Если функция $f(x)$, определенная на X, имеет производную $f'(x)$ во всех точках X, то эта производная сама является функцией аргумента x: $f'(x) = g(x)$, и можно ставить вопрос о ее производной $g'(x)$. Производная от первой производной функции $f(x)$, т.е. $(f'(x))'$, называется **второй производной**, или **производной второго порядка**, и обозначается $f''(x)$ или y''. Итак,

$$f''(x) = (f'(x))', \quad \text{или} \quad y'' = (y')'.$$

Аналогично определяется **производная третьего порядка**:

$$f'''(x) = (f''(x))', \quad \text{или} \quad y''' = (y'')' \quad \text{и т.д.}$$

Производная *n*-го порядка обозначается $f^{(n)}(x)$ (или $y^{(n)}$) и определяется в соответствии с описанной схемой так:

$$f^{(n)}(x) = (f^{(n-1)}(x))',$$

где $f^{(n-1)}(x)$ – производная $(n-1)$-го порядка.

Пример **16.6:**
1) $y = e^{kx}$, $y' = ke^{kx}$, $y'' = k^2e^{kx}$, ..., $y^{(n)} = k^ne^{kx}$;
2) $y = \sin x$, $y' = \cos x$, $y'' = -\sin x$, $y''' = -\cos x$, $y^{(4)} = \sin x$, ...

Нетрудно показать, что $(\sin x)^{(n)} = \sin\left(x + \frac{\pi}{2}n\right)$.

Дифференциалом второго порядка (или **вторым дифференциалом**) функции $y = f(x)$ называется дифференциал от дифференциала этой функции, т.е. $d(dy)$, и обозначается d^2y:

$$d^2y = d(dy). \tag{16.7}$$

Очевидно, $d^2y = d(dy) = d(f'(x)dx) = (f'(x)dx)'dx = f''(x)dx^2$, где $dx^2 = (dx)^2$. Здесь мы при дифференцировании считаем dx постоянной, так как $dx = \Delta x$ не зависит от x. Итак,

$$d^2y = f''(x)dx^2. \tag{16.8}$$

Аналогично определяется **третий дифференциал** $d^3y = d(d^2y)$ и т.д.; наконец, **дифференциал *n*-го порядка (*n*-й дифференциал)** – это дифференциал от дифференциала $(n-1)$-го порядка:

$$d^ny = d(d^{n-1}y). \tag{16.9}$$

Выражение для d^ny находим аналогично тому, как это сделано для d^2y:

$$d^ny = f^{(n)}(x)dx^n. \tag{16.10}$$

Отсюда

$$f^{(n)}(x) = \frac{d^ny}{dx^n}.$$

Следует заметить, что дифференциалы второго и более высоких порядков *не обладают* свойством инвариантности формы в отличие от дифференциала первого порядка. Убедимся в этом.

Пусть $y = f(u)$, $u = \varphi(x)$, или $y = f(\varphi(x))$, тогда

$$d^2y = d^2f(\varphi(x)) = d(df(\varphi(x))).$$

По условию $df(\varphi(x)) = f'(u)du$, где $u = \varphi(x)$. Отсюда

$$d(f'(u)du) = d(f'(u))du + f'(u)d(du). \tag{**}$$

Если обозначить $f'(u) = g(u)$, то

$$d(f'(u)) = dg(u) = g'(u)du = (f'(u))'du = f''(u)du.$$

Кроме того, $d(du) = d^2u = d^2\varphi(x)$.

Таким образом, из (**) получаем

$$d^2f(u) = f''(u)(du)^2 + f'(u)d^2u, \quad u = \varphi(x). \tag{16.11}$$

Очевидно, что второй дифференциал сложной функции $f(\varphi(x))$ существует, если функции $f(u)$ и $\varphi(x)$ имеют конечные производные до второго порядка включительно.

Из формулы (16.11) следует, что второй дифференциал сложной функции не обладает инвариантностью формы:

- если $y = f(x)$, где x — независимый аргумент, то

$$dy = f''(x)dx^2;$$

- если $y = f(u)$, где u — зависимый аргумент, $u = \varphi(x)$, то

$$dy = f''(u)du^2 + f'(u)d^2u.$$

Вопросы

1. В чем заключается геометрический смысл производной?
2. Пусть для функции $y = f(x)$ производная $f'(3) = \sqrt{3}$. Под каким углом к оси Ox расположена касательная к графику функции при $x = 3$?
3. Что такое предельная производительность? Как связано это понятие с понятием производной?
4. В чем заключается «золотое» правило экономики?
5. Что такое предельная себестоимость продукции?
6. Как определяется понятие дифференцируемости функции $y = f(x)$ в точке x_0?
7. Почему функцию $f(x)$, имеющую производную в точке x_0, называют дифференцируемой в этой точке?
8. Будет ли функция, дифференцируемая в точке $x = 2$, непрерывной в этой точке?
9. Пусть функция $y = f(x)$ непрерывна в точке $x = 5$. Можно ли утверждать, что эта функция имеет производную в указанной точке?
10. Будет ли прямая $y = 4x - 4$ касательной к параболе $y = x^2$? А прямая $y = -4x - 4$?
11. Равносильны ли утверждения: «функция $y = f(x)$ дифференцируема в точке x_0» и «функция $y = f(x)$ имеет в точке x_0 конечную производную»?

12. Каков алгоритм нахождения производной произвольной функции?
13. Можно ли утверждать, что любая из основных элементарных функций дифференцируема в каждой точке, в которой она определена?
14. В соответствии с каким правилом дифференцируется сложная функция? Приведите примеры, отличные от указанных в книге.
15. Каков геометрический смысл дифференциала?
16. Может ли дифференциал функции $f(x)$ быть больше приращения этой функции?
17. Производная y' нередко обозначается в виде $\frac{dy}{dx}$. Каков смысл этого обозначения?
18. На чем основано применение дифференциала в приближенных вычислениях?
19. В чем заключается инвариантность формы дифференциала сложной функции?

Глава 17
СВОЙСТВА ДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЙ

17.1. ОСНОВНЫЕ ТЕОРЕМЫ ДИФФЕРЕНЦИАЛЬНОГО ИСЧИСЛЕНИЯ

***Теорема* 17.1 (теорема Ферма).** Пусть функция $y = f(x)$ определена на интервале (a, b) и имеет наибольшее (наименьшее) значение в некоторой точке $x_0 \in (a, b)$. Тогда если в этой точке существует конечная производная $f'(x_0)$, то $f'(x_0) = 0$.

Доказательство. Докажем теорему для случая, когда функция имеет в точке x_0 наибольшее значение (для наименьшего значения доказательство аналогично). В этом случае для всех $x \in (a, b)$ выполняется неравенство $f(x) \le f(x_0)$, что означает $\Delta y = f(x_0 + \Delta x) - f(x_0) \le 0$ для любой точки $x = (x_0 + \Delta x) \in (a, b)$. Если $\Delta x > 0$, то $\dfrac{\Delta y}{\Delta x} \le 0$, следовательно,

$$\lim_{\Delta x \to 0+} \frac{\Delta y}{\Delta x} \le 0; \tag{*}$$

если же $\Delta x < 0$, то $\dfrac{\Delta y}{\Delta x} \ge 0$, следовательно,

$$\lim_{\Delta x \to 0-} \frac{\Delta y}{\Delta x} \ge 0. \tag{**}$$

По определению производной

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x},$$

причем этот предел не зависит от того, стремится Δx к нулю, будучи положительным или будучи отрицательным. Но односторонние пределы (*) и (**) совпадают только в случае, когда они равны нулю:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0+} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0-} \frac{\Delta y}{\Delta x} = 0,$$

откуда следует, что $f'(x_0) = 0$. Теорема доказана.

Геометрический смысл теоремы Ферма очевиден: если в точке x_0 дифференцируемая функция принимает наибольшее (наименьшее) значение, то в точке $M(x_0, f(x_0))$ касательная к графику этой функции параллельна оси Ox (рис. 17.1).

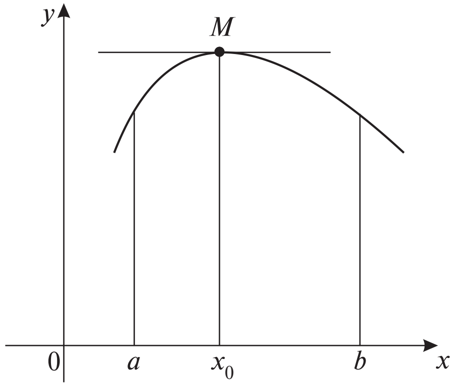


Рис. 17.1. Геометрический смысл теоремы Ферма

Теорема **17.2 (теорема Ролля).** Пусть функция $f(x)$ удовлетворяет следующим трем условиям:

1) непрерывна на отрезке $[a, b]$;
2) дифференцируема на интервале (a, b);
3) на концах отрезка принимает равные значения: $f(a) = f(b)$.

Тогда внутри отрезка существует хотя бы одна точка $\xi \in (a, b)$, в которой производная равна нулю:

$$f'(\xi) = 0.$$

Доказательство. Так как $f(x)$ непрерывна на $[a, b]$, то в силу второй теоремы Вейерштрасса (см. § 15.2) она достигает на $[a, b]$ своего наибольшего значения M и своего наименьшего значения m.

Возможны два случая:

1. $M = m$. Тогда $f(x) = M = m = \text{const}$; поэтому $f'(x) = 0$ во всех точках, так что в качестве ξ можно взять любую точку из (a, b).

2. $M \neq m$. Оба эти значения не могут достигаться на концах отрезка (так как $f(a) = f(b)$). Следовательно, хотя бы одно из этих значений достигается в некоторой внутренней точке $\xi \in (a, b)$, и в силу теоремы Ферма $f'(\xi) = 0$. Теорема доказана.

Геометрический смысл теоремы Ролля заключается в следующем: если крайние ординаты кривой $y = f(x)$ равны, то на кривой найдется точка, где касательная параллельна оси Ox (рис. 17.2).

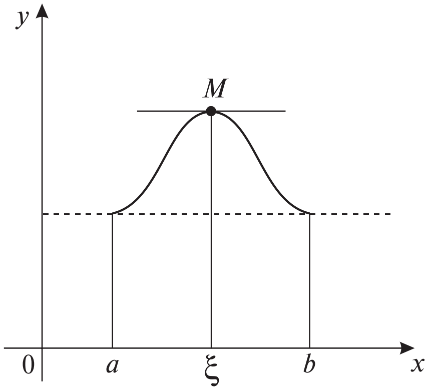


Рис. 17.2. Геометрический смысл теоремы Ролля

Следует заметить, что все условия теоремы Ролля существенны, и при невыполнении хотя бы одного из них заключение теоремы может оказаться неверным.

***Теорема* 17.3 (теорема Лагранжа).** Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема на интервале (a, b). Тогда существует такая точка $\xi \in (a, b)$, что для нее выполняется равенство

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

Доказательство. Рассмотрим вспомогательную функцию

$$F(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Функция $F(x)$ удовлетворяет всем условиям теоремы Ролля: она непрерывна на $[a, b]$ (так как $f(x)$ непрерывна), дифференцируема на (a, b):

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

а также принимает на концах отрезка $[a, b]$ одинаковые значения:

$$F(a) = F(b) = 0.$$

Согласно теореме Ролля существует такая точка $\xi \in (a, b)$, что $F'(\xi) = 0$, т.е.

$$f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0.$$

Отсюда

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

Теорема доказана.

Заметим, что из теоремы Лагранжа вытекает равенство

$$f(b) - f(a) = f'(\xi)(b - a), \tag{17.1}$$

называемое **формулой Лагранжа**.

Теорема Ролля является частным случаем теоремы Лагранжа.

Геометрический смысл теоремы Лагранжа виден из рис. 17.3. Хорда, проходящая через точки $M_1(a, f(a))$ и $M_2(b, f(b))$, имеет угловой коэффициент, равный

$$\operatorname{tg}\alpha = \frac{M_2N}{M_1N} = \frac{f(b) - f(a)}{b - a}.$$

Теорема Лагранжа утверждает, что существует такая точка $M(\xi, f(\xi))$ в интервале (a, b), где касательная к графику функции параллельна хорде M_1M_2: ее угловой коэффициент $f'(\xi)$ равен угловому коэффициенту хорды M_1M_2.

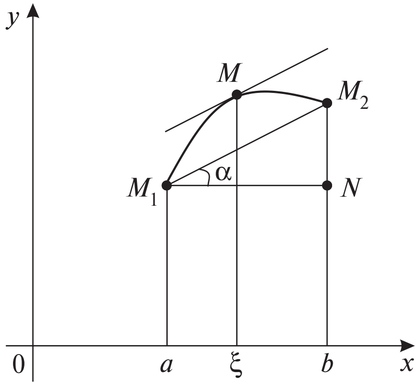


Рис. 17.3. Геометрический смысл теоремы Лагранжа

Теорема **17.4 (теорема Коши).** Пусть функции $f(x)$ и $g(x)$ непрерывны на $[a, b]$ и дифференцируемы в (a, b), причем

$g'(x) \neq 0$. Тогда существует такая точка $\xi \in (a, b)$, что справедлива формула

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}. \qquad (17.2)$$

Д о к а з а т е л ь с т в о. Прежде всего убедимся в том, что знаменатель в левой части равенства (17.2) не равен нулю. Действительно, если бы было $g(b) - g(a) = 0$, т.е. $g(b) = g(a)$, то по теореме Ролля $g'(\xi) = 0$ в некоторой точке $\xi \in (a, b)$, а это противоречит условию доказываемой теоремы.

Рассмотрим вспомогательную функцию

$$F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)}[g(x)-g(a)].$$

Нетрудно убедиться в том, что эта функция на $[a, b]$ удовлетворяет всем условиям теоремы Ролля: она непрерывна на $[a, b]$ (в силу непрерывности $f(x)$ и $g(x)$), дифференцируема в (a, b) – ее производная имеет вид

$$F'(x) = f'(x) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(x)$$

и $F(a) = F(b) = 0$. Стало быть, существует такая точка $\xi \in (a, b)$, что $F'(\xi) = 0$, т.е.

$$f'(\xi) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(\xi) = 0.$$

Отсюда (с учетом того, что $g'(\xi) \neq 0$) получаем формулу

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Теорема доказана.

Формула (17.2) называется **формулой Коши**.

Теорема Лагранжа есть частный случай теоремы Коши при $g(x) = x$.

17.2. ПРАВИЛО ЛОПИТАЛЯ

Неопределенность вида $\frac{0}{0}$

Рассмотрим отношение двух функций $\frac{f(x)}{g(x)}$. Будем говорить, что это отношение при $x \to a$ есть неопределенность вида $\frac{0}{0}$, если

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0.$$

Если существует $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}$, то вычисление этого предела называют *раскрытием* упомянутой неопределенности.

***Теорема* 17.5.** Пусть функции $f(x)$ и $g(x)$ на отрезке $[a, b]$ удовлетворяют условиям теоремы Коши и пусть $f(a) = g(a) = 0$. Тогда если существует $\lim\limits_{x\to a}\dfrac{f'(x)}{g'(x)}$, то существует и $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}$, причем

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}. \tag{17.3}$$

Доказательство. Возьмем $[a, x]\subset[a, b]$. Применим теорему Коши к функциям $f(x)$ и $g(x)$ на отрезке $[a, x]$; получаем

$$\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f'(\xi)}{g'(\xi)},$$

где ξ – некоторая точка, расположенная между a и x. Но по условию теоремы $f(a)=g(a)=0$, поэтому

$$\frac{f(x)}{g(x)}=\frac{f'(\xi)}{g'(\xi)}.$$

Пусть $x\to a$. Тогда $\xi\to a$ (так как $a<\xi<x$). Если при этом существует $\lim\limits_{x\to a}\dfrac{f'(x)}{g'(x)}=K$, то $\lim\limits_{x\to a}\dfrac{f'(\xi)}{g'(\xi)}$ также существует и равен K. Поэтому

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(\xi)}{g'(\xi)}=\lim_{\xi\to a}\frac{f'(\xi)}{g'(\xi)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}=K.$$

Итак,

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)},$$

что и требовалось доказать.

Теорему 17.5 обычно называют **правилом Лопиталя**.

***Замечание* 1.** Теорема 17.5 остается справедливой и тогда, когда $f(x)$ и $g(x)$ *не определены* при $x = a$, но $\lim\limits_{x\to a}f(x)=\lim\limits_{x\to a}g(x)=0$. В таком случае достаточно *доопределить* функции $f(x)$ и $g(x)$ в точке a так, чтобы они стали непрерывными в этой точке, положив $f(a)=\lim\limits_{x\to a}f(x)=0$, $g(a)=\lim\limits_{x\to a}g(x)=0$, после чего все условия теоремы выполнены.

Замечание **2.** Правило Лопиталя можно применять повторно, если $f'(x)$ и $g'(x)$ удовлетворяют тем же требованиям, что и исходные функции $f(x)$ и $g(x)$.

Пример **17.1.** Найти пределы:

а) $\lim\limits_{x\to 0}\dfrac{\sin 3x}{4x}$; б) $\lim\limits_{x\to 0}\dfrac{e^{3x}-1}{x}$; в) $\lim\limits_{x\to 0}\dfrac{e^{x}+e^{-x}-2}{1-\cos x}$.

Решение:

а) $\lim\limits_{x\to 0}\dfrac{\sin 3x}{4x}=\lim\limits_{x\to 0}\dfrac{(\sin 3x)'}{(4x)'}=\lim\limits_{x\to 0}\dfrac{3\cos 3x}{4}=\dfrac{3}{4}$;

б) $\lim\limits_{x\to 0}\dfrac{e^{3x}-1}{x}=\lim\limits_{x\to 0}\dfrac{3e^{3x}}{1}=3$;

в) $\lim\limits_{x\to 0}\dfrac{e^{x}+e^{-x}-2}{1-\cos x}=\lim\limits_{x\to 0}\dfrac{e^{x}-e^{-x}}{\sin x}=\lim\limits_{x\to 0}\dfrac{e^{x}+e^{-x}}{\cos x}=\dfrac{2}{1}=2.$

Замечание **3.** Правило Лопиталя применимо и в том случае, если

$$\lim_{x\to\infty} f(x)=0 \quad \text{и} \quad \lim_{x\to\infty} g(x)=0.$$

Чтобы убедиться в этом, положим $x=\dfrac{1}{z}$. Тогда $z\to 0$ при $x\to\infty$, следовательно,

$$\lim_{z\to 0} f\left(\frac{1}{z}\right)=0,\quad \lim_{z\to 0} g\left(\frac{1}{z}\right)=0.$$

К функциям $f\left(\dfrac{1}{z}\right)$ и $g\left(\dfrac{1}{z}\right)$ от новой переменной z можно применить теорему 17.5. Получаем

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=\lim_{z\to 0}\frac{f\left(\frac{1}{z}\right)}{g\left(\frac{1}{z}\right)}=\lim_{z\to 0}\frac{f'\left(\frac{1}{z}\right)\cdot\left(-\frac{1}{z^2}\right)}{g'\left(\frac{1}{z}\right)\cdot\left(-\frac{1}{z^2}\right)}=\lim_{z\to 0}\frac{f'\left(\frac{1}{z}\right)}{g'\left(\frac{1}{z}\right)}=\lim_{x\to\infty}\frac{f'(x)}{g'(x)},$$

т.е.

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=\lim_{x\to\infty}\frac{f'(x)}{g'(x)}. \qquad (17.4)$$

Неопределенность вида $\dfrac{\infty}{\infty}$

В случае когда функции $f(x)$ и $g(x)$ *стремятся к бесконечности* при $x \to a$, также применимо правило Лопиталя.

Пусть $\lim\limits_{x\to a} f(x) = \infty$, $\lim\limits_{x\to a} g(x) = \infty$ и пусть отношение $\dfrac{f'(x)}{g'(x)}$ имеет предел: $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} = K$. Тогда отношение $\dfrac{f(x)}{g(x)}$ также имеет предел при $x \to a$ и справедливо равенство (17.3):

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

(Примем это утверждение без доказательства.)

Заметим, что для функций $f(x)$ и $g(x)$, стремящихся к бесконечности при $x \to \infty$, правило Лопиталя также справедливо:

$$\lim_{x\to \infty} \frac{f(x)}{g(x)} = \lim_{x\to \infty} \frac{f'(x)}{g'(x)}.$$

17.3. ФОРМУЛА ТЕЙЛОРА

Пусть функция $y = f(x)$ имеет в точке $x = x_0$ и в некоторой ее окрестности производные до $(n + 1)$-го порядка включительно.

Назовем **многочленом Тейлора** порядка n функции $y = f(x)$ следующее выражение:

$$P_n(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots +$$
$$+ \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \qquad (17.5)$$

Этот многочлен и его производные до n-й включительно в точке $x = x_0$ имеют те же значения, что и функция $f(x)$, и ее производные соответственно:

$$\begin{aligned} &f(x_0) = P_n(x_0), \quad f'(x_0) = P_n'(x_0), \\ &f''(x_0) = P_n''(x_0), \ \ldots, \ \ f^{(n)}(x_0) = P_n^{(n)}(x_0) \end{aligned} \qquad (17.6)$$

(эти равенства легко проверяются непосредственно). Поэтому имеются основания считать, что многочлен (17.5) дает приближенное представление функции $f(x)$. Степень этого приближения, очевидно, измеряется разностью $R_n(x) = f(x) - P_n(x)$. Имеем

$$f(x) = P_n(x) + R_n(x)$$

или в развернутом виде:

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\ldots+$$
$$+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+R_n(x). \qquad (17.7)$$

Формула (17.7) называется **формулой Тейлора**, а $R_n(x)$ – **остаточным членом**.

Наша дальнейшая задача – оценить, насколько функция $f(x)$ отличается от многочлена $P_n(x)$ при различных значениях x, т.е. оценить величину $R_n(x)$.

Запишем остаточный член в форме

$$R_n(x)=\frac{Q(x)}{(n+1)!}(x-x_0)^{n+1}, \qquad (17.8)$$

где $Q(x)$ – некоторая функция, которую предстоит найти.

Формула (17.7) с учетом (17.8) выглядит теперь так:

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\ldots+$$
$$+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\frac{Q(x)}{(n+1)!}(x-x_0)^{n+1}. \qquad (17.9)$$

Зафиксируем x. Пусть для определенности $x > x_0$. При этом функция $Q(x)$ имеет фиксированное числовое значение; обозначаем его через Q.

Обозначим через t переменную величину, изменяющуюся от x_0 до x, и рассмотрим на отрезке $[x_0, x]$ вспомогательную функцию

$$F(t)=f(x)-f(t)-\frac{f'(t)}{1!}(x-t)-\frac{f''(t)}{2!}(x-t)^2-\ldots-$$
$$-\frac{f^{(n)}(t)}{n!}(x-t)^n-\frac{Q}{(n+1)!}(x-t)^{n+1}, \qquad (17.10)$$

где Q имеет числовое значение, определяемое соотношением (17.9) при упомянутом фиксированном x.

Найдем производную $F'(t)$:

$$F'(t)=-f'(t)-\frac{f''(t)}{1!}(x-t)+f'(t)-\frac{f'''(t)}{2!}(x-t)^2+$$
$$+\frac{2f''(t)}{2!}(x-t)-\ldots-\frac{f^{(n+1)}(t)}{n!}(x-t)^n+\frac{nf^{(n)}(t)}{n!}(x-t)^{n-1}+$$
$$+\frac{(n+1)Q}{(n+1)!}(x-t)^n.$$

Здесь соответствующие слагаемые с противоположными знаками взаимно уничтожаются. Получаем

$$F'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{Q}{n!}(x-t)^n. \tag{17.11}$$

Итак, функция $F(t)$ имеет производную (17.11) на отрезке $[x_0, x]$. Кроме того, из (17.10) следует, что $F(x) = F(x_0) = 0$. Поэтому к функции $F(t)$ на $[x_0, x]$ применима теорема Ролля; следовательно, существует такое $\xi \in (x_0, x)$, что $F'(\xi) = 0$. Отсюда на основании равенства (17.11)

$$-\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n + \frac{Q}{n!}(x-\xi)^n = 0,$$

следовательно,

$$Q = f^{(n+1)}(\xi).$$

Подставляя это выражение в (17.8), получаем

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}. \tag{17.12}$$

Выражение (17.12) представляет собой **остаточный член в форме Лагранжа**.

Подставим теперь $R_n(x)$ в (17.7):

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots +$$
$$+ \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}. \tag{17.13}$$

Формула (17.13) называется **формулой Тейлора с остаточным членом в форме Лагранжа**.

Формула (17.13) применяется в тех случаях, когда при определенных значениях x, отличных от x_0, мы желаем заменить приближенно функцию $f(x)$ многочленом $P_n(x)$ и численно оценить погрешность, возникающую при такой замене. Однако бывают случаи, когда нас не интересуют определенные значения x, но важно знать поведение остаточного члена при стремлении x к x_0, точнее говоря, нас интересует *порядок малости остаточного члена*. Для этого остаточный член применяется в иной форме.

Д о к а ж е м, что остаточный член $R_n(x)$ при $x \to x_0$ есть бесконечно малая более высокого порядка, чем $(x-x_0)^n$:

$$R_n(x) = o((x-x_0)^n). \tag{17.14}$$

Это **остаточный член в форме Пеано**.

Пусть в некоторой окрестности точки x_0 существуют производные функции $f(x)$ до n-го порядка включительно, и $f^{(n)}(x)$ непрерывна в точке x_0.

Заменим в формуле (17.13) n на $n-1$:

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots +$$
$$+\frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1} + \frac{f^{(n)}(\xi)}{n!}(x-x_0)^n, \tag{17.15}$$

где ξ содержится между x_0 и x. Представим последний член в виде

$$\frac{f^{(n)}(\xi)}{n!} = \frac{f^{(n)}(x_0)}{n!} + \alpha(x). \tag{17.16}$$

Если $x \to x_0$, то $\xi \to x_0$ (так как ξ содержится между x_0 и x). Тогда $f^{(n)}(\xi) \to f^{(n)}(x_0)$, так как $f^{(n)}(x)$ непрерывна. В силу (17.16) $\alpha(x) \to 0$. А это и означает, что $\alpha(x)(x-x_0)^n = o((x-x_0)^n)$.

Следовательно,

$$\frac{f^{(n)}(\xi)}{n!}(x-x_0)^n = \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + o((x-x_0)^n).$$

Теперь из (17.15) получаем

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots +$$
$$+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + o((x-x_0)^n). \tag{17.17}$$

Здесь

$$R_n(x) = o((x-x_0)^n),$$

что и требовалось доказать.

Положим $x_0 = 0$ в формуле (17.13), получим **формулу Маклорена** (являющуюся частным случаем формулы Тейлора):

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots +$$
$$+\frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}, \tag{17.18}$$

где ξ – точка, расположенная между 0 и x.

Формула Маклорена с остаточным членом в форме Пеано выглядит так:

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+o(x^n). \qquad (17.19)$$

Разложение по формуле Маклорена некоторых элементарных функций

Наиболее простыми элементарными функциями, как известно, являются многочлены. Формулы Тейлора и Маклорена дают возможность представить функцию $f(x)$ в виде многочлена, коэффициенты которого вычисляются сравнительно просто. Эти разложения применяются и для приближенного вычисления функций.

В частности, имеют место следующие приближенные равенства (при $x \to 0$):

$$e^x \approx 1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!};$$

$$\ln(1+x)\approx x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots+(-1)^{n+1}\frac{x^n}{n};$$

$$\sin x \approx x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots+\frac{(-1)^k x^{2k+1}}{(2k+1)!};$$

$$\cos x \approx 1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots+\frac{(-1)^k x^{2k}}{(2k)!};$$

$$(1+x)^{\alpha}\approx 1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\ldots+\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n.$$

Здесь в каждом случае погрешность является бесконечно малой более высокого порядка, чем x^n (в случае синуса $n=2k+1$, в случае косинуса $n=2k$).

Рассмотрим **разложение экспоненты и синуса** более подробно.

1. $\boldsymbol{f(x)=e^x}$. Очевидно, $f'(x)=e^x$, $f''(x)=e^x$, ..., $f^{(n)}(x)=e^x$; $f(0)=1$, $f'(0)=1$, ..., $f^{(n)}(0)=1$. Подставляя эти выражения в (17.19), получаем

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+o(x^n).$$

2. $f(x) = \sin x$. Последовательно дифференцируя и подставляя $x = 0$, получаем: $f(0) = 0$; $f'(x) = \cos x$, $f'(0) = 1$; $f''(x) = -\sin x$, $f''(0) = 0$; $f'''(x) = -\cos x$, $f'''(0) = -1$; ...; $f^{(n)}(x) = \sin\left(x + n\frac{\pi}{2}\right)$, $f^{(n)}(0) = \sin n\frac{\pi}{2}$.

Подставляя в (17.19), получаем

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o(x^{2k+1}).$$

Рассмотрим еще одну форму записи формулы Тейлора. В формуле (17.17) перенесем $f(x_0)$ в левую часть равенства и обозначим, как обычно, $x - x_0 = \Delta x$. Тогда разность $f(x) - f(x_0) = f(x_0 + \Delta x) - f(x_0) = \Delta f(x_0)$. Получаем

$$\Delta f(x_0) = f'(x_0)\Delta x + \frac{f''(x_0)}{2!}\Delta x^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}\Delta x^n + o(\Delta x^n). \qquad (17.17')$$

Полученная формула есть обобщение формулы (16.2′):

$$\Delta f(x_0) = f'(x_0)\Delta x + o(\Delta x),$$

которая, очевидно, получается из (17.17′), если положить $n = 1$. Аналогичным образом преобразуется формула (17.13):

$$\Delta f(x_0) = f'(x_0)\Delta x + \frac{f''(x_0)}{2!}\Delta x^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}\Delta x^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}\Delta x^{n+1}. \qquad (17.13')$$

Если заменить в формулах (17.13′) и (17.17′) приращение независимого аргумента Δx на dx (а мы знаем, что $dx = \Delta x$) и принять во внимание, что

$$f'(x_0)dx = df(x_0),\ \ f''(x_0)dx^2 = d^2 f(x_0),\ \ldots,$$
$$f^{(n)}(x_0)dx^n = d^n f(x_0),\ \ f^{(n+1)}(\xi)dx^{n+1} = d^{n+1} f(\xi),$$

то, подставляя эти выражения в (17.13′) и (17.17′), получаем соответственно

$$\Delta f(x_0) = df(x_0) + \frac{1}{2!}d^2 f(x_0) + \ldots + \frac{1}{n!}d^n f(x_0) + \frac{1}{(n+1)!}d^{n+1} f(\xi), \qquad (17.13'')$$

$$\Delta f(x_0) = df(x_0) + \frac{1}{2!}d^2 f(x_0) + \ldots + \frac{1}{n!}d^n f(x_0) + o(\Delta x^n). \qquad (17.17'')$$

Таким образом, при $\Delta x \to 0$ формулы (17.13″) и (17.17″) дают возможность из бесконечно малого приращения функции $\Delta f(x_0)$ выделить не только его главный член – первый дифференциал, но и члены более высоких порядков малости – ими оказываются последовательные дифференциалы второго, третьего и высших порядков с коэффициентами соответственно $\frac{1}{2!}, \frac{1}{3!}, \ldots, \frac{1}{n!}$.

Каждая из формул (17.13″) и (17.17″) называется **формулой Тейлора в дифференциальной форме**. При изучении функций нескольких переменных будем пользоваться именно таким представлением формулы Тейлора.

Вопросы

1. В чем заключается геометрический смысл теоремы Лагранжа?
2. Является ли теорема Лагранжа частным случаем теоремы Коши?
3. Пусть $\lim\limits_{x\to 2} f(x) = 1$, $\lim\limits_{x\to 2} g(x) = 0$. Можно ли применить правило Лопиталя для нахождения предела $\lim\limits_{x\to 2} \frac{f(x)}{g(x)}$?
4. Пусть $\lim\limits_{x\to 1} f(x) = +\infty$, $\lim\limits_{x\to 1} g(x) = -\infty$. Можно ли применить правило Лопиталя для нахождения предела $\lim\limits_{x\to 1} \frac{f(x)}{g(x)}$?
5. Что такое многочлен Тейлора? Каковы его свойства?
6. Как связан многочлен Тейлора функции $f(x)$ с формулой Тейлора для этой функции?
7. Что такое формула Маклорена?
8. Как выглядит формула Тейлора для функции $f(x)$ в дифференциальной форме?

Глава 18
ИССЛЕДОВАНИЕ ФУНКЦИЙ С ПОМОЩЬЮ ПЕРВОЙ ПРОИЗВОДНОЙ

18.1. ПРИЗНАК МОНОТОННОСТИ ФУНКЦИИ

***Теорема* 18.1.** Пусть функция $f(x)$ непрерывна на промежутке X и имеет внутри него конечную производную. Для того чтобы $f(x)$ была монотонно возрастающей (убывающей) на X, достаточно условие $f'(x) > 0$ ($f'(x) < 0$) внутри X.

Доказательство проведем для случая возрастания. Пусть $f'(x) > 0$; $x_1, x_2 \in X$, $x_2 > x_1$. Применим к $f(x)$ на $[x_1, x_2]$ теорему Лагранжа:

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1),$$

где $x_1 < c < x_2$. Так как $f'(c) > 0$, то $f(x_2) > f(x_1)$, следовательно, $f(x)$ – возрастающая функция.

Заметим, что доказанное условие не является необходимым. Функция может быть возрастающей (убывающей), например, в том случае, когда производная $f'(x)$ обращается в нуль в конечном числе внутренних точек промежутка X.

18.2. ЭКСТРЕМУМ ФУНКЦИИ

Определение. Точка x_0 называется **точкой локального максимума** функции $f(x)$, если в некоторой окрестности точки x_0 выполняется неравенство $f(x_0) > f(x)$.

Точка x_0 называется **точкой локального минимума** функции $f(x)$, если в некоторой окрестности точки x_0 выполняется неравенство $f(x_0) < f(x)$.

Если x_0 – точка локального максимума (минимума), то значение функции $f(x_0)$ называется **локальным максимумом (минимумом)**.

Общий термин для локального максимума и локального минимума – **локальный экстремум**.

Необходимое условие экстремума дифференцируемой функции следует из теоремы Ферма, доказанной в § 17.1: *для того чтобы дифференцируемая функция $f(x)$ имела в точке x_0 локальный экстремум, необходимо, чтобы в этой точке выполнялось равенство* $f'(x_0) = 0$.

Действительно, так как x_0 – точка экстремума, то существует интервал, содержащий точку x_0, на котором значение $f(x_0)$ является наибольшим или наименьшим. Тогда по теореме Ферма $f'(x_0) = 0$.

Заметим, что условие $f'(x_0) = 0$ не является достаточным условием экстремума. Так, функция $y = x^3$ возрастает на всей числовой прямой и не имеет экстремумов, но ее производная в точке $x_0 = 0$ равна нулю: $f'(x_0) = 3x_0^2 = 0$.

Кроме того, функция может иметь экстремум в некоторой точке, но не быть дифференцируемой в этой точке.

Точки, в которых производная функции обращается в нуль или не существует, называются **критическими** (или **стационарными**). Очевидно, если в какой-либо точке имеется экстремум*, то эта точка – критическая.

18.3. ПЕРВОЕ ДОСТАТОЧНОЕ УСЛОВИЕ ЭКСТРЕМУМА

***Теорема* 18.2.** Пусть функция $f(x)$ непрерывна в некотором интервале, содержащем критическую точку x_0, и дифференцируема во всех точках этого интервала, кроме, может быть, самой точки x_0. Если при переходе через точку x_0 производная меняет знак с «плюса» на «минус», то в точке x_0 имеется локальный максимум, а если с «минуса» на «плюс», то минимум.

Доказательство. Пусть для определенности производная меняет знак с «плюса» на «минус»: $f'(x) > 0$ при $x < x_0$, $f'(x) < 0$ при $x > x_0$ (для всех x, принадлежащих упомянутому интервалу). Применим теорему Лагранжа к $f(x)$ на отрезке $[x, x_0]$:

$$f(x_0) - f(x) = f'(c)(x_0 - x), \quad c \in (x, x_0).$$

Так как $f'(c) > 0$ и $x_0 - x > 0$, то $f(x) < f(x_0)$.

Применяя теперь теорему Лагранжа на отрезке $[x_0, x]$, т.е. при $x > x_0$, получаем

$$f(x) - f(x_0) = f'(c)(x - x_0), \quad c \in (x_0, x).$$

---

* Нередко мы говорим просто экстремум (максимум, минимум), имея в виду локальный экстремум (локальный максимум, локальный минимум).

Так как точка c находится теперь справа от x_0, то $f'(c) < 0$. Кроме того, $x - x_0 > 0$. Поэтому $f(x) - f(x_0) < 0$. Снова получаем $f(x) < f(x_0)$. Итак, для всех x из упомянутого интервала, содержащего точку x_0, выполняется неравенство

$$f(x_0) > f(x).$$

Следовательно, в точке x_0 имеется локальный максимум.

Случай локального минимума рассматривается аналогично.

На основании теорем 18.1 и 18.2 применяется следующая **схема исследования функции на экстремум с помощью первой производной**.

1. Найти производную $y' = f'(x)$.
2. Найти критические точки.
3. Исследовать знак производной слева и справа от каждой критической точки и сделать вывод о наличии локальных экстремумов функции.
4. Найти значения функции в точках локального экстремума.

Пример **18.1.** Исследовать функцию $f(x) = 3x^4 - 4x^3 - 12x^2 + 10$.

Решение. Находим производную:

$$f'(x) = 12x^3 - 12x^2 - 24x = 12x(x^2 - x - 2) = 12x(x + 1)(x - 2).$$

Решая уравнение $f'(x) = 0$, т.е. $x(x + 1)(x - 2) = 0$, находим критические точки: $x_1 = -1$, $x_2 = 0$, $x_3 = 2$. Исследовав знак производной (рис. 18.1), получаем: $x = -1$, $x = 2$ — точки локального минимума, $f(-1) = 5$, $f(2) = -22$ — минимальные значения функции; $x = 0$ — точка локального максимума, $f(0) = 10$ — максимальное значение функции в этой точке.

Рис. 18.1

18.4. НАИБОЛЬШЕЕ И НАИМЕНЬШЕЕ ЗНАЧЕНИЯ ФУНКЦИИ НА ОТРЕЗКЕ

Многие экономические задачи формулируются как задачи на нахождение наибольшего (наименьшего) значения функции на некотором множестве. Рассмотрим наиболее простой случай, когда требуется найти наибольшее (наименьшее) значение на отрезке $[a, b]$. В силу второй теоремы Вейерштрасса, если функция непрерывна на отрезке $[a, b]$, то она принимает на нем наибольшее и наименьшее значения. Отметим, что наибольшее или наименьшее значение функции может достигаться как в точках локального экстремума, так и на концах отрезка.

Обычно пользуются следующей **схемой для отыскания наибольшего и наименьшего значений функции на отрезке**.

1. Найти производную $f'(x)$.
2. Найти критические точки.
3. Найти значения функции в критических точках и на концах отрезка и выбрать из них наибольшее и наименьшее.

Заметим, что при этом нет необходимости исследовать на локальный экстремум функцию в критических точках.

Пример **18.2.** Найти наибольшее и наименьшее значения функции $f(x) = x^2 e^x$ на отрезке $[-3, 1]$.

Решение. 1. $f'(x) = 2xe^x + x^2 e^x = x(x+2)e^x$.

2. $f'(x) = 0$: $x(x+2)e^x = 0$. Критические точки: $x_1 = 0$, $x_2 = -2$.

3. $f(-3) = 9e^{-3}$, $f(-2) = 4e^{-2}$, $f(0) = 0$, $f(1) = e$;

$f_{\text{наиб}} = f(1) = e$, $f_{\text{наим}} = f(0) = 0$.

Итак, наибольшее значение достигается на правом конце отрезка, а наименьшее – в одной из критических точек.

Пример **18.3.** В пункте A находится месторождение сырья. Расстояние от пункта A до ближайшей точки B на железной дороге равно 200 км (рис. 18.2). Железная дорога проходит через город C, в котором расположен завод по переработке упомянутого сырья. Расстояние от B до C равно 1000 км. Для доставки сырья на завод строится шоссе AD, соединяющее месторождение с железной дорогой. Стоимость перевозок по шоссе вдвое больше, чем по железной дороге. На каком расстоянии от A должен находиться пункт D, чтобы общая стоимость перевозок сырья с месторождения A в город C по маршруту ADC была минимальной?

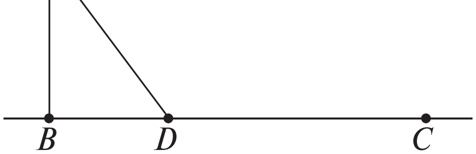


Рис. 18.2

Р е ш е н и е. Обозначим $BD = x$. Тогда $DC = 1000 - x$. Пусть a денежных единиц стоит перевозка одной тонны груза по железной дороге. Тогда перевозка одной тонны по шоссе стоит $2a$. По теореме Пифагора вычисляем длину шоссе AD: $AD = \sqrt{x^2 + 200^2}$. Стоимость перевозки одной тонны по маршруту ADC составляет

$$f(x) = 2a\sqrt{x^2 + 200^2} + a(1000 - x).$$

Из практических соображений ясно, что необходимо найти наименьшее значение этой функции на отрезке [0, 1000]. При этом значению $x = 0$ соответствует маршрут ABC, а $x = 1000$ – маршрут AC.

Вычисляем производную:

$$f'(x) = \frac{2ax}{\sqrt{x^2 + 200^2}} - a.$$

Находим критические точки, приравнивая производную к нулю:

$$\frac{2ax}{\sqrt{x^2 + 200^2}} - a = 0,$$

$$\frac{2ax - a\sqrt{x^2 + 200^2}}{\sqrt{x^2 + 200^2}} = 0,$$

$$2x = \sqrt{x^2 + 200^2},$$

$$3x^2 = 200^2.$$

Нас интересует только неотрицательное значение x:

$$x = \frac{200}{\sqrt{3}}.$$

Это означает, что $BD \approx 115{,}4$ км.

Надо убедиться, что значение функции в точке $x = \dfrac{200}{\sqrt{3}}$ – наименьшее. Для этого вычисляем значения $f(x)$ в указанной точке, в точках $x = 0$, $x = 1000$ и сравниваем их:

$$f\left(\frac{200}{\sqrt{3}}\right) \approx 1346a,$$

$$f(0) = 1400a,$$

$$f(1000) = 2a\sqrt{1040000} > 2000a.$$

Итак, наименьшее значение достигается в критической точке $x = \dfrac{200}{\sqrt{3}}$.

Вопросы

1. Пусть функция $y = f(x)$ возрастает на $[0, +\infty)$. Можно ли утверждать, что $f'(x) > 0$ для всех $x \in [0, +\infty)$?
2. Что такое локальный экстремум?
3. Какая точка называется критической (стационарной) точкой данной функции?
4. Может ли функция иметь два локальных минимума?
5. Имеет ли функция $y = 4 - x^2$ локальный минимум?
6. Пусть функция $f(x)$ непрерывна на X, $x_0 \in X$, $f'(x) < 0$ при $x < x_0$ и $f'(x) > 0$ при $x > x_0$, а в точке $x = x_0$ производной $f'(x)$ не существует. Имеется ли в точке x_0 экстремум, и если имеется, то какой?
7. Сколько экстремумов имеет функция $y = \sin x$ на отрезке $[0, 2\pi]$?
8. Пусть производная функции $y = f(x)$ равна единице на интервале $(-1, 3)$. Будет ли функция возрастающей на этом интервале?
9. Функция $y = f(x)$ дифференцируема на (a, b) и в шести точках этого интервала $f'(x) = 0$. Может ли $f(x)$ иметь на (a, b) четыре минимума?
10. Если функция $y = f(x)$ имеет в точке x_0 максимум, то будет ли иметь максимум функция $y = (f(x))^2$ в этой точке?
11. Имеет ли функция $y = 3x - 4$ экстремумы?
12. Может ли функция $y = f(x)$ в некоторой точке $x \in (a, b)$ иметь значение меньшее, чем любой из минимумов этой функции на (a, b)?
13. Может ли наименьшее значение функции $y = f(x)$, $x \in [a, b]$ находиться в точке $x = b$?
14. Пусть функция $y = f(x)$ имеет на $[a, b]$ локальный максимум и локальный минимум. Может ли ее наибольшее значение не совпадать с локальным максимумом, а наименьшее – с локальным минимумом?
15. Можно ли найти наибольшее и наименьшее значения функции на отрезке, не исследуя ее на локальный экстремум, а зная только ее значения в критических точках?

Глава 19
ИССЛЕДОВАНИЕ ФУНКЦИЙ С ПОМОЩЬЮ ВТОРОЙ ПРОИЗВОДНОЙ. ПОЛНОЕ ИССЛЕДОВАНИЕ ФУНКЦИЙ И ПОСТРОЕНИЕ ГРАФИКОВ

19.1. ВТОРОЕ ДОСТАТОЧНОЕ УСЛОВИЕ ЭКСТРЕМУМА

***Теорема* 19.1.** Пусть $f(x)$ и ее производные $f'(x)$ и $f''(x)$ существуют и непрерывны в некоторой окрестности точки x_0 и $f'(x_0) = 0$. Тогда:

1) если $f''(x_0) < 0$, то x_0 есть точка локального максимума;

2) если $f''(x_0) > 0$, то x_0 есть точка локального минимума.

Доказательство. Пусть $f'(x_0) = 0$, $f''(x_0) < 0$. Так как $f''(x)$ непрерывна, то $f''(x) < 0$ не только в самой точке x_0, но и в некоторой ее окрестности. Но $f''(x)$ есть производная от первой производной. Поэтому из неравенства $f''(x) < 0$ следует, что первая производная убывает в упомянутой выше окрестности. Однако в точке x_0 она обращается в нуль, стало быть, слева от x_0 производная $f'(x)$ положительна, а справа – отрицательна. В силу теоремы 18.2 в точке x_0 имеется локальный максимум.

Случай $f''(x_0) > 0$ рассматривается аналогично.

Заметим, что если в точке x_0 обе производные обращаются в нуль: $f'(x_0) = 0$ и $f''(x_0) = 0$, то доказанная теорема не дает ответа на вопрос о максимуме или минимуме. В этом случае можно либо применить первое достаточное условие экстремума, либо привлечь высшие производные.

19.2. ВЫПУКЛОСТЬ И ВОГНУТОСТЬ ГРАФИКА ФУНКЦИИ. ТОЧКИ ПЕРЕГИБА

Рассмотрим на плоскости кривую $y = f(x)$, являющуюся графиком функции $f(x)$.

Будем говорить, что кривая $y=f(x)$ имеет на (a, b) **выпуклость, направленную вверх**, если все точки кривой лежат н и ж е любой ее касательной на этом интервале.

Будем говорить, что кривая $y=f(x)$ имеет на (b, c) **выпуклость, направленную вниз**, если все точки кривой лежат в ы ш е любой ее касательной на этом интервале.

На рис. 19.1 выпуклость кривой на (a, b) направлена вверх, а на (b, c) – вниз.

Если выпуклость кривой направлена в в е р х, то кривая называется **выпуклой**; если же выпуклость направлена в н и з – **вогнутой**.

Кривая, изображенная на рис. 19.1, является выпуклой на (a, b) и вогнутой на (b, c).

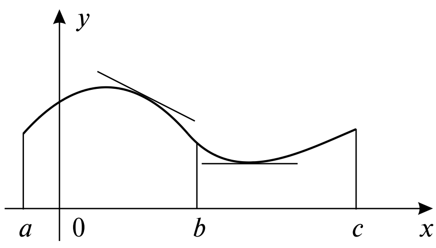


Рис. 19.1. Направление выпуклости графика

***Теорема* 19.2.** Если функция $y=f(x)$ имеет на интервале (a, b) вторую производную и $f''(x) > 0$ на (a, b), то график этой функции имеет на (a, b) выпуклость, направленную вниз; если же $f''(x) < 0$ на (a, b), то график имеет на (a, b) выпуклость, направленную вверх.

Д о к а з а т е л ь с т в о. Для определенности рассмотрим случай $f''(x) < 0$. Пусть x_0 – произвольная точка интервала (a, b). Напишем уравнение касательной к графику функции $y=f(x)$, проходящей через точку $M_0(x_0, f(x_0))$:

$$Y=f(x_0)+f'(x_0)(x-x_0). \tag{19.1}$$

Здесь Y – текущая ордината касательной.

Разложим функцию $f(x)$ в окрестности точки x_0 по формуле Тейлора при $n=1$:

$$y=f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(\xi)}{2!}(x-x_0)^2, \quad \xi\in(a,b). \tag{19.2}$$

(Здесь $R_n=R_1=\dfrac{f''(\xi)}{2!}(x-x_0)^2$.)

Из (19.2) вычтем (19.1):

$$y - Y = \frac{f''(\xi)}{2!}(x - x_0)^2. \qquad (19.3)$$

По условию $f''(x) < 0$ на (a, b), следовательно, правая часть равенства (19.3) отрицательна, т.е. $y < Y$ для всех $x \in (a, b)$. А это и означает, что кривая $y = f(x)$ находится ниже касательной. Теорема доказана.

Определение. Точка, отделяющая выпуклую часть кривой от вогнутой, называется **точкой перегиба**.

Необходимое условие перегиба в точке x_0 для графика функции $f(x)$, имеющей в этой точке непрерывную вторую производную, заключается в том, что

$$f''(x_0) = 0.$$

Действительно, предположив противное, например, что $f''(x_0) > 0$, получаем, что в окрестности точки x_0 кривая имеет выпуклость, направленную вниз, и точка x_0 не может быть точкой перегиба.

Достаточным условием перегиба является смена знака второй производной функции $y = f(x)$ при переходе через точку x_0. Иначе говоря, если вторая производная имеет разные знаки слева и справа от x_0, то график функции имеет перегиб при $x = x_0$.

Действительно, в этом случае направления выпуклости графика слева и справа от x_0 различны, а это и означает наличие перегиба в точке x_0.

Из сказанного следует, что **схема нахождения точек перегиба** такова:

1) найти точки, в которых вторая производная $f''(x)$ равна нулю или не существует (такие точки также называют *критическими*);

2) исследовать знак второй производной слева и справа от каждой такой точки.

***Пример* 19.1.** Найти точки перегиба и направления выпуклости графика функции $f(x) = (1 - x)e^x$.

Решение. Находим первую и вторую производные:

$$f'(x) = -e^x + (1 - x)e^x = -xe^x, \qquad f''(x) = -e^x - xe^x = -(x + 1)e^x.$$

Приравнивая вторую производную к нулю, находим критическую точку $x = -1$. Очевидно, слева от этой точки вторая производная положительна, а справа отрицательна. Следовательно, при $x < -1$ выпуклость графика направлена вниз, а при $x > -1$ – вверх. Точка $x = -1$ есть точка перегиба.

19.3. АСИМПТОТЫ

Определение. Прямая линия называется **асимптотой** графика функции $y = f(x)$, если расстояние от точки M, лежащей на графике, до этой прямой стремится к нулю при неограниченном удалении точки M от начала координат.

Различают три вида асимптот: *вертикальные*, *горизонтальные* и *наклонные*.

Прямая $x = a$ называется **вертикальной асимптотой** графика функции $y = f(x)$, если хотя бы одно из предельных значений $\lim\limits_{x\to a+} f(x)$ или $\lim\limits_{x\to a-} f(x)$ равно $+\infty$ или $-\infty$.

Прямая $y = b$ называется **горизонтальной асимптотой** графика функции $y = f(x)$ при $x \to +\infty$ $(x \to -\infty)$, если $\lim\limits_{x\to+\infty} f(x)=b$ $\left(\lim\limits_{x\to-\infty} f(x)=b\right)$.

Прямая $y = kx + b$ называется **наклонной асимптотой** графика функции $y=f(x)$ при $x \to +\infty$ $(x \to -\infty)$, если $f(x)$ можно представить в виде $f(x) = kx + b + \alpha(x)$, где $\alpha(x) \to 0$ при $x \to +\infty$ $(x \to -\infty)$.

Наличие наклонной асимптоты обусловлено одновременным существованием двух пределов:

$$\lim_{x\to\pm\infty} \frac{f(x)}{x} = k, \qquad \lim_{x\to\pm\infty}\left[f(x)-kx\right]=b$$

(при этом следует отдельно рассматривать $x \to +\infty$ и $x \to -\infty$).

Примеры вертикальных и горизонтальных асимптот хорошо известны из школьного курса математики. В частности, график функции $y=\dfrac{1}{x}$ имеет вертикальную асимптоту $x = 0$ и горизонтальную асимптоту $y = 0$; график функции $y = \operatorname{tg} x$ имеет бесконечно много вертикальных асимптот: $x=\pm\dfrac{\pi}{2},\ x=\pm\dfrac{3\pi}{2},\ \ldots$.

***Пример* 19.2.** Найти наклонную асимптоту графика функции $y=\dfrac{x^3}{x^2-x+1}$.

Решение. Наклонная асимптота имеет вид $y = kx + b$, найдем k и b:

$$k = \lim_{x\to+\infty}\frac{f(x)}{x} = \lim_{x\to+\infty}\frac{x^3}{x(x^2-x+1)} = \lim_{x\to+\infty}\frac{x^2}{x^2-x+1} = 1,$$

$$b=\lim_{x\to+\infty}\left[f(x)-kx\right]=\lim_{x\to+\infty}\left(\frac{x^3}{x^2-x+1}-x\right)=\lim_{x\to+\infty}\frac{x^2-x}{x^2-x+1}=1.$$

Уравнение асимптоты: $y = x + 1$.

Пример **19.3.** В модели потребительского спроса используются, в частности, *функции Торнквиста*, моделирующие связь между величиной дохода и величиной спроса потребителей на: а) товары первой необходимости; б) товары второй необходимости; в) предметы роскоши:

а) $y=\dfrac{b_1(x-a_1)}{x-c_1}\quad (x>a_1)$; б) $y=\dfrac{b_2(x-a_2)}{x-c_2}\quad (x>a_2)$;

в) $y=\dfrac{b_3x(x-a_3)}{x-c_3}\quad (x>a_3)$.

Графики первых двух из этих функций имеют горизонтальные асимптоты $y=b_1$ и $y=b_2$ соответственно:

$$\lim_{x\to+\infty}\frac{b_1(x-a_1)}{x-c_1}=b_1;\qquad \lim_{x\to+\infty}\frac{b_2(x-a_2)}{x-c_2}=b_2,$$

а график последней – наклонную асимптоту:

$$k=\lim_{x\to+\infty}\frac{f(x)}{x}=\lim_{x\to+\infty}\frac{b_3x(x-a_3)}{x(x-c_3)}=b_3,$$

$$b=\lim_{x\to+\infty}\left[\frac{b_3x(x-a_3)}{x-c_3}-b_3x\right]=\lim_{x\to+\infty}\frac{b_3x^2-b_3a_3x-b_3x^2+b_3c_3x}{x-c_3}=$$

$$=b_3c_3-b_3a_3.$$

Уравнение наклонной асимптоты: $y=b_3x+b_3(c_3-a_3)$.

Здесь мы традиционно обозначили аргумент через x, а функцию – через y. Следует заметить, что обычно для этих функций применяются другие обозначения:

а) $x=\dfrac{\alpha I}{I+\beta}$; б) $x=\dfrac{\alpha(I-\gamma)}{I+\beta}$; в) $x=\dfrac{\alpha I(I-\gamma)}{I+\beta}$.

19.4. ОБЩАЯ СХЕМА ИССЛЕДОВАНИЯ ФУНКЦИЙ И ПОСТРОЕНИЯ ГРАФИКОВ

Приведем **схему исследования функции и построения ее графика**.

1. Найти область определения функции.
2. Найти точки разрыва функции.

3. Определить интервалы возрастания и убывания функции.

4. Найти максимумы и минимумы.

5. Определить направление выпуклости графика функции, точки перегиба.

6. Найти асимптоты.

Кроме того, учитываются четность (или нечетность) функции, ее периодичность, точки пересечения графика с осями координат.

На основании проведенного исследования строится график функции. При этом полезно намечать элементы графика параллельно с исследованием.

Пример **19.4.** Исследовать функцию $y=\dfrac{x^3}{2(x-1)^2}$ и построить ее график.

Решение. 1. Область определения функции: $(-\infty, 1)\cup(1, +\infty)$, т.е. $x\neq 1$.

2. $x = 1$ – точка разрыва второго рода, так как

$$\lim_{x\to 1-}\frac{x^3}{2(x-1)^2}=\lim_{x\to 1+}\frac{x^3}{2(x-1)^2}=+\infty.$$

3. Вычислим производную:

$$f'(x)=\frac{3x^2(x-1)^2-2x^3(x-1)}{2(x-1)^4}=\frac{3x^2(x-1)-2x^3}{2(x-1)^3}=$$

$$=\frac{x^3-3x^2}{2(x-1)^3}=\frac{x^2(x-3)}{2(x-1)^3}.$$

Определим области возрастания и убывания функции:

$x\in(-\infty, 1)\Rightarrow f'(x)>0\Rightarrow$ функция возрастает;

$x\in(1, 3)\Rightarrow f'(x)<0\Rightarrow$ функция убывает;

$x\in(3, +\infty)\Rightarrow f'(x)>0\Rightarrow$ функция возрастает.

4. Приравнивая производную к нулю, находим критическую точку $x = 3$. В точке $x = 3$ производная меняет знак с минуса на плюс ($f'(x)<0$ при $1<x<3$; $f'(x)>0$ при $x>3$). Следовательно, в точке $x = 3$ имеется минимум и $f_{\min}=f(3)=\dfrac{27}{8}$.

5. Вычислим вторую производную:

$$f''(x)=\frac{(3x^2-6x)(x-1)^3-3(x^3-3x^2)(x-1)^2}{2(x-1)^6}=$$

$$=\frac{(3x^2-6x)(x-1)-3(x^3-3x^2)}{2(x-1)^4}=\frac{3x}{(x-1)^4}.$$

Определим направление выпуклости и точки перегиба:

$x < 0 \Rightarrow f''(x) < 0 \Rightarrow$ выпуклость направлена вверх;

$x > 0 \Rightarrow f''(x) > 0 \Rightarrow$ выпуклость направлена вниз;

$x = 0 \Rightarrow f''(x) = 0 \Rightarrow$ точка $(0, 0)$ – точка перегиба.

6. Найдем асимптоты графика. Очевидно, $x = 1$ – вертикальная асимптота.

Найдем наклонную асимптоту:

$$k = \lim_{x\to\pm\infty}\frac{f(x)}{x} = \lim_{x\to\pm\infty}\frac{x^3}{2x(x-1)^2} = \lim_{x\to\pm\infty}\frac{x^2}{2(x-1)^2} = \frac{1}{2},$$

$$b = \lim_{x\to\pm\infty}\left[\frac{x^3}{2(x-1)^2} - \frac{1}{2}x\right] = \lim_{x\to\pm\infty}\frac{2x^2 - x}{2(x-1)^2} = 1.$$

Итак, $y = \frac{1}{2}x + 1$ – наклонная асимптота.

График исследуемой функции изображен на рис. 19.2.

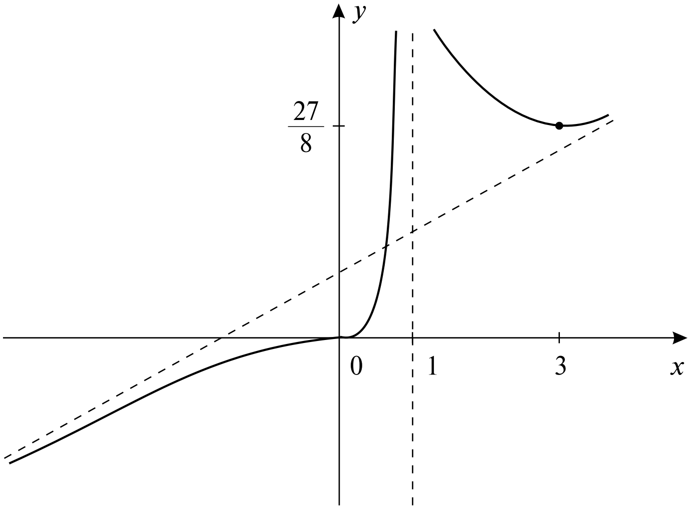


Рис. 19.2. График функции $y = \dfrac{x^3}{2(x-1)^2}$

Пример **19.5.** В теории вероятностей и в математической статистике важную роль играет дифференциальная функция нормального распределения:

$$y = \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Исследуем эту функцию методами дифференциального исчисления по приведенной выше схеме и построим ее график. Заметим, что этот график называют *нормальной кривой* (*кривой Гаусса*).

Р е ш е н и е. 1. Область определения функции – вся ось Ox.

2. Функция непрерывна на всей оси Ox.

3. Вычислим первую производную:

$$f'(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}}\left(-\frac{(x-a)^2}{2\sigma^2}\right)'=-\frac{x-a}{\sigma^3\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Легко видеть, что $f'(x)>0$ при $x<a$, $f'(x)<0$ при $x>a$. Следовательно, на интервале $(-\infty, a)$ функция возрастает, а на интервале $(a, +\infty)$ – убывает.

4. Приравнивая производную к нулю, находим критическую точку $x=a$. В точке $x=a$ производная меняет знак с плюса на минус, следовательно, в ней имеется максимум:

$$f_{\max}=f(a)=\frac{1}{\sigma\sqrt{2\pi}}.$$

5. Вычислим вторую производную:

$$f''(x)=-\frac{1}{\sigma^3\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}}-\frac{x-a}{\sigma^3\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}}\left(-\frac{(x-a)^2}{2\sigma^2}\right)'=$$

$$=-\frac{1}{\sigma^3\sqrt{2\pi}}\,e^{-\frac{(x-a)^2}{2\sigma^2}}\left(1-\frac{(x-a)^2}{\sigma^2}\right).$$

Легко видеть, что вторая производная равна нулю, когда

$$1-\frac{(x-a)^2}{\sigma^2}=0,$$

т.е. при $x=a+\sigma$ и $x=a-\sigma$.

Имеем:

$x\in(-\infty, a-\sigma)\Rightarrow f''(x)>0\Rightarrow$ выпуклость направлена вниз;

$x\in(a-\sigma, a+\sigma)\Rightarrow f''(x)<0\Rightarrow$ выпуклость направлена вверх;

$x\in(a+\sigma, +\infty)\Rightarrow f''(x)>0\Rightarrow$ выпуклость направлена вниз.

При переходе через точки $x=a+\sigma$, $x=a-\sigma$ вторая производная меняет знак. Значение функции в обеих этих точках одно и то же:

$$f(a+\sigma)=f(a-\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{1}{2}}=\frac{1}{\sigma\sqrt{2\pi e}}.$$

Таким образом, точками перегиба графика являются точки

$$\left(a-\sigma,\ \frac{1}{\sigma\sqrt{2\pi e}}\right) \quad \text{и} \quad \left(a+\sigma,\ \frac{1}{\sigma\sqrt{2\pi e}}\right).$$

6. Вертикальных асимптот, очевидно, нет. Предел функции при $x \to \pm\infty$ равен нулю:

$$\lim_{x\to+\infty} f(x) = \lim_{x\to-\infty} f(x) = 0.$$

Следовательно, ось Ox есть горизонтальная асимптота графика (очевидно, $\lim\limits_{x\to+\infty} \frac{f(x)}{x} = 0$, и наклонных асимптот нет).

При построении графика учтем дополнительно, что при всех значениях аргумента $f(x) > 0$, т.е. кривая расположена выше оси Ox, а также тот факт, что кривая симметрична относительно прямой $x = a$ (рис. 19.3), так как в аналитическом выражении функции разность $x - a$ содержится в квадрате.

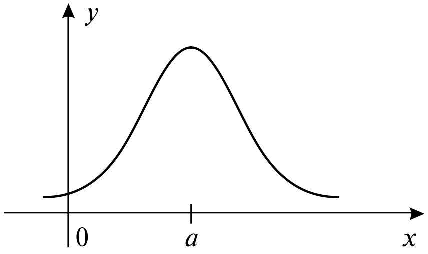


Рис. 19.3. Нормальная кривая

19.5. ИССЛЕДОВАНИЕ ФУНКЦИИ НА МАКСИМУМ И МИНИМУМ С ПОМОЩЬЮ ПРОИЗВОДНЫХ ВЫСШИХ ПОРЯДКОВ

Если в некоторой точке x_0 обращаются в нуль и первая, и вторая производные: $f'(x_0) = 0$, $f''(x_0) = 0$, то в этой критической точке может быть максимум или минимум либо нет ни того, ни другого. В этом случае можно воспользоваться высшими производными.

Пусть в окрестности точки $x = x_0$ функция $f(x)$ имеет производные до n-го порядка включительно и n-я производная $f^{(n)}(x)$ непрерывна в этой точке. Пусть все производные до $(n-1)$-го порядка включительно в этой точке обращаются в нуль:

$$f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0, \qquad (*)$$

и пусть при этом $f^{(n)}(x_0) \neq 0$. Разложим разность $f(x) - f(x_0)$ по степеням разности $x - x_0$ по формуле Тейлора с остаточным членом в форме Пеано:

$$f(x)-f(x_0)=\frac{f'(x_0)}{1!}(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\ldots+$$

$$+\frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1}+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+o\Big((x-x_0)^n\Big).$$

Здесь $o\Big((x-x_0)^n\Big)$ есть $\alpha(x)(x - x_0)^n$, где $\alpha(x) \to 0$ при $x \to x_0$. Кроме того, в силу условия (*) первые $(n - 1)$ слагаемых в правой части последнего равенства обращаются в нуль. Поэтому

$$f(x)-f(x_0)=\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\alpha(x)(x-x_0)^n.$$

Положим $\alpha(x)=\dfrac{\beta(x)}{n!}$. Очевидно, $\beta(x)$ также есть бесконечно малая при $x \to x_0$: $\lim\limits_{x\to x_0}\beta(x)=0.$ Получаем

$$f(x)-f(x_0)=\frac{f^{(n)}(x_0)+\beta(x)}{n!}(x-x_0)^n. \qquad (**)$$

Так как $\beta(x) \to 0$ при $x \to x_0$, то для значений x, достаточно близких к x_0, знак суммы $f^{(n)}(x_0) + \beta(x)$, стоящей в числителе, совпадает со знаком $f^{(n)}(x_0)$ как для $x < x_0$, так и для $x > x_0$. Рассмотрим два случая:

1) n – нечетное число, $n = 2k + 1$. Тогда при переходе от значений x, меньших x_0, к значениям x, бóльшим x_0, выражение $(x - x_0)^{2k+1}$ изменит знак на противоположный:

$$(x - x_0)^{2k+1} < 0 \text{ при } x < x_0,$$

$$(x - x_0)^{2k+1} > 0 \text{ при } x > x_0.$$

При этом знак первого множителя в (**), совпадающий со знаком $f^{(n)}(x_0)$, не изменится. Следовательно, знак разности $f(x) - f(x_0)$ изменится. Поэтому в точке x_0 функция $f(x)$ не может иметь экстремум, так как вблизи этой точки принимает значения как меньше $f(x_0)$, так и больше $f(x_0)$;

2) n – четное число, $n = 2k$. В этом случае разность $f(x) - f(x_0)$ не меняет знака при переходе от x, меньших x_0, к зна-

чениям, бóльшим x_0, так как, очевидно, $(x-x_0)^{2k}>0$ при всех значениях x. Очевидно, вблизи x_0 как слева, так и справа знак разности $f(x)-f(x_0)$ совпадает со знаком $f^{(n)}(x_0)$. Поэтому если $f^{(n)}(x_0)>0$, то $f(x)>f(x_0)$ в некоторой окрестности точки x_0, следовательно, функция $f(x)$ имеет в точке x_0 минимум; если же $f^{(n)}(x_0)<0$, то функция имеет максимум.

Мы вывели следующее ***правило***: если при $x=x_0$

$$f'(x_0)=f''(x_0)=\ldots=f^{(n-1)}(x_0)=0, \tag{*}$$

а $f^{(n)}(x_0)\neq 0$ и при этом n – нечетное, то функция $f(x)$ не имеет в точке $x=x_0$ ни максимума, ни минимума.

Если же первая из производных, не равных нулю в точке x_0, есть производная четного порядка, то функция имеет в точке x_0 экстремум: максимум, если $f^{(n)}(x_0)<0$, и минимум, если $f^{(n)}(x_0)>0$.

***Пример* 19.6.** Исследовать на максимум и минимум функцию

$$f(x)=x^4+8x^3+24x^2+24x.$$

Решение. Найдем критические точки:

$$f'(x)=4x^3+24x^2+48x+24=4(x^3+6x^2+12x+8).$$

Из уравнения $4(x^3+6x^2+12x+8)=0$ получаем единственную критическую точку $x=-2$.

Рассмотрим значения производных в точке $x=-2$:

$$f''(x)=12x^2+48x+48, \qquad f''(-2)=0;$$

$$f'''(x)=24x+48, \qquad f'''(-2)=0;$$

$$f^{(4)}(x)=24>0.$$

Следовательно, функция $f(x)$ имеет в точке $x=-2$ минимум.

***Пример* 19.7.** Исследовать на экстремум функцию

$$f(x)=e^x-e^{-x}-2\sin x.$$

Решение. Вычислим производную:

$$f'(x)=e^x+e^{-x}-2\cos x.$$

Очевидно, точка $x=0$ является критической: $f'(0)=0$. Имеем:

$$f''(x)=e^x-e^{-x}+2\sin x, \qquad f''(0)=0;$$

$$f'''(x)=e^x+e^{-x}+2\cos x, \qquad f'''(0)=4.$$

Здесь первая не обратившаяся в нуль в критической точке производная имеет третий, т.е. нечетный, порядок. Следовательно, в этой критической точке экстремума нет.

Вопросы

1. Пусть x_0 – критическая точка функции $y = f(x)$ и пусть $f''(x_0) = 0$. Можно ли утверждать, что в точке x_0 есть экстремум?
2. Пусть график функции $y = f(x)$ имеет выпуклость, направленную вверх. Куда направлена выпуклость кривой $y = \lambda f(x)$: а) при $\lambda > 0$; б) при $\lambda < 0$?
3. Пусть график функции $y = f(x)$ имеет на интервале (a, b) три точки перегиба: x_1, x_2 и x_3 $(x_1 < x_2 < x_3)$, и пусть $y = f(x)$ является выпуклой кривой на (a, x_1). Выпуклой или вогнутой является эта кривая на (x_3, b)?
4. Пусть $f''(x_0) = 0$. Можно ли утверждать, что x_0 – точка перегиба?
5. Пусть $y = f(x)$ имеет горизонтальную асимптоту при $x \to +\infty$. Чему равен предел $\lim\limits_{x\to+\infty} \dfrac{f(x)}{x}$?
6. Могут ли у графика функции $y = f(x)$ быть две разные наклонные асимптоты?
7. Как исследовать на экстремум в точке x_0 функцию $y = f(x)$, если $f'(x_0) = 0$ и $f''(x_0) = 0$?

Глава 20
ПРИЛОЖЕНИЯ ПРОИЗВОДНОЙ В ЭКОНОМИЧЕСКОЙ ТЕОРИИ

20.1. МАКСИМИЗАЦИЯ ПРИБЫЛИ

Рассмотрим экономическую интерпретацию теоремы Ферма.

Пусть $S = S(x)$ – функция издержек, $D = D(x)$ – функция дохода, $P = P(x)$ – функция прибыли. Тогда $P(x) = D(x) - S(x)$. *Оптимальным уровнем производства*, очевидно, является такой, при котором прибыль м а к с и м а л ь н а, т.е. такое значение выпуска x_0, при котором функция прибыли $P(x)$ имеет максимум. В силу теоремы Ферма в этой точке $x = x_0$ производная равна нулю: $P'(x_0) = 0$. Но $P'(x) = D'(x) - S'(x)$, поэтому

$$D'(x_0) = S'(x_0). \qquad (*)$$

Производная $S'(x)$ выражает предельные издержки MS, а производная $D'(x)$ – предельный доход MD. Таким образом, равенство (*), полученное с помощью теоремы Ферма, приобретает вид

$$MS(x_0) = MD(x_0).$$

Последнее равенство есть выражение одного из базовых законов микроэкономики: *максимум прибыли достигается при р а в е н с т в е предельных издержек* и *предельного дохода*.

20.2. ЭЛАСТИЧНОСТЬ

Перейдем теперь к логарифмической производной и ее приложениям. Пусть функция $y = f(x)$ положительна и дифференцируема в точке x. Как уже отмечалось, в частности, при выводе формулы производной для показательной функции (см. § 16.5), производная от функции $\ln y = \ln f(x)$ имеет вид

$$\left[\ln f(x)\right]' = \frac{1}{f(x)} f'(x) = \frac{f'(x)}{f(x)}, \qquad \text{или} \qquad (\ln y)' = \frac{y'}{y}.$$

Это выражение называется **логарифмической производной** функции $f(x)$. Логарифмическую производную называют также **темпом изменения** T_y функции y:

$$T_y = (\ln y)' = \frac{y'}{y}. \tag{20.1}$$

Во многих прикладных задачах используется понятие *эластичности* функции.

Определение. **Эластичностью** $E_x(y)$ функции $y = f(x)$ называется предел отношения относительного приращения функции y к относительному приращению аргумента x при $\Delta x \to 0$:

$$E_x(y) = \lim_{\Delta x \to 0}\left(\frac{\Delta y}{y} : \frac{\Delta x}{x}\right) = \frac{x}{y}\lim_{\Delta x \to 0}\frac{\Delta y}{\Delta x} = \frac{x}{y}y' = x\frac{y'}{y}. \tag{20.2}$$

Эластичность функции приближенно выражает процентное изменение функции $y = f(x)$ при изменении аргумента x на 1%.

Из формулы (20.2) следует, что *эластичность функции равна произведению независимой переменной* x *на темп изменения функции* T_y:

$$E(y) = E_x(y) = xT_y. \tag{20.3}$$

Отметим **свойства эластичности**:

$$E(uv) = E(u) + E(v), \tag{20.4}$$

$$E\left(\frac{u}{v}\right) = E(u) - E(v), \tag{20.5}$$

очевидным образом следующие из соответствующих свойств логарифмов.

Эластичность функции применяется при анализе спроса и предложения. Пусть $D = D(p)$ – функция спроса от цены товара p.

Эластичность спроса относительно цены определяется отношением:

$$E = \frac{\text{Процентное изменение спроса}}{\text{Процентное изменение цены}}. \tag{20.6}$$

Процентное изменение спроса – это $\frac{\Delta D}{D}\cdot 100$, а процентное изменение цены – это $\frac{\Delta p}{p}\cdot 100$. Поэтому

$$E = \left(\frac{\Delta D}{D}\cdot 100\right) : \left(\frac{\Delta p}{p}\cdot 100\right),$$

или

$$E = \frac{p}{D}\frac{\Delta D}{\Delta p}. \tag{20.7}$$

При непрерывной зависимости ΔD от Δp разностное отношение в выражении (20.7) заменяют пределом при $\Delta p \to 0$:

$$E(D) = p\frac{D'(p)}{D(p)}. \tag{20.8}$$

Ввиду того, что функция спроса $D = D(p)$ является убывающей (см. рис. 13.14), ее производная отрицательна и эластичность спроса также отрицательна. (Некоторые авторы определяют эластичность как положительную величину, ставя перед правой частью формул (20.6)–(20.8) знак «минус».)

Различают **три вида спроса**:

1) *эластичный*, если $|E(D)| > 1$;

2) *нейтральный*, если $|E(D)| = 1$;

3) *неэластичный*, если $|E(D)| < 1$.

Пример **20.1.** Функция спроса имеет вид $D = \sqrt{240 - p}$. Найти эластичность спроса при цене $p = 176$.

Решение: $E(D) = p\dfrac{D'}{D} = p\dfrac{1}{-2(240-p)}$. Подставляя $p = 176$,

получаем $E(D) = -\dfrac{176}{128} = -1{,}375$;

$|E(D)| = 1{,}375 > 1$ – спрос эластичен.

Аналогично вводится понятие **эластичности предложения** как отношение процентного изменения предложения к процентному изменению цены. Так как функция предложения $S = S(p)$ – возрастающая (см. рис. 13.15), то

$$E(S) = p\frac{S'}{S}$$

есть положительная величина.

20.3. ОПТИМИЗАЦИЯ НАЛОГООБЛОЖЕНИЯ

Пусть t – налог с единицы выпускаемой продукции, $S = S(x)$ – функция издержек, $D = D(x)$ – функция дохода, $P = P(x)$ – функция прибыли. Тогда функция прибыли имеет вид

$$P(x) = D(x) - S(x) - tx.$$

Пусть, например, цена на продукцию $v(x) = a - bx$, т.е. линейно уменьшается с увеличением объема продукции, а функция издержек имеет вид $S(x) = x^2 + c$. Здесь a, b, c – некоторые положительные константы. Функция прибыли в этом случае имеет вид

$$P(x) = x(a - bx) - x^2 - c - tx.$$

Желая максимизировать прибыль, фирма ищет оптимальный объем производства. Условие максимума прибыли: $P'(x) = 0$, т.е. $a - t - 2bx - 2x = 0$, откуда

$$x_0 = \frac{a-t}{2b+2}.$$

При таком значении объема продукции суммарный налог T имеет вид $T = \dfrac{t(a-t)}{2(b+1)}$. Интересы государства заключаются в том, чтобы величина T была максимальной. Дифференцируем T и, приравнивая производную к нулю: $T' = 0$, $a - 2t = 0$, получаем

$$t_0 = \frac{a}{2}.$$

Рассмотрим эту задачу при конкретных числовых значениях констант a, b и c. Пусть $a = 80$, $b = 1$, $c = 10$. Тогда $t_0 = 40$, $x_0 = 10$. При этих значениях максимальная величина прибыли $P_0 = 190$, а доход государства $T_0 = t_0 x_0 = 400$. (Заметим, что при отсутствии налогов максимальная прибыль достигалась бы при вдвое большем объеме производства $x_0 = 20$ и составляла бы $P_0 = 790$.)

Вопросы

1. При каком соотношении между предельными издержками и предельным доходом достигается максимум прибыли?
2. Что такое темп изменения функции?
3. Что называется эластичностью функции? Как эластичность связана с темпом изменения функции?
4. Как определяется эластичность спроса относительно цены?
5. В каких случаях спрос считается эластичным?
6. Как определяется понятие эластичности предложения? Положительной или отрицательной величиной является эластичность предложения?

Раздел VI
ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ

Глава 21
НЕОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ. МЕТОДЫ ИНТЕГРИРОВАНИЯ

21.1. ПЕРВООБРАЗНАЯ И НЕОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

Определение. Функция $y = F(x)$ называется **первообразной** для функции $f(x)$ на промежутке X, если для любого $x \in X$ выполняется равенство

$$F'(x) = f(x).$$

Например:

функция $F(x) = x^3$ является первообразной для функции $f(x) = 3x^2$ на всей числовой прямой, так как для любого x выполняется равенство $(x^3)' = 3x^2$;

функция $F(x) = \arcsin x$ является первообразной для функции $f(x) = \dfrac{1}{\sqrt{1-x^2}}$ на интервале $(-1, 1)$, так как в каждой точке этого интервала выполняется равенство $(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$.

Очевидно, если для данной функции $f(x)$ существует первообразная, то эта первообразная не является единственной; если $F(x)$ – первообразная для $f(x)$ и C – произвольная константа, то $F(x) + C$ также есть первообразная для $f(x)$.

Лемма. Если $f'(x) = 0$ на всем промежутке X, то $f(x)$ постоянна на промежутке X.

Доказательство. Достаточно доказать, что для любых двух различных точек $x_1, x_2 \in X$ выполняется равенство $f(x_1) = f(x_2)$. Применим к $f(x)$ на $[x_1, x_2]$ теорему Лагранжа (см. § 17.1):

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1), \quad c \in (x_1, x_2).$$

Так как $f'(x) = 0$ в любой точке, то, в частности, $f'(c) = 0$, следовательно, $f(x_1) = f(x_2)$, что и требовалось доказать.

***Теорема* 21.1.** Если $F(x)$ – первообразная для функции $f(x)$ на промежутке X, то любая другая первообразная для $f(x)$ на X может быть представлена в виде $F(x) + C$, где C – некоторое число.

Доказательство. Пусть $F'(x) = f(x)$ и $\Phi'(x) = f(x)$ для всех $x \in X$. Тогда

$$[\Phi(x) - F(x)]' = 0.$$

В силу доказанной выше леммы $\Phi(x) - F(x) = C = \text{const}$, т.е. $\Phi(x) = F(x) + C$, что и требовалось доказать.

Определение. Если $F(x)$ – первообразная для $f(x)$, то выражение $F(x) + C$, где C – произвольная постоянная, называется **неопределенным интегралом** от функции $f(x)$ и обозначается $\int f(x)\,dx$.

Итак, если $F'(x) = f(x)$, то

$$\int f(x)\,dx = F(x) + C. \tag{21.1}$$

Неопределенный интеграл от функции $f(x)$ есть совокупность всех первообразных для функции $f(x)$.

В выражении (21.1) знак $\int$ называется *знаком интеграла*, $f(x)$ – подынтегральной функцией, $f(x)\,dx$ – *подынтегральным выражением*, C – *постоянной интегрирования*.

Операция нахождения неопределенного интеграла от данной функции называется **интегрированием** этой функции.

Интегрирование является операцией, обратной дифференцированию. Правильность интегрирования проверяется дифференцированием.

Свойства неопределенного интеграла:

1. $\left(\int f(x)dx\right)' = f(x),$

т.е. *производная от неопределенного интеграла равна подынтегральной функции*.

2. $d\int f(x)\,dx = f(x)\,dx,$

т.е. *дифференциал от неопределенного интеграла равен подынтегральному выражению*.

3. $\int dF(x) = F(x) + C.$

Эти свойства непосредственно следуют из определения интеграла.

Правила интегрирования:

I. $\int cf(x)\,dx = c\int f(x)\,dx$, где $c = \text{const}$, $c \neq 0$.

II. $\int [f(x) \pm g(x)]\,dx = \int f(x)\,dx \pm \int g(x)\,dx$.

III. Если $F'(x) = f(x)$, то $\int f(ax+b)dx = \frac{1}{a}F(ax+b)+C$.

Правила I и II следуют из соответствующих правил дифференцирования (см. § 16.4). Убедимся в справедливости равенства III:

$$\left(\frac{1}{a}F(ax+b)+C\right)' = \frac{1}{a}f(ax+b)(ax+b)' =$$

$$= \frac{1}{a}f(ax+b)a = f(ax+b).$$

Таблица интегралов

1. $\int 0\,dx = C$.

2. $\int x^{\alpha}dx = \frac{x^{\alpha+1}}{\alpha+1}+C \quad (\alpha \neq -1)$.

3. $\int \frac{dx}{x} = \ln|x|+C$.

4. $\int a^x dx = \frac{a^x}{\ln a}+C$.

4′. $\int e^x dx = e^x + C$.

5. $\int \cos x\,dx = \sin x + C$.

6. $\int \sin x\,dx = -\cos x + C$.

7. $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$.

8. $\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$.

9. $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C$.

9′. $\int \frac{dx}{\sqrt{a^2-x^2}} = \arcsin\frac{x}{a} + C$.

10. $\int \frac{dx}{1+x^2} = \operatorname{arctg} x + C$.

10′. $\int \frac{dx}{a^2+x^2} = \frac{1}{a}\operatorname{arctg}\frac{x}{a} + C$.

11. $\int \operatorname{ctg} x\,dx = \ln|\sin x| + C$.

12. $\int \operatorname{tg} x\,dx = -\ln|\cos x| + C$.

13. $\int \frac{dx}{\sqrt{x^2+A}} = \ln\left|x+\sqrt{x^2+A}\right| + C$.

14. $\int \frac{dx}{x^2-a^2} = \frac{1}{2a}\ln\left|\frac{x-a}{x+a}\right| + C \quad (a \neq 0)$.

Большинство формул, приведенных в таблице интегралов, непосредственно следует из таблицы производных (см. § 16.5). Справедливость любой из формул легко проверяется дифференцированием правой части.

Проверим формулу 10′. По правилу дифференцирования сложной функции

$$\left(\frac{1}{a}\operatorname{arctg}\frac{x}{a}+C\right)'=\frac{1}{a}\,\frac{1}{1+\left(\dfrac{x}{a}\right)^2}\left(\frac{x}{a}\right)'=\frac{1}{a}\,\frac{1}{1+\dfrac{x^2}{a^2}}\,\frac{1}{a}=\frac{1}{a^2+x^2}.$$

Аналогично проверяется формула 9′.

Проверим формулу 11. Если $\sin x>0$, то

$$\left(\ln|\sin x|+C\right)'=(\ln\sin x)'=\frac{1}{\sin x}(\sin x)'=\frac{\cos x}{\sin x}=\operatorname{ctg}x.$$

Если $\sin x<0$, то

$$\left(\ln|\sin x|+C\right)'=[\ln(-\sin x)]'=\frac{1}{-\sin x}(-\sin x)'=\frac{-\cos x}{-\sin x}=\operatorname{ctg}x.$$

Аналогично проверяется формула 12.

Проверим формулу 13. Пусть $x+\sqrt{x^2+A}>0.$ Тогда

$$\left(\ln\left|x+\sqrt{x^2+A}\right|+C\right)'=\left[\ln\left(x+\sqrt{x^2+A}\right)\right]'=$$

$$=\frac{1}{x+\sqrt{x^2+A}}\left(x+\sqrt{x^2+A}\right)'=\frac{1}{x+\sqrt{x^2+A}}\left(1+\frac{x}{\sqrt{x^2+A}}\right)=$$

$$=\frac{1}{x+\sqrt{x^2+A}}\,\frac{\sqrt{x^2+A}+x}{\sqrt{x^2+A}}=\frac{1}{\sqrt{x^2+A}}.$$

Аналогично рассматривается случай $x+\sqrt{x^2+A}<0.$ Также аналогично проверяется формула 14.

21.2. ОСНОВНЫЕ МЕТОДЫ ИНТЕГРИРОВАНИЯ

Непосредственное интегрирование

Вычисление интегралов, основанное на применении формул 1–14 и правил I–III (см. § 21.1), называется *непосредственным интегрированием*.

Рассмотрим ***примеры***:

1. $\int\left(6x^2+2\cos x-\frac{3}{x}\right)dx=6\int x^2dx+2\int\cos x\,dx-3\int\frac{dx}{x}=$

$=2x^3+2\sin x-3\ln|x|+C.$

Следует заметить, что в конце решения записывают одну общую постоянную C, не выписывая постоянных от интегрирования отдельных слагаемых.

2. $\int x\sqrt{x}\,dx=\int x^{\frac{3}{2}}dx=\frac{x^{\frac{5}{2}}}{\frac{5}{2}}+C=\frac{2x^2\sqrt{x}}{5}+C.$

3. $\int\frac{3x^2+1}{x^2(x^2+1)}dx=\int\frac{x^2+1+2x^2}{x^2(x^2+1)}dx=$

$=\int\frac{x^2+1}{x^2(x^2+1)}dx+\int\frac{2x^2}{x^2(x^2+1)}dx=$

$=\int\frac{dx}{x^2}+2\int\frac{dx}{x^2+1}=-\frac{1}{x}+2\,\mathrm{arctg}\,x.$

Метод подстановки (метод замены переменной)

Замена переменной интегрирования – один из самых распространенных и эффективных методов сведения неопределенного интеграла к комбинации табличных.

Пусть дан интеграл $\int f(x)dx.$ Введем новую переменную по формуле $x=\varphi(t)$, где $\varphi(t)$ – дифференцируемая функция, тогда $dx=\varphi'(t)\,dt.$ Подставим эти выражения в интеграл:

$$\int f(x)dx=\int f(\varphi(t))\,\varphi'(t)\,dt. \qquad (21.2)$$

Формула (21.2) называется **формулой замены переменной**.

***Пример* 21.1.** Найти: а) $\int\frac{\sqrt{x}}{x+1}dx$; б) $\int x(x-1)^9dx.$

Решение. а) Сделаем замену переменной: $x=t^2$, $t=\sqrt{x}$. Тогда $dx=2tdt.$ Получаем

$$\int\frac{\sqrt{x}}{x+1}dx=\int\frac{t\cdot 2t\,dt}{t^2+1}=2\int\frac{t^2}{t^2+1}dt=2\int\frac{t^2+1-1}{t^2+1}dt=$$

$$=2\left(\int dt-\int\frac{dt}{t^2+1}\right)=2(t-\mathrm{arctg}\,t)+C=2\left(\sqrt{x}-\mathrm{arctg}\sqrt{x}\right)+C.$$

б) Положим $x = t + 1$. Тогда $x - 1 = t$, $dx = dt$. Получаем

$$\int x(x-1)^9\,dx = \int (t+1)\,t^9 dt = \int \left(t^{10} + t^9\right) dt =$$

$$= \frac{t^{11}}{11} + \frac{t^{10}}{10} + C = \frac{(x-1)^{11}}{11} + \frac{(x-1)^{10}}{10} + C.$$

Нередко замену переменной делают не в виде $x = \varphi(t)$, а в виде $t = \psi(x)$.

***Пример* 21.2.** Найти: а) $\int \sin^2 x \cos x\,dx$; б) $\int \frac{2x\,dx}{1+x^4}$.

Решение. а) Сделаем замену $t = \sin x$. Тогда $dt = \cos x\,dx$. Получаем

$$\int \sin^2 x \cos x\,dx = \int t^2 dt = \frac{t^3}{3} + C = \frac{\sin^3 x}{3} + C.$$

б) Положим $t = x^2$. Тогда $dt = 2x\,dx$. Получаем

$$\int \frac{2x\,dx}{1+x^4} = \int \frac{dt}{1+t^2} = \operatorname{arctg} t + C = \operatorname{arctg} x^2 + C.$$

Заметим, что новую переменную можно и не выписывать явно. В таких случаях говорят о *подведении под знак дифференциала*.

В частности, вычисление последнего интеграла в примере 21.2 можно записать в таком виде:

$$\int \frac{2x\,dx}{1+x^4} = \int \frac{d(x^2)}{1+(x^2)^2} = \operatorname{arctg} x^2 + C.$$

***Пример* 21.3.** $\int \frac{\ln x\,dx}{x} = \int \ln x\,d(\ln x) = \frac{\ln^2 x}{2} + C.$

Интегрирование по частям

Пусть $u = u(x)$ и $v = v(x)$ – дифференцируемые функции. Тогда $d(uv) = vdu + udv$, или

$$udv = d(uv) - vdu.$$

Интегрируя обе части последнего равенства, получаем **формулу интегрирования по частям**:

$$\int udv = uv - \int vdu. \tag{21.3}$$

***Пример* 21.4.** Найти: а) $\int x e^x dx$; б) $\int (2x + 3) \cos x\,dx$; в) $\int x \ln x\,dx$.

Р е ш е н и е. а) Положим $x = u$, $e^x dx = dv$. Тогда $du = dx$, $u = e^x$. По формуле (21.3) получаем

$$\int x e^x dx = xe^x - \int e^x dx = xe^x - e^x + C = e^x(x-1) + C.$$

б) Положим $u = 2x + 3$, $\cos x\, dx = dv$. Тогда $du = 2dx$, $v = \sin x$. Получаем

$$\int (2x+3)\cos x\, dx = (2x+3)\sin x - 2\int \sin x\, dx =$$
$$= (2x+3)\sin x + 2\cos x + C.$$

в) Положим $u = \ln x$, $dv = x dx$. Тогда $du = \frac{dx}{x}$, $v = \frac{x^2}{2}$. Получаем

$$\int x \ln x dx = \frac{x^2}{2}\ln x - \int \frac{x^2}{2}\cdot\frac{1}{x} dx = \frac{x^2}{2}\ln x - \frac{1}{2}\int x dx = \frac{x^2}{2}\ln x - \frac{x^2}{4} + C.$$

В ряде случаев формулу интегрирования по частям применяют несколько раз, постепенно упрощая подынтегральное выражение.

***Пример* 21.5.** Найти $\int x^2 \cos x\, dx$.

Р е ш е н и е. Положим $u = x^2$, $\cos x\, dx = dv$. Тогда $du = 2x dx$, $v = \sin x$. Получаем

$$\int x^2 \cos x\, dx = x^2 \sin x - \int 2x \sin x\, dx.$$

Возникший интеграл не является табличным, но он проще исходного, так как степень переменной x понизилась. Применим повторно формулу интегрирования по частям, положив $u = x$, $\sin x\, dx = dv$. Тогда $du = dx$, $v = -\cos x$. Получаем

$$\int x^2 \cos x\, dx = x^2 \sin x - \int 2x \sin x\, dx =$$
$$= x^2 \sin x - 2(-x\cos x + \int \cos x\, dx) =$$
$$= x^2 \sin x + 2x\cos x - 2\sin x + C.$$

Укажем наиболее распространенные типы интегралов, для нахождения которых применяется формула интегрирования по частям:

1. $\int P_n(x) e^{ax} dx$, $\int P_n(x) \sin ax\, dx$, $\int P_n(x) \cos ax\, dx$.
2. $\int P_n(x) \ln x\, dx$, $\int P_n(x) \arcsin x\, dx$, $\int P_n(x) \arccos x\, dx$,
   $\int P_n(x)\, \mathrm{arctg}\, x\, dx$, $\int P_n(x)\, \mathrm{arcctg}\, x\, dx$.

Здесь $P_n(x)$ – многочлен n-й степени.

Для нахождения интегралов п е р в о й группы полагают $P_n(x) = u$ (остальные множители составляют dv). Для нахождения интегралов в т о р о й группы полагают $P_n(x)\,dx = dv$ (остальные множители берут в качестве u).

21.3. ИНТЕГРИРОВАНИЕ РАЦИОНАЛЬНЫХ ДРОБЕЙ

Выражение $\dfrac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ – многочлены, называется **рациональной дробью**. Рациональная дробь называется **правильной**, если степень числителя меньше степени знаменателя. Если же степень числителя больше или равна степени знаменателя, то дробь называется **неправильной**.

Неправильную дробь можно представить в виде суммы многочлена и правильной дроби, поделив числитель на знаменатель:

$$\frac{P(x)}{Q(x)} = R(x) + \frac{S(x)}{Q(x)}.$$

Здесь $R(x)$ – некоторый многочлен, а второе слагаемое – правильная дробь.

Например, $\dfrac{x^5+3x^4+x^3+2x^2+3}{x^3+x^2-x+1} = x^2+2x+\dfrac{3x^2-2x+3}{x^3+x^2-x+1}$.

Для того чтобы проинтегрировать правильную дробь, ее разлагают на простые дроби, предварительно разложив знаменатель на элементарные множители.

Приведем без доказательства **формулу разложения правильной дроби**. Пусть знаменатель $Q(x)$ разлагается на множители $(x-a)^{\alpha}(x^2+px+q)^{\beta}$. Здесь $x = a$ – действительный корень $Q(x)$ кратности α, x^2+px+q – квадратный трехчлен с отрицательным дискриминантом. Тогда правильная дробь разлагается на сумму элементарных дробей с помощью так называемого *метода неопределенных коэффициентов* следующим образом:

$$\frac{P(x)}{(x-a)^{\alpha}(x^2+px+q)^{\beta}} = \frac{A_1}{(x-a)^{\alpha}} + \frac{A_2}{(x-a)^{\alpha-1}} + \ldots + \frac{A_{\alpha}}{x-a} + \ldots +$$

$$+\frac{M_1x+N_1}{(x^2+px+q)^{\beta}} + \frac{M_2x+N_2}{(x^2+px+q)^{\beta-1}} + \ldots + \frac{M_{\beta}x+N_{\beta}}{x^2+px+q}, \qquad (21.4)$$

где коэффициенты подлежат вычислению в процессе разложения дроби.

В связи с упомянутым разложением необходимо рассмотреть так называемые **простые дроби**:

I. $\dfrac{A}{x-a}$. **III.** $\dfrac{Mx+N}{x^2+px+q}$.

II. $\dfrac{A}{(x-a)^{\alpha}}$. **IV.** $\dfrac{Mx+N}{(x^2+px+q)^{\beta}}$.

Здесь a, p, q, A, M, N – действительные числа; α, β – натуральные числа; кроме того, предполагается, что знаменатели дробей III и IV типов не имеют действительных корней, т.е. $\left(\dfrac{p}{2}\right)^2 - q < 0$.

Рассмотрим **интегралы от простых дробей**.

Дроби I и II типов интегрируются просто:

$$\int \frac{A}{x-a}dx = A\ln|x-a| + C,$$

$$\int \frac{A}{(x-a)^{\alpha}}dx = -\frac{1}{\alpha-1}\,\frac{A}{(x-a)^{\alpha-1}} + C.$$

Для вычисления интеграла от дроби III типа из трехчлена в знаменателе выделяют полный квадрат:

$$x^2+px+q = \left(x+\frac{p}{2}\right)^2 + \left(q-\frac{p^2}{4}\right) = \left(x+\frac{p}{2}\right)^2 + a^2;$$

$$\int \frac{Mx+N}{x^2+px+q}dx = \int \frac{\frac{M}{2}(2x+p) + \left(N - \frac{Mp}{2}\right)}{x^2+px+q}dx =$$

$$= \frac{M}{2}\int \frac{2x+p}{x^2+px+q}dx + \left(N-\frac{Mp}{2}\right)\int \frac{dx}{\left(x+\frac{p}{2}\right)^2 + \left(q-\frac{p^2}{4}\right)} =$$

$$= \frac{M}{2}\ln(x^2+px+q) + \frac{2N-Mp}{\sqrt{4q-p^2}}\ \mathrm{arctg}\frac{2x+p}{\sqrt{4q-p^2}} + C.$$

***Пример* 21.6.** Найти $\displaystyle\int \frac{3x+4}{x^2+2x+5}dx$.

Р е ш е н и е. Можно воспользоваться выведенной выше формулой для интеграла от дроби III типа, но мы еще раз повторим процесс ее вывода на этом конкретном примере:

$$\int \frac{3x+4}{x^2+2x+5}dx = \frac{3}{2}\int \frac{2x+2}{x^2+2x+5}dx + \int \frac{dx}{(x+1)^2+4} =$$

$$= \frac{3}{2}\int \frac{d(x^2+2x+5)}{x^2+2x+5} + \int \frac{dx}{(x+1)^2+4} = \frac{3}{2}\ln(x^2+2x+5) + \frac{1}{2}\operatorname{arctg}\frac{x+1}{2} + C.$$

Перейдем к вычислению интеграла от дроби IV типа, т.е. интеграла $\int \frac{Mx+N}{(x^2+px+q)^\beta}dx$ при $\beta \geq 2$.

Применим ту же подстановку $x + \frac{p}{2} = t$, что и в случае с дробью III типа:

$$\int \frac{Mx+N}{(x^2+px+q)^\beta}dx = \int \frac{Mt + \left(N - \frac{Mp}{2}\right)}{(t^2+a^2)^\beta}dt =$$

$$= \frac{M}{2}\int \frac{2t\,dt}{(t^2+a^2)^\beta} + \left(N - \frac{Mp}{2}\right)\int \frac{dt}{(t^2+a^2)^\beta}.$$

Первый из полученных интегралов берется подстановкой $t^2 + a^2 = z$, $2t\,dt = dz$:

$$\int \frac{2t\,dt}{(t^2+a^2)^\beta} = \int \frac{dz}{z^\beta} = -\frac{1}{\beta-1}\,\frac{1}{z^{\beta-1}} + C = -\frac{1}{\beta-1}\,\frac{1}{(t^2+a^2)^{\beta-1}} + C.$$

Вычисление второго из оставшихся интегралов требует определенных усилий. Итак, нам нужно вычислить интеграл

$$J_\beta = \int \frac{dt}{(t^2+a^2)^\beta} \qquad (\beta = 1, 2, 3, \ldots).$$

Применим формулу интегрирования по частям. Положим $u = \frac{1}{(t^2+a^2)^\beta}$, $dv = dt$; отсюда $du = -\frac{2\beta t\,dt}{(t^2+a^2)^{\beta+1}}$, $v = t$. Получаем

$$J_\beta = \frac{t}{(t^2+a^2)^\beta} + 2\beta\int \frac{t^2}{(t^2+a^2)^{\beta+1}}dt. \qquad (*)$$

Преобразовываем последний интеграл:

$$\int \frac{t^2}{(t^2+a^2)^{\beta+1}}dt = \int \frac{t^2+a^2-a^2}{(t^2+a^2)^{\beta+1}}dt =$$

$$= \int \frac{dt}{(t^2+a^2)^\beta} - a^2\int \frac{dt}{(t^2+a^2)^{\beta+1}} = J_\beta - a^2 J_{\beta+1}.$$

Подставим последнее выражение в равенство (*):

$$J_\beta = \frac{t}{(t^2+a^2)^\beta} + 2\beta J_\beta - 2\beta a^2 J_{\beta+1},$$

откуда

$$J_{\beta+1} = \frac{1}{2\beta a^2}\,\frac{t}{(t^2+a^2)^\beta} + \frac{2\beta-1}{2\beta}\,\frac{1}{a^2}J_\beta. \qquad (**)$$

Полученная формула сводит вычисление интеграла

$$J_{\beta+1} = \int \frac{dt}{(t^2+a^2)^{\beta+1}}$$

к вычислению интеграла

$$J_\beta = \int \frac{dt}{(t^2+a^2)^\beta}.$$

(Напомним, что β – натуральное число.)

Очевидно,

$$J_1 = \int \frac{dt}{t^2+a^2} = \frac{1}{a}\operatorname{arctg}\frac{t}{a} + C.$$

Возьмем одно из его значений, а именно при $C = 0$. По формуле (**) при $\beta = 1$ найдем

$$J_2 = \int \frac{dt}{(t^2+a^2)^2} = \frac{1}{2a^2}\,\frac{t}{t^2+a^2} + \frac{1}{2a^3}\operatorname{arctg}\frac{t}{a},$$

при $\beta = 2$ получим

$$J_3 = \int \frac{dt}{(t^2+a^2)^3} = \frac{1}{4a^2}\,\frac{t}{(t^2+a^2)^2} + \frac{3}{4a^2}J_2 =$$

$$= \frac{1}{4a^2}\,\frac{t}{(t^2+a^2)^2} + \frac{3}{8a^4}\,\frac{t}{t^2+a^2} + \frac{3}{8a^5}\operatorname{arctg}\frac{t}{a}$$

и т.д. Теперь остается только вспомнить, что $t = x + \frac{p}{2}$, и вернуться к переменной x. Заметим, что формулы вида (**), позволяющие сводить вычисление $J_{\beta+1}$ к вычислению J_β с меньшим на единицу знаком, называются *рекуррентными*. Мы встречались с рекуррентными соотношениями и раньше, в линейной алгебре – там, где сводили вычисление определителя порядка n к вычислению определителей порядка $n - 1$ [см. формулу (3.7)].

Рассмотрим теперь **интегрирование рациональных дробей, не являющихся простыми**. Как уже говорилось, рациональная дробь, если она не является правильной, преобразуется и представляется в виде суммы многочлена и правильной дроби. Затем правильная дробь разлагается на простые дроби. При этом применяется метод неопределенных коэффициентов. После того как дробь представлена в виде суммы простых дробей, вычисляют интеграл от нее как сумму интегралов от этих простых дробей. Рассмотрим это на примерах.

***Пример* 21.7.** Вычислить интеграл $\int \frac{2x^3 - x^2 + 2x + 1}{x^4 - 2x^3 + 2x^2 - 2x + 1} dx.$

Решение. Подынтегральная функция – правильная дробь. Разложим ее на простые дроби. Для этого сначала разложим знаменатель на множители: $x^4 - 2x^3 + 2x^2 - 2x + 1 = (x - 1)^2(x^2 + 1)$. (Это можно сделать, например, найдя подбором корень $x = 1$ и деля на $(x - 1)$.) Применяем формулу разложения (21.4):

$$\frac{2x^3 - x^2 + 2x + 1}{(x-1)^2(x^2+1)} = \frac{A_1}{(x-1)^2} + \frac{A_2}{x-1} + \frac{Mx+N}{x^2+1}.$$

Умножим обе части этого равенства на знаменатель левой части:

$$2x^3 - x^2 + 2x + 1 = A_1(x^2+1) + A_2(x-1)(x^2+1) + (Mx+N)(x-1)^2.$$

После приведения подобных членов получаем

$$2x^3 - x^2 + 2x + 1 = (A_2 + M)x^3 + (A_1 - A_2 - 2M + N)x^2 +$$

$$+ (A_2 + M - 2N)x + A_1 - A_2 + N.$$

Приравниваем коэффициенты при одинаковых степенях x слева и справа:

$$\begin{array}{l|l} x^3 & \quad A_2 + M \qquad\quad = 2, \\ x^2 & A_1 - A_2 - 2M + N = -1, \\ x^1 & \quad A_2 + M - 2N = 2, \\ x^0 & A_1 - A_2 \qquad\; + N = 1. \end{array}$$

Решая систему, находим: $A_1 = 2$, $A_2 = 1$, $M = 1$, $N = 0$. Итак,

$$\frac{2x^3 - x^2 + 2x + 1}{(x-1)^2(x^2+1)} = \frac{2}{(x-1)^2} + \frac{1}{x-1} + \frac{x}{x^2+1}.$$

Интегрируем:

$$\int \frac{2x^3-x^2+2x+1}{x^4-2x^3+2x^2-2x+1}dx = 2\int \frac{dx}{(x-1)^2} + \int \frac{dx}{x-1} + \int \frac{xdx}{x^2+1} =$$

$$= -\frac{2}{x-1} + \ln|x-1| + \frac{1}{2}\ln(x^2+1) + C.$$

***Пример* 21.8.** Вычислить интеграл

$$I = \int \frac{2x^6-5x^5+6x^4-10x^3+8x^2-3x+15}{x^5-2x^4+2x^3-4x^2+x-2}dx.$$

Решение. Подынтегральная функция – неправильная рациональная дробь. Деля числитель на знаменатель, выделяем целую часть:

$$2x^6 - 5x^5 + 6x^4 - 10x^3 + 8x^2 - 3x + 15 =$$

$$= (2x - 1)(x^5 - 2x^4 + 2x^3 - 4x^2 + x - 2) + 2x^2 + 2x + 13.$$

Следовательно, подынтегральная функция имеет вид

$$2x-1+\frac{2x^2+2x+13}{x^5-2x^4+2x^3-4x^2+x-2}.$$

Знаменатель оставшейся правильной дроби разложим на множители:

$$x^5 - 2x^4 + 2x^3 - 4x^2 + x - 2 = x^4(x - 2) + 2x^2(x - 2) + x - 2 =$$

$$= (x - 2)(x^4 + 2x^2 + 1) = (x - 2)(x^2 + 1)^2.$$

(Можно было бы поступить иначе – подбором найти корень знаменателя $x = 2$, а затем разделить знаменатель на $(x - 2)$.)

Итак,

$$I = \int (2x-1)\,dx + \int \frac{2x^2+2x+13}{(x-2)(x^2+1)^2}dx.$$

Первый из этих двух интегралов вычисляется сразу:

$$\int (2x - 1)\,dx = x^2 - x + C,$$

а для вычисления второго интеграла разложим его подынтегральную функцию, являющуюся правильной дробью, на простые дроби:

$$\frac{2x^2+2x+13}{(x-2)(x^2+1)^2} = \frac{A}{x-2} + \frac{M_1x+N_1}{(x^2+1)^2} + \frac{M_2x+N_2}{x^2+1}.$$

Приводя дроби в правой части к общему знаменателю, приравниваем числители:

$$2x^2 + 2x + 13 = A(x^2 + 1)^2 + (M_1x + N_1)(x - 2) +$$
$$+ (M_2x + N_2)(x^2 + 1)(x - 2).$$

Приравнивая коэффициенты при одинаковых степенях x слева и справа, приходим к системе из пяти линейных уравнений:

$$\begin{array}{c|ccccccccccc}
x^4 & A & & & + & M_2 & & & & & = & 0, \\
x^3 & & & & - & 2M_2 & & & + & N_2 & = & 0, \\
x^2 & 2A & + & M_1 & + & M_2 & & & - & 2N_2 & = & 2, \\
x^1 & & - & 2M_1 & - & 2M_2 & + & N_1 & + & N_2 & = & 2, \\
x^0 & A & & & & & - & 2N_1 & - & 2N_2 & = & 13.
\end{array}$$

Решая эту систему, находим:

$$A = 1,\ M_1 = -3,\ N_1 = -4,\ M_2 = -1, N_2 = -2.$$

Получаем

$$\frac{2x^2+2x+13}{(x-2)(x^2+1)^2} = \frac{1}{x-2} - \frac{3x+4}{(x^2+1)^2} - \frac{x+2}{x^2+1}.$$

Вычисляем интегралы от каждого из слагаемых, стоящих справа:

$$\int \frac{dx}{x-2} = \ln|x-2| + C,$$

$$\int \frac{x+2}{x^2+1}dx = \frac{1}{2}\ln(x^2+1) + 2\operatorname{arctg} x + C,$$

$$\int \frac{3x+4}{(x^2+1)^2}dx = \frac{3}{2}\int \frac{2x\,dx}{(x^2+1)^2} + 4\int \frac{dx}{(x^2+1)^2} = -\frac{3}{2(x^2+1)} + 4\int \frac{dx}{(x^2+1)^2}.$$

Последний интеграл вычисляется с помощью рекуррентной формулы (**):

$$\int \frac{dx}{(x^2+1)^2} = \frac{1}{2}\frac{x}{x^2+1} + \frac{1}{2}\operatorname{arctg} x.$$

С учетом коэффициентов после очевидных преобразований получаем

$$\int \frac{2x^6 - 5x^5 + 6x^4 - 10x^3 + 8x^2 - 3x + 15}{x^5 - 2x^4 + 2x^3 - 4x^2 + x - 2}dx =$$

$$= x^2 - x + \frac{1}{2}\frac{3-4x}{x^2+1} + \frac{1}{2}\ln\frac{(x-2)^2}{x^2+1} - 4\operatorname{arctg} x + C.$$

21.4. ИНТЕГРИРОВАНИЕ ИРРАЦИОНАЛЬНЫХ ФУНКЦИЙ

Рассмотрим случаи, когда замена переменной позволяет свести интегралы от иррациональных функций к интегралам от рациональных функций (т.е. рационализирует интеграл). Обозначим через $R(u, v)$ рациональную функцию от u и v, т.е. функцию, которая получена с использованием лишь арифметических операций над переменными u и v.

1. Рассмотрим интеграл вида $\int \boldsymbol{R}\left(\boldsymbol{x}, \boldsymbol{x}^{\frac{m}{n}}, \boldsymbol{x}^{\frac{r}{s}}, \ldots\right) \boldsymbol{dx}$. Пусть k – общий знаменатель дробей $\frac{m}{n}, \frac{r}{s}, \ldots$. Сделаем подстановку:

$$x = t^k, \; dx = kt^{k-1}dt.$$

Тогда, очевидно, подынтегральная функция преобразуется в рациональную функцию от t.

***Пример* 21.9.** Вычислить интеграл $\int \frac{\sqrt{x}\,dx}{\sqrt[4]{x^3}+1}$.

Решение. Наименьшее общее кратное показателей корней равно 4. Поэтому делаем подстановку $x = t^4$, $dx = 4t^3dt$:

$$\int \frac{\sqrt{x}\,dx}{\sqrt[4]{x^3}+1} = 4\int \frac{t^2}{t^3+1}t^3dt = 4\int \frac{t^5}{t^3+1}dt = 4\int\left(t^2 - \frac{t^2}{t^3+1}\right)dt =$$

$$= 4\int t^2dt - \frac{4}{3}\int \frac{d(t^3+1)}{t^3+1} = \frac{4}{3}t^3 - \frac{4}{3}\ln\left|t^3+1\right| + C =$$

$$= \frac{4}{3}\left[\sqrt[4]{x^3} - \ln\left(\sqrt[4]{x^3}+1\right)\right] + C.$$

2. Рассмотрим теперь интеграл вида

$$\int \boldsymbol{R}\left(\boldsymbol{x}, \left(\frac{\boldsymbol{ax}+\boldsymbol{b}}{\boldsymbol{cx}+\boldsymbol{d}}\right)^{\frac{m}{n}}, \left(\frac{\boldsymbol{ax}+\boldsymbol{b}}{\boldsymbol{cx}+\boldsymbol{d}}\right)^{\frac{r}{s}}, \ldots\right)\boldsymbol{dx}.$$

Такие интегралы вычисляются подстановкой

$$\frac{ax+b}{cx+d} = t^k,$$

где k – общий знаменатель дробей $\frac{m}{n}, \frac{r}{s}, \ldots$.

Пример **21.10.** Вычислить $\int \frac{dx}{\sqrt{x+5}\left(\sqrt[4]{x+5}+1\right)}$.

Решение. Здесь $a=1, b=5, c=0, d=1$; $\frac{m}{n}=\frac{1}{4}$, $\frac{r}{s}=\frac{1}{2}$. Делаем подстановку $x+5=t^4$, $dx=4t^3dt$. Получаем

$$\int \frac{dx}{\sqrt{x+5}\left(\sqrt[4]{x+5}+1\right)}=4\int \frac{t^3dt}{t^2(t+1)}=4\int \frac{t\,dt}{t+1}=4\int\left(1-\frac{1}{t+1}\right)dt=$$

$$=4(t-\ln|t+1|)+C=4\left[\sqrt[4]{x+5}-\ln\left(\sqrt[4]{x+5}+1\right)\right]+C.$$

Некоторые интегралы от иррациональных функций рационализируются тригонометрическими подстановками. В частности, при вычислении интеграла вида $\int \boldsymbol{R}\left(\boldsymbol{x},\ \sqrt{\boldsymbol{a}^2-\boldsymbol{x}^2}\right)\boldsymbol{dx}$ применяется подстановка $x=a\sin t$, а интеграла вида $\int \boldsymbol{R}\left(\boldsymbol{x},\ \sqrt{\boldsymbol{a}^2+\boldsymbol{x}^2}\right)\boldsymbol{dx}$ – подстановка $x=a\,\mathrm{tg}\,t$.

Пример **21.11.** Найти $\int \frac{dx}{\sqrt{(4-x^2)^3}}$.

Решение. Положим $x=2\sin t$. Тогда $dx=2\cos t\,dt$,

$$\int \frac{dx}{\sqrt{(4-x^2)^3}}=\int \frac{2\cos t\,dt}{\sqrt{(4-4\sin^2 t)^3}}=\int \frac{2\cos t\,dt}{8\cos^3 t}=\frac{1}{4}\int \frac{dt}{\cos^2 t}=$$

$$=\frac{1}{4}\,\mathrm{tg}\,t+C=\frac{1}{4}\frac{\sin t}{\cos t}+C=\frac{1}{4}\frac{\sin t}{\sqrt{1-\sin^2 t}}+C=\frac{1}{4}\frac{x}{\sqrt{4-x^2}}+C.$$

21.5. ИНТЕГРИРОВАНИЕ ТРИГОНОМЕТРИЧЕСКИХ ФУНКЦИЙ

Рассмотрим интеграл вида $\int \boldsymbol{R}(\mathbf{sin}\,\boldsymbol{x},\ \mathbf{cos}\,\boldsymbol{x})\boldsymbol{dx}$. Покажем, что он сводится к интегралу от рациональной функции с помощью подстановки

$$t=\mathrm{tg}\,\frac{x}{2},$$

называемой *универсальной тригонометрической подстановкой*.

Действительно, выражая $\sin x$, $\cos x$ и dx через $\operatorname{tg}\frac{x}{2}=t$, получаем:

$$\sin x=\frac{2t}{1+t^2},\quad \cos x=\frac{1-t^2}{1+t^2},\quad x=2\operatorname{arctg}t,\quad dx=\frac{2dt}{1+t^2}.$$

Подставляя полученные выражения в интеграл, получаем

$$\int R(\sin x,\cos x)\,dx=\int R\left(\frac{2t}{1+t^2},\frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2}=\int R_1(t)\,dt.$$

***Пример* 21.12.** Найти $\int\frac{dx}{1+\sin x}$.

Решение. Применяя универсальную подстановку $t=\operatorname{tg}\frac{x}{2}$, после очевидных преобразований получаем:

$$\int\frac{dx}{1+\sin x}=2\int\frac{dt}{(1+t)^2}=-\frac{2}{1+t}+C=-\frac{2}{1+\operatorname{tg}\frac{x}{2}}+C.$$

Следует заметить, однако, что универсальная тригонометрическая подстановка часто приводит к весьма сложным рациональным функциям. Поэтому во многих случаях вместо универсальной подстановки применяют другие подстановки, позволяющие быстрее и проще достичь цели.

***Пример* 21.13.** Найти $\int\sin^4 x\cdot\cos^3 x\,dx$.

Решение. Положим $t=\sin x$. Тогда $dt=\cos x\,dx$. Получаем

$$\int\sin^4 x\cdot\cos^3 x\,dx=\int\sin^4 x\cdot\cos^2 x\cdot\cos x\,dx=\int t^4(1-t^2)\,dt=$$

$$=\frac{t^5}{5}-\frac{t^7}{7}+C=\frac{\sin^5 x}{5}-\frac{\sin^7 x}{7}+C.$$

Вообще, интеграл вида $\int\boldsymbol{\sin^m x\cdot\cos^n x\,dx}$, где m и n – натуральные числа, удобнее вычислять с помощью следующих подстановок:

а) если m – четное, n – нечетное, то подстановка $t=\sin x$;

б) если m – нечетное, n – четное, то подстановка $t=\cos x$;

в) если m и n – нечетные, то любая из подстановок «а» или «б»;

г) если m и n – четные, то применяются формулы понижения степени:

$$\sin^2 x=\frac{1-\cos 2x}{2},\qquad \cos^2 x=\frac{1+\cos 2x}{2}.$$

Интегралы следующих видов: $\int \cos mx \cdot \cos nx\, dx$, $\int \sin mx \cdot \cos nx\, dx$, $\int \sin mx \cdot \sin nx\, dx$ берутся с помощью формул, известных из тригонометрии:

$$\cos\alpha\, \cos\beta = \frac{1}{2}\left[\cos(\alpha+\beta) + \cos(\alpha-\beta)\right],$$

$$\sin\alpha\, \cos\beta = \frac{1}{2}\left[\sin(\alpha+\beta) + \sin(\alpha-\beta)\right],$$

$$\sin\alpha\, \sin\beta = \frac{1}{2}\left[-\cos(\alpha+\beta) + \cos(\alpha-\beta)\right].$$

Пример **21.14.** Найти $\int \cos 5x \cdot \cos 3x\, dx$.

Решение. Воспользуемся формулой произведения косинусов:

$$\int \cos 5x \cdot \cos 3x\, dx = \frac{1}{2}\int (\cos 8x + \cos 2x)\, dx = \frac{\sin 8x}{16} + \frac{\sin 2x}{4} + C.$$

«Неберущиеся» интегралы

Известно, что операция дифференцирования не выводит функцию из класса элементарных функций. С операцией интегрирования дело обстоит сложнее. Не всякий интеграл от элементарной функции можно выразить в конечном виде через элементарные функции. Такие интегралы называют **«неберущимися»**. Укажем некоторые из них:

$\int \frac{\sin x}{x} dx$ – интегральный синус;

$\int \frac{\cos x}{x} dx$ – интегральный косинус;

$\int \frac{dx}{\ln x}$ – интегральный логарифм;

$\int e^{-x^2} dx$ – интеграл Пуассона;

$\int \sin x^2 dx$, $\int \cos x^2 dx$ – интегралы Френеля.

Эти интегралы играют важную роль в прикладных науках. Для вычисления «неберущихся» интегралов применяют приближенные методы, позволяющие оценивать и вычислять такие интегралы с любой степенью точности.

Вопросы

1. Чему равна производная от первообразной для данной функции?
2. Как могут различаться две первообразные одной и той же функции?
3. Как отличается первообразная данной функции от неопределенного интеграла от этой функции?
4. Чему равна производная от неопределенного интеграла?
5. Чему равен дифференциал от неопределенного интеграла?
6. Какие существуют основные методы интегрирования?
7. На какой формуле основан метод замены переменной в неопределенном интеграле?
8. Что такое метод интегрирования по частям? Какова формула интегрирования по частям?
9. Что такое подведение под знак дифференциала?
10. Какая функция называется рациональной дробью?
11. Какая рациональная дробь называется правильной, а какая – неправильной?
12. Как свести интегрирование неправильной рациональной дроби к интегрированию правильной рациональной дроби?
13. Какие рациональные дроби называются простыми?
14. Каков порядок действий при интегрировании рациональной дроби, не являющейся простой?
15. Какой метод применяется при разложении правильной рациональной дроби на простые?
16. Что такое универсальная тригонометрическая подстановка?
17. Почему для интегрирования тригонометрических выражений недостаточно владеть только универсальной тригонометрической подстановкой?
18. Всегда ли интеграл от элементарной функции выражается в конечном виде через элементарные функции?
19. Что такое «неберущиеся» интегралы? Приведите примеры «неберущихся» интегралов.

Глава 22
ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ И ЕГО СВОЙСТВА

22.1. ПОНЯТИЕ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

Пусть функция $f(x)$ определена на отрезке $[a, b]$. Разобьем отрезок $[a, b]$ произвольным образом на n частей точками:

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b.$$

Выберем в каждом из частичных отрезков $[x_{i-1}, x_i]$ произвольную точку ξ_i:

$$x_{i-1} \le \xi_i \le x_i, \qquad 1 \le i \le n.$$

Обозначим через Δx_i длину i-го частичного отрезка: $\Delta x_i = x_i - x_{i-1}$, $1 \le i \le n$. Рассмотрим сумму

$$\sigma_n = f(\xi_1)\,\Delta x_1 + f(\xi_2)\,\Delta x_2 + \ldots + f(\xi_n)\,\Delta x_n = \sum_{i=1}^{n} f(\xi_i)\Delta x_i. \qquad (22.1)$$

Сумма (22.1) называется **интегральной суммой** для функции $f(x)$ на отрезке $[a, b]$.

Если $f(x) > 0$, то интегральная сумма σ_n представляет собой сумму площадей прямоугольников с основаниями Δx_i и высотами $f(\xi_i)$, $i = 1, 2, \ldots, n$, т.е. площадь ступенчатой фигуры, образованной этими прямоугольниками (рис. 22.1).

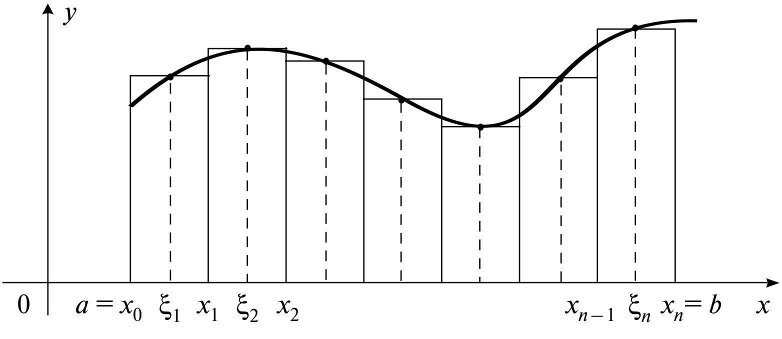


Рис. 22.1. Ступенчатая фигура

Обозначим через λ длину наибольшего частичного отрезка данного разбиения:

$$\lambda = \max_{1 \le i \le n} \Delta x_i.$$

О п р е д е л е н и е. Конечный предел интегральной суммы (22.1) при $\lambda \to 0$, если он существует и не зависит ни от способа разбиения отрезка $[a, b]$, ни от выбора точек ξ_i, называется **определенным интегралом** от функции $f(x)$ на отрезке $[a, b]$ и обозначается $\int\limits_a^b f(x)\,dx$:

$$\int\limits_a^b f(x)\,dx = \lim_{\lambda \to 0} \sigma_n = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i. \qquad (22.2)$$

Если определенный интеграл (22.2) существует, то функция $f(x)$ называется **интегрируемой** на отрезке $[a, b]$. Число a в формуле (22.2) называется *нижним пределом*, число b – *верхним пределом* интеграла, $f(x)$ – *подынтегральной функцией*, x – *переменной интегрирования*, а отрезок $[a, b]$ – *отрезком интегрирования*.

Отметим р а з л и ч и я в понятиях определенного и неопределенного интегралов: неопределенный интеграл $\int f(x)\,dx$ есть *семейство функций*, а определенный интеграл $\int\limits_a^b f(x)\,dx$ есть *определенное число*.

Давая определение понятия определенного интеграла, мы предполагали, что $a < b$. Положим по определению:

$$\int\limits_b^a f(x)\,dx = -\int\limits_a^b f(x)\,dx. \qquad (22.3)$$

Геометрический смысл определенного интеграла

В соответствии с определением понятия определенного интеграла, в случае, когда функция $f(x)$ неотрицательна на $[a, b]$, интеграл $\int\limits_a^b f(x)\,dx$ численно равен *площади* S фигуры, ограниченной сверху кривой $y = f(x)$, снизу – осью Ox, с боков – прямыми $x = a$, $x = b$ (рис. 22.2).

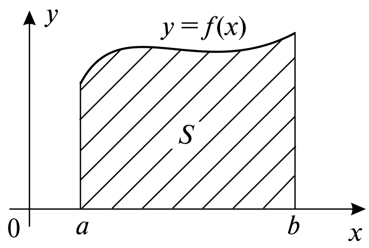


Рис. 22.2. Геометрический смысл определенного интеграла

Экономический смысл определенного интеграла

Пусть функция $z = z(t)$ описывает производительность в зависимости от времени t. Тогда объем v продукции, произведенной за промежуток времени с момента $t = t_0$ до момента $t = T$, выражается интегралом от $z(t)$ на отрезке $[t_0, T]$:

$$v = \int_{t_0}^{T} z(t)\,dt.$$

Классы интегрируемых функций

Достаточное условие существования определенного интеграла дает следующая теорема (приведем ее без доказательства).

***Теорема* 22.1.** Если функция $y = f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на этом отрезке.

Как видно из теоремы 22.1, класс интегрируемых функций шире класса дифференцируемых функций. Известно, что всякая дифференцируемая функция непрерывна, но не всякая непрерывная функция дифференцируема. Итак, непрерывности функции недостаточно для ее дифференцируемости, но достаточно для интегрируемости. Более того, существуют классы интегрируемых функций, не являющихся непрерывными. Приведем без доказательства теоремы об этих функциях.

***Теорема* 22.2.** Если функция $f(x)$ ограничена на отрезке $[a, b]$ и имеет на нем лишь конечное число точек разрыва, то она интегрируема на этом отрезке.

***Теорема* 22.3.** Монотонная на отрезке $[a, b]$ функция $f(x)$ интегрируема на этом отрезке.

Ограниченность интегрируемой функции

***Теорема* 22.4.** Если функция $f(x)$ интегрируема на $[a, b]$, то она ограничена на $[a, b]$.

Доказательство. Предположим, что $f(x)$ не ограничена на $[a, b]$. Тогда она не ограничена хотя бы на одном из частичных отрезков $[x_{i-1}, x_i]$. В этом случае за счет выбора точки ξ_i можно сделать произведение $f(\xi_i)\Delta x_i$ сколь угодно большим, а следовательно, и интегральную сумму σ_n сколь угодно большой; при этих условиях σ_n не имеет конечного предела, следовательно, функция $f(x)$ не интегрируема. Отсюда делаем вывод, что предположение не верно.

22.2. СВОЙСТВА ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

Рассмотрим сначала **свойства определенного интеграла**, **выражаемые равенствами**.

1. По определению полагаем

$$\int_a^a f(x)\,dx = 0, \tag{22.4}$$

т.е. *определенный интеграл с одинаковыми пределами интегрирования равен нулю*.

2. По определению *при перестановке верхнего и нижнего пределов интегрирования интеграл меняет знак на противоположный*:

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx. \tag{22.5}$$

3. *Постоянный множитель можно выносить за знак интеграла*:

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx. \tag{22.6}$$

4. *Интеграл от алгебраической суммы двух функций равен алгебраической сумме интегралов от этих функций*:

$$\int_a^b (f(x) \pm g(x))\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx. \tag{22.7}$$

5. Для любых чисел a, b и c имеет место равенство

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx. \tag{22.8}$$

Сформулируем свойства 3–5 более подробно и докажем их.

***Свойство* 3.** Если функция $f(x)$ интегрируема на $[a, b]$ и $c = \text{const}$, то функция $cf(x)$ также интегрируема на $[a, b]$ и справедливо равенство

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx.$$

Доказательство. Для интегральных сумм справедливо очевидное равенство

$$\sum_{i=1}^n cf(\xi_i)\Delta x_i = c\sum_{i=1}^n f(\xi_i)\Delta x_i.$$

Это равенство справедливо при любом разбиении отрезка $[a, b]$ на частичные отрезки и любом выборе точек ξ_i. Обозначая, как и раньше, $\lambda = \max\limits_{1\le i\le n} \Delta x_i$, переходим к пределу при $\lambda \to 0$:

$$\int_a^b cf(x)\,dx = \lim_{\lambda\to 0}\sum_{i=1}^n cf(\xi_i)\Delta x_i = c\lim_{\lambda\to 0}\sum_{i=1}^n f(\xi_i)\Delta x_i = c\int_a^b f(x)\,dx.$$

***Свойство* 4.** Если $f(x)$ и $g(x)$ интегрируемы на $[a, b]$, то их алгебраическая сумма также интегрируема на $[a, b]$ и справедливо равенство

$$\int_a^b (f(x)\pm g(x))\,dx = \int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx.$$

Доказательство. При любом разбиении отрезка $[a, b]$ и любом выборе точек ξ_i для интегральных сумм выполняется равенство

$$\sum_{i=1}^n (f(\xi_i)\pm g(\xi_i))\Delta x_i = \sum_{i=1}^n f(\xi_i)\Delta x_i \pm \sum_{i=1}^n g(\xi_i)\Delta x_i,$$

поэтому

$$\int_a^b (f(x)\pm g(x))\,dx = \lim_{\lambda\to 0}\sum_{i=1}^n (f(\xi_i)\pm g(\xi_i))\Delta x_i =$$

$$= \lim_{\lambda\to 0}\left(\sum_{i=1}^n f(\xi_i)\Delta x_i \pm \sum_{i=1}^n g(\xi_i)\Delta x_i\right) =$$

$$= \lim_{\lambda\to 0}\sum_{i=1}^n f(\xi_i)\Delta x_i \pm \lim_{\lambda\to 0}\sum_{i=1}^n g(\xi_i)\Delta x_i = \int_a^b f(x)dx \pm \int_a^b g(x)dx.$$

Свойство 5. Для любых трех чисел a, b и c имеет место равенство

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx,$$

если все эти три интеграла существуют.

Доказательство. Рассмотрим сначала случай, когда точка c расположена между точками a и b, т.е. $a < c < b$. Составим интегральную сумму для функции $f(x)$ на отрезке $[a, b]$.

Так как для интегрируемой функции предел интегральной суммы не зависит от способа разбиения отрезка на частичные отрезки, будем разбивать отрезок $[a, b]$ на части так, чтобы одной из точек деления (одним из концов частичного отрезка) была точка c.

Обозначим через $\sum\limits_a^b$ интегральную сумму, соответствующую отрезку $[a, b]$, через $\sum\limits_a^c$ интегральную сумму, соответствующую отрезку $[a, c]$, а через $\sum\limits_c^b$ — соответствующую отрезку $[c, b]$. Тогда, очевидно, справедливо равенство:

$$\sum_a^b f(\xi_i)\Delta x_i = \sum_a^c f(\xi_i)\Delta x_i + \sum_c^b f(\xi_i)\Delta x_i.$$

Переходя в последнем равенстве к пределу при $\lambda \to 0$, получаем

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Пусть теперь точка c находится справа от точки b: $a < b < c$. Тогда на основании доказанного справедливо равенство

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

Отсюда

$$\int_a^b f(x)dx = \int_a^c f(x)dx - \int_b^c f(x)dx.$$

Однако в соответствии со свойством 2

$$\int_b^c f(x)dx = -\int_c^b f(x)dx,$$

поэтому

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Аналогично свойство 5 доказывается и для случая, когда точка c находится слева от отрезка $[a, b]$, и вообще для любого другого расположения точек a, b и c.

Замечание. Мы сформулировали и доказали свойство 5 в предположении, что все три рассматриваемых интеграла существуют. Можно было бы ослабить это требование и доказать свойство 5 в предположении, что интеграл существует только для наибольшего из трех рассматриваемых отрезков.

Теперь рассмотрим **свойства определенного интеграла, выражаемые неравенствами**.

6. Если функция $f(x)$ интегрируема на $[a, b]$, $a < b$ и $f(x) \geq 0$, то

$$\int_a^b f(x)dx \geq 0. \qquad (22.9)$$

(Это следует из того, что все слагаемые в интегральной сумме неотрицательны.)

7. Если функции $f(x)$ и $g(x)$ интегрируемы на $[a, b]$, $a < b$ и $f(x) \geq g(x)$, то

$$\int_a^b f(x)dx \geq \int_a^b g(x)dx. \qquad (22.10)$$

Доказательство. Так как $f(x) - g(x) \geq 0$, то из свойства 6 следует

$$\int_a^b (f(x) - g(x))dx \geq 0.$$

Отсюда, учитывая свойство 4, получаем

$$\int_a^b f(x)dx - \int_a^b g(x)dx \geq 0, \quad \text{или} \quad \int_a^b f(x)dx \geq \int_a^b g(x)dx.$$

8. Пусть функция $f(x)$ интегрируема на $[a, b]$ и удовлетворяет на $[a, b]$ условию $m \leq f(x) \leq M$. Тогда

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a). \qquad (22.11)$$

Доказательство. В силу свойства 7

$$\int_a^b m dx \le \int_a^b f(x)dx \le \int_a^b M dx,$$

но $\int_a^b m dx = m\int_a^b dx = m(b-a)$, $\int_a^b M dx = M\int_a^b dx = M(b-a)$, поэтому

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a).$$

9 (теорема о среднем). Если функция $f(x)$ непрерывна на $[a, b]$, то существует такое $\xi \in [a, b]$, что

$$\int_a^b f(x)dx = f(\xi)(b-a). \tag{22.12}$$

Доказательство. По второй теореме Вейерштрасса (см. § 15.2) непрерывная функция $f(x)$ достигает на $[a, b]$ своего наибольшего значения M и своего наименьшего значения m. Так как $m \le f(x) \le M$, то выполняется неравенство (22.11). Деля это неравенство почленно на $(b - a)$, получаем

$$m \le \frac{1}{b-a}\int_a^b f(x)dx \le M.$$

По второй теореме Больцано – Коши (см. § 15.2) функция $f(x)$ принимает на $[a, b]$ все промежуточные значения между m и M. В частности, существует такое $\xi \in [a, b]$, что

$$f(\xi) = \frac{1}{b-a}\int_a^b f(x)dx.$$

Отсюда получаем (22.12).

22.3. ОСНОВНАЯ ФОРМУЛА ИНТЕГРАЛЬНОГО ИСЧИСЛЕНИЯ

Интеграл с переменным верхним пределом

Если функция интегрируема на отрезке $[a, b]$, то она интегрируема на любом отрезке $[a, x]$, где $x \in [a, b]$.

Заметим, что не имеет значения, какой буквой обозначена переменная интегрирования в определенном интеграле:

$$\int_a^b f(x)dx = \int_a^b f(z)dz = \int_a^b f(t)dt = \dots ,$$

так как смена обозначений никак не влияет на значение интеграла.

Рассмотрим функцию аргумента x:

$$\Phi(x) = \int_a^x f(t)dt. \tag{22.13}$$

Назовем функцию $\Phi(x)$ **интегралом с переменным верхним пределом**. (В формуле (22.13) переменная интегрирования обозначена через t, чтобы не смешивать ее с верхним пределом x.)

Ранее было установлено, что всякая дифференцируемая функция непрерывна (см. теорему 16.2), однако функция, непрерывная в точке, может не иметь производную в этой точке. Докажем теперь, что *каждая непрерывная на данном отрезке функция* $f(x)$ *имеет на этом отрезке первообразную*.

***Теорема* 22.5.** Если функция $f(x)$ непрерывна на $[a, b]$, то функция $\Phi(x)$ является первообразной для $f(x)$:

$$\Phi'(x) = \left(\int_a^x f(t)dt \right)' = f(x).$$

Доказательство. Пусть Δx таково, что $(x + \Delta x) \in [a, b]$. Тогда

$$\Delta\Phi = \Phi(x + \Delta x) - \Phi(x) = \int_a^{x+\Delta x} f(t)dt - \int_a^x f(t)dt = \int_x^{x+\Delta x} f(t)dt.$$

По теореме о среднем [см. формулу (22.12)]

$$\int_x^{x+\Delta x} f(t)dt = f(\xi)\Delta x, \quad \xi \in [x, x + \Delta x],$$

отсюда

$$\frac{\Delta\Phi}{\Delta x} = \frac{\Phi(x + \Delta x) - \Phi(x)}{\Delta x} = f(\xi).$$

При $\Delta x \to 0$, очевидно, $\xi \to x$, и так как $f(x)$ непрерывна в точке x, то $\lim\limits_{\xi \to x} f(\xi) = f(x)$. Получаем

$$\Phi'(x) = \lim_{\Delta x \to 0} \frac{\Delta\Phi}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Phi(x + \Delta x) - \Phi(x)}{\Delta x} = \lim_{\Delta x \to 0} f(\xi) = \lim_{\xi \to x} f(\xi) = f(x).$$

Теорема доказана.

Доказанную ***теорему*** можно сформулировать также в следующем виде: *производная от определенного интеграла по переменному верхнему пределу равна подынтегральной функции* (в которую вместо переменной интегрирования подставлено значение верхнего предела).

Формула Ньютона – Лейбница

Согласно теореме 22.5 для непрерывной на $[a, b]$ функции $f(x)$ интеграл

$$\Phi(x) = \int_a^x f(t)dt$$

является первообразной. Пусть $F(x)$ – любая первообразная для $f(x)$. Тогда

$$\Phi(x) = F(x) + C.$$

Постоянную C найдем, положив $x = a$ (очевидно, $\Phi(a) = 0$):

$$0 = \Phi(a) = F(a) + C, \quad \text{откуда} \quad C = -F(a).$$

Тогда

$$\Phi(x) = F(x) - F(a).$$

При $x = b$ получаем $\Phi(b) = F(b) - F(a)$, т.е.

$$\int_a^b f(t)dt = F(b) - F(a),$$

или, что то же самое,

$$\int_a^b f(x)dx = F(b) - F(a). \tag{22.14}$$

Формула (22.14) называется **формулой Ньютона – Лейбница**. Это *основная формула интегрального исчисления*.

Разность $F(b) - F(a)$ записывают символом $F(x)\Big|_a^b$ («двойная подстановка от a до b»). Тогда формула (22.14) принимает вид

$$\int_a^b f(x)dx = F(x)\Big|_a^b. \tag{22.15}$$

***Пример* 22.1.**

1) $\displaystyle\int_0^3 x^2 dx = \frac{x^3}{3}\Big|_0^3 = \frac{3^3}{3} - \frac{0^3}{3} = 9;$

2) $\int\limits_1^{e} \frac{dx}{x} = \ln x \Big|_1^{e} = \ln e - \ln 1 = 1;$

3) $\int\limits_0^1 \frac{dx}{1+x^2} = \operatorname{arctg} x \Big|_0^1 = \operatorname{arctg} 1 - \operatorname{arctg} 0 = \frac{\pi}{4}.$

22.4. ЗАМЕНА ПЕРЕМЕННОЙ И ИНТЕГРИРОВАНИЕ ПО ЧАСТЯМ В ОПРЕДЕЛЕННОМ ИНТЕГРАЛЕ

Замена переменной

Пусть дан интеграл $\int\limits_a^b f(x)dx$, где $f(x)$ – непрерывная на отрезке $[a, b]$ функция. Введем новую переменную, положив $x = \varphi(t)$.

Пусть $\varphi(\alpha) = a$, $\varphi(\beta) = b$ и значения $\varphi(t)$ не выходят за пределы $[a, b]$, когда t изменяется на $[\alpha, \beta]$. Пусть, кроме того, $\varphi(t)$ и $\varphi'(t)$ непрерывны на $[\alpha, \beta]$. Тогда имеет место формула

$$\int\limits_a^b f(x)dx = \int\limits_\alpha^\beta f(\varphi(t))\, \varphi'(t)\, dt. \tag{22.16}$$

Справедливость этой формулы следует из (22.14). Действительно, если $F'(x) = f(x)$, то одновременно

$$\int\limits_a^b f(x)dx = F(x)\Big|_a^b = F(b) - F(a),$$

$$\int\limits_\alpha^\beta f(\varphi(t))\, \varphi'(t)\, dt = F(\varphi(t))\Big|_\alpha^\beta = F(\varphi(\beta)) - F(\varphi(\alpha)) = F(b) - F(a),$$

откуда и вытекает доказываемое равенство (22.16).

Следует заметить, что при вычислении определенного интеграла с помощью замены переменной нет необходимости возвращаться к старой переменной, т.е. делать обратную замену.

***Пример* 22.2.** Вычислить $\int\limits_{1/2}^1 \frac{\sqrt{1-x^2}}{x^2} dx.$

Решение. Применим подстановку $x = \sin t$. Тогда $dx = \cos t\, dt$, $t = \arcsin x$, $t = \frac{\pi}{6}$ при $x = \frac{1}{2}$ и $t = \frac{\pi}{2}$ при $x = 1$. Получаем

$$\int_{1/2}^{1}\frac{\sqrt{1-x^2}}{x^2}dx=\int_{\pi/6}^{\pi/2}\frac{\sqrt{1-\sin^2 t}}{\sin^2 t}\cos t\,dt=\int_{\pi/6}^{\pi/2}\frac{\cos^2 t}{\sin^2 t}dt=\int_{\pi/6}^{\pi/2}\frac{1-\sin^2 t}{\sin^2 t}dt=$$

$$=(-\operatorname{ctg} t-t)\Big|_{\pi/6}^{\pi/2}=-\frac{\pi}{2}+\sqrt{3}+\frac{\pi}{6}=\sqrt{3}-\frac{\pi}{3}.$$

Интегрирование по частям

Пусть функции $u=u(x)$ и $v=v(x)$ имеют непрерывные производные на отрезке $[a, b]$. Тогда

$$(uv)'=u'v+uv'.$$

Интегрируем обе части этого равенства:

$$\int_a^b (uv)'dx=\int_a^b u'v\,dx+\int_a^b uv'dx.$$

Так как $\int (uv)'dx=uv+C$, то $\int_a^b (uv)'dx=uv\Big|_a^b$. Получаем

$$uv\Big|_a^b=\int_a^b vdu+\int_a^b udv.$$

Следовательно,

$$\int_a^b udv=uv\Big|_a^b-\int_a^b vdu. \tag{22.17}$$

Равенство (22.17) называется **формулой интегрирования по частям в определенном интеграле**.

***Пример* 22.3.** Вычислить $\int_1^e x\ln x\,dx$.

Решение. Положим $u=\ln x$, $dv=xdx$. Тогда $du=\frac{dx}{x}$, $v=\frac{x^2}{2}$. Получаем

$$\int_1^e x\ln x\,dx=\frac{x^2}{2}\ln x\Big|_1^e-\int_1^e\frac{x^2}{2}\frac{dx}{x}=\frac{e^2}{2}-\frac{1}{4}x^2\Big|_1^e=\frac{e^2}{2}-\frac{e^2}{4}+\frac{1}{4}=\frac{e^2+1}{4}.$$

22.5. ПРИБЛИЖЕННОЕ ВЫЧИСЛЕНИЕ ОПРЕДЕЛЕННЫХ ИНТЕГРАЛОВ

При решении ряда прикладных задач часто приходится иметь дело с определенными интегралами от функций, для которых первообразные не являются элементарными функциями. Для вычисления таких интегралов существуют различные методы п р и б л и ж е н н о г о вычисления. Приведем наиболее простые из них: *формулу прямоугольников*, *формулу трапеций* и *формулу Симпсона*.

Пусть на отрезке $[a, b]$ задана непрерывная функция $y = f(x)$. Требуется вычислить определенный интеграл $\int\limits_a^b f(x)\,dx$.

1. Формула прямоугольников. Разделим отрезок $[a, b]$ точками $a = x_0, x_1, x_2, ..., x_n = b$ на n р а в н ы х частей длиной Δx:

$$\Delta x = \frac{b-a}{n}.$$

Обозначим: $y_0 = f(x_0), y_1 = f(x_1), ..., y_n = f(x_n)$.

Составим суммы

$$y_0\Delta x + y_1\Delta x + ... + y_{n-1}\Delta x,$$

$$y_1\Delta x + y_2\Delta x + ... + y_n\Delta x,$$

каждая из которых является интегральной суммой для $f(x)$ на $[a, b]$. Поэтому

$$\int\limits_a^b f(x)\,dx \approx \frac{b-a}{n}(y_0 + y_1 + ... + y_{n-1}), \qquad (22.18)$$

$$\int\limits_a^b f(x)\,dx \approx \frac{b-a}{n}(y_1 + y_2 + ... + y_n). \qquad (22.18')$$

Каждая из формул (22.18) и (22.18′) называется **формулой прямоугольников**.

Ошибка, совершаемая при вычислении интеграла по формуле прямоугольников, будет тем меньше, чем больше число n, т.е. чем меньше частичные отрезки, на которые разбивается отрезок $[a, b]$.

2. Формула трапеций:

$$\int\limits_a^b f(x)\,dx \approx \frac{b-a}{n}\left(\frac{y_0 + y_n}{2} + y_1 + y_2 + ... + y_{n-1}\right). \qquad (22.19)$$

3. Формула Симпсона. Разделим отрезок $[a, b]$ на ч е т н о е число равных частей $n = 2m$:

$$\int\limits_a^b f(x)\,dx \approx \frac{b-a}{6m}[y_0 + y_{2m} + 2(y_2 + y_4 + ... + y_{2m-2}) +$$

$$+ 4(y_1 + y_3 + ... + y_{2m-1})]. \qquad (22.20)$$

Заметим, что при одном и том же шаге деления отрезка $\Delta x = \dfrac{b-a}{n}$ формула трапеций дает несколько более точное значение определенного интеграла, чем формула прямоугольников, а формула Симпсона – значительно более точное значение, чем формула трапеций.

***Пример* 22.4.** Рассмотрим известный интеграл

$$\int_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4} = 0{,}785398\ldots.$$

Разобьем отрезок [0, 1] на четыре равные части: $x_0 = 0$; $x_1 = 0{,}25$; $x_2 = 0{,}5$; $x_3 = 0{,}75$; $x_4 = 1$. Тогда $y_0 = 1{,}0000$; $y_1 \approx 0{,}9412$; $y_2 = 0{,}8000$; $y_3 = 0{,}6400$; $y_4 = 0{,}5000$.

По формулам прямоугольников имеем:

$$\frac{1}{4}(1+0{,}9412+0{,}8+0{,}64) = 0{,}8453,$$

$$\frac{1}{4}(0{,}9412+0{,}8+0{,}64+0{,}5) = 0{,}7203,$$

по формуле трапеций

$$\frac{1}{4}\left(\frac{1+0{,}5}{2}+0{,}9412+0{,}8+0{,}64\right) = 0{,}7828,$$

по формуле Симпсона

$$\frac{1}{12}(1+0{,}5+3{,}76471+1{,}6+2{,}56) = 0{,}78539.$$

(Мы взяли с точностью до 0,00001 значение $4y_1 \approx 3{,}76471$.)

Таким образом, формула Симпсона дает весьма точный результат: все пять знаков верны. Формула трапеций дает ошибку уже в третьем знаке. Если бы мы разделили отрезок [0, 1] на 10 частей, то формула трапеций дала бы результат, отличающийся от истинного значения менее чем на 0,0005. Для того же, чтобы получить удовлетворительный результат с помощью формулы прямоугольников, необходимо разбить отрезок на значительно большее число частей.

Вообще, для того чтобы знать, сколько точек деления надо взять для вычисления интеграла с заданной степенью точности, нужно воспользоваться формулами оценки погрешности. Эти оценки можно найти в более подробных курсах математического анализа.

Вопросы

1. Что такое интегральная сумма для данной функции на данном отрезке?
2. Что называется определенным интегралом от функции $f(x)$ на отрезке $[a, b]$?
3. В чем различия понятий определенного и неопределенного интегралов?
4. В чем заключается геометрический смысл определенного интеграла?
5. Каков экономический смысл определенного интеграла?
6. Всякая ли интегрируемая функция дифференцируема? Всякая ли дифференцируемая функция интегрируема?
7. Чему равна производная от определенного интеграла по его переменному верхнему пределу?
8. Чем отличается применение метода замены переменной при вычислении определенного интеграла от применения этого метода при вычислении неопределенного интеграла?
9. Какие методы применяются для вычисления определенных интегралов от функций, не имеющих первообразных, которые выражались бы через элементарные функции?
10. Какая из приближенных формул дает более точное значение определенного интеграла при одном и том же шаге деления отрезка интегрирования: формула прямоугольников, формула трапеций или формула Симпсона?

Глава 23
ПРИЛОЖЕНИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

23.1. ГЕОМЕТРИЧЕСКИЕ И МЕХАНИЧЕСКИЕ ПРИЛОЖЕНИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

Площадь плоской фигуры

Как отмечалось ранее, для непрерывной на отрезке $[a, b]$ функции $f(x) \geq 0$ площадь S криволинейной трапеции, ограниченной линиями $y = f(x)$, $y = 0$, $x = a$, $x = b$ (см. рис. 22.2), выражается интегралом

$$S = \int_a^b f(x)\,dx. \tag{23.1}$$

Пример **23.1.** Вычислить площадь фигуры, ограниченной графиком функции $y = \ln x$, осью Ox и прямыми $x = \mathrm{e}$, $x = \mathrm{e}^2$ (рис. 23.1).

Решение. Согласно формуле (23.1) искомая площадь

$$S = \int_{\mathrm{e}}^{\mathrm{e}^2} \ln x\,dx = x(\ln x - 1)\Big|_{\mathrm{e}}^{\mathrm{e}^2} = \mathrm{e}^2.$$

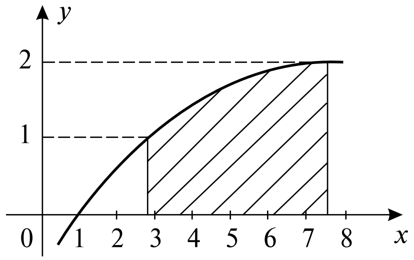


Рис. 23.1

Пусть $y = f_1(x)$, $y = f_2(x)$ – непрерывные на $[a, b]$ функции и пусть $f_2(x) \geq f_1(x)$ на указанном отрезке. Тогда площадь S фигуры, ограниченной графиками функций $y=f_1(x)$, $y=f_2(x)$ и вертикальными прямыми $x=a$, $x=b$, вычисляется по формуле

$$S=\int_a^b (f_2(x)-f_1(x))\,dx. \tag{23.2}$$

Д о к а з а т е л ь с т в о. 1. Пусть $f_1(x)\ge 0$, $f_2(x)\ge 0$. Тогда формула (23.2) является очевидным следствием того, что площадь фигуры равна разности площадей криволинейных трапеций (рис. 23.2):

$$S=\int_a^b f_2(x)\,dx-\int_a^b f_1(x)\,dx=\int_a^b (f_2(x)-f_1(x))\,dx.$$

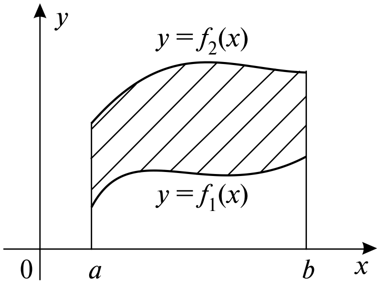


Рис. 23.2. Площадь фигуры, ограниченной линиями $y=f_1(x)$, $y=f_2(x)$, $x=a$, $x=b$

2. Пусть графики функций $y=f_1(x)$, $y=f_2(x)$ полностью или частично расположены ниже оси Ox. Так как эти функции ограничены, то существует такое число M, что $f_1(x)+M\ge 0$, $f_2(x)+M\ge 0$. Очевидно, фигура, ограниченная линиями $y=f_1(x)+M$, $y=f_2(x)+M$ (находящимися над Ox), получена параллельным переносом фигуры, ограниченной линиями $y=f_1(x)$, $y=f_2(x)$, и имеет ту же площадь:

$$S=\int_a^b (f_2(x)+M)\,dx-\int_a^b (f_1(x)+M)\,dx=\int_a^b (f_2(x)-f_1(x))\,dx.$$

Пример **23.2.** Вычислить площадь фигуры, ограниченной линиями $y=x^2-2x$, $y=4x-x^2$.

Р е ш е н и е. Приравняв правые части этих уравнений, находим абсциссы точек пересечения указанных кривых: $x_1=0$, $x_2=3$. Следовательно,

$$S=\int_0^3 (4x-x^2-(x^2-2x))\,dx=\int_0^3 (6x-2x^2)\,dx=\left(3x^2-\frac{2}{3}x^3\right)\Bigg|_0^3=9.$$

Объем тела вращения

Рассмотрим тело, которое образовано вращением вокруг оси Ox криволинейной трапеции, ограниченной графиком непрерывной на отрезке $[a, b]$ неотрицательной функции $y = f(x)$ и линиями $y = 0$, $x = a$, $x = b$. Объем этого тела выражается формулой

$$V = \pi \int_a^b f^2(x)\,dx. \tag{23.3}$$

Докажем это. Разобьем отрезок $[a, b]$ произвольным образом на n частей точками $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ и на каждом отрезке $[x_{i-1}, x_i]$ произвольно выберем точку ξ_i. При вращении кривой $y = f(x)$ вокруг оси Ox каждый прямоугольник с основанием $\Delta x_i = x_i - x_{i-1}$ и высотой $f(\xi_i)$ описывает цилиндр радиусом $f(\xi_i)$ и высотой Δx_i (рис. 23.3).

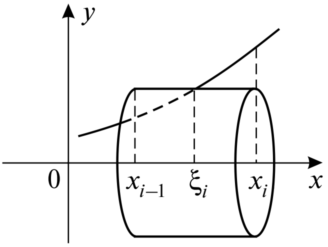


Рис. 23.3. Вычисление объема тела вращения

Сумма объемов таких цилиндров имеет вид

$$V_n = \sum_{i=1}^{n} \pi f^2(\xi_i)\Delta x_i. \tag{23.4}$$

При более мелком разбиении отрезка эта сумма дает приближенное значение искомого объема. С другой стороны, эта сумма есть интегральная сумма для непрерывной функции $y = \pi f^2(x)$. Переходя к пределу при $\max \Delta x_i \to 0$, получаем формулу (23.3).

***Пример* 23.3.** Вычислить объем конуса радиусом R и высотой H.

Решение. Конус, объем которого мы вычисляем, можно получить в результате вращения треугольника OAB вокруг оси Ox. Здесь $AB = R$, $OB = H$ (рис. 23.4).

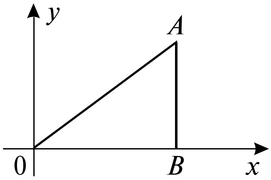


Рис. 23.4

Очевидно, $\operatorname{tg} AOB = \dfrac{R}{H}$. Поэтому уравнение прямой OA имеет вид $y = \dfrac{R}{H}x$. По формуле (23.3) получаем

$$V = \pi \int_0^H \left(\frac{R}{H} x \right)^2 dx = \frac{\pi R^2}{H^2} \int_0^H x^2 dx = \frac{\pi R^2}{3H^2} \cdot x^3 \Big|_0^H = \frac{\pi R^2 H}{3}.$$

Эта формула хорошо известна из школьного курса геометрии.

Пример **23.4.** Вычислить объем тела, образуемого вращением фигуры, ограниченной линиями $y = \mathrm{e}^x$, $y = 0$, $x = -1$, $x = 0$, вокруг оси Ox (рис. 23.5).

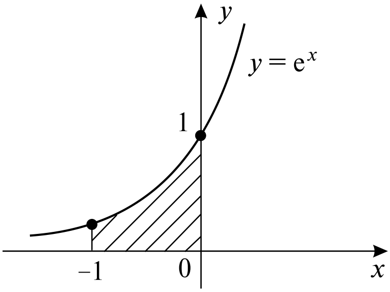


Рис. 23.5

Решение: $V = \pi \int_{-1}^{0} (\mathrm{e}^x)^2 dx = \pi \cdot \dfrac{1}{2} \mathrm{e}^{2x} \Big|_{-1}^{0} = \dfrac{\pi}{2}\left(1 - \dfrac{1}{\mathrm{e}^2}\right).$

Длина дуги плоской кривой

Пусть кривая на плоскости Oxy задана уравнением $y = f(x)$ и $f(x)$ имеет непрерывную производную $f'(x)$ на отрезке $[a, b]$. Тогда длина l дуги этой кривой равна

$$l = \int_a^b \sqrt{1 + (y')^2}\, dx, \tag{23.5}$$

где a и b – абсциссы концов дуги.

(Примем это утверждение без доказательства.)

Пример **23.5.** Найти длину дуги кривой $y = \ln\sin x$ от $x = \frac{\pi}{3}$ до $x = \frac{2\pi}{3}$.

Решение. Вычисляем производную $y' = \operatorname{ctg} x$ и подставляем ее в формулу (23.5):

$$l = \int\limits_{\pi/3}^{2\pi/3} \sqrt{1+\operatorname{ctg}^2 x}\, dx = \int\limits_{\pi/3}^{2\pi/3} \frac{dx}{\sin x} = \ln \operatorname{tg}\frac{x}{2}\Bigg|_{\frac{\pi}{3}}^{\frac{2\pi}{3}} = \ln 3.$$

Механические и физические приложения интеграла

Эти приложения для экономистов малоинтересны, поэтому бегло упомянем некоторые из них.

Хорошо известно, что пройденный путь есть интеграл от скорости движения. Механическая работа также вычисляется с помощью интеграла. Пусть под воздействием некоторой силы F материальная точка движется по прямой Os, причем направление силы совпадает с направлением движения. Требуется найти работу силы F по перемещению точки из положения $s = a$ в положение $s = b$. Если сила F постоянна, то работа A равна произведению силы F на длину пути: $A = F \cdot (b - a)$. Если же сила непрерывно изменяется, т.е. $F = F(s)$ есть непрерывная функция на $[a, b]$, то работа A выражается формулой

$$A = \int\limits_a^b F(s)\,ds.$$

23.2. ПРИЛОЖЕНИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА В ЭКОНОМИКЕ

В § 22.1 мы отмечали, что, зная функцию производительности труда, можно с помощью определенного интеграла выразить объем произведенной продукции. Рассмотрим пример.

Пример **23.6.** Найти дневную выработку P за рабочий день с 8 до 14 часов, если производительность труда задана эмпирической формулой

$$P = P(t) = -\frac{t^2}{4} + 5t - 15.$$

(Эта формула отражает процесс работы, при котором производительность растет на протяжении первых двух часов, а затем падает.)

Р е ш е н и е. Рассматривая функцию производительности (рис. 23.6) на отрезке [8, 14], выразим дневную выработку интегралом:

$$P=\int_8^{14}\left(-\frac{t^2}{4}+5t-15\right)dt=\left(-\frac{t^3}{12}+\frac{5t^2}{2}-15t\right)\Bigg|_8^{14}=54.$$

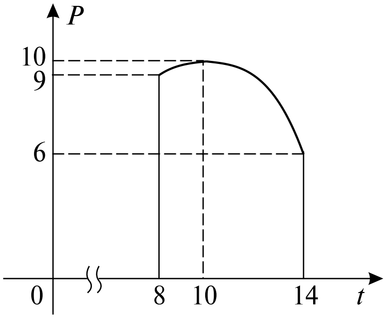


Рис. 23.6

Итак, за указанный промежуток времени произведено 54 единицы продукции.

В § 20.1 мы рассматривали, в частности, предельные издержки, задаваемые производной функции издержек $S(x)$: $MS=S'(x)$. Эта производная характеризует затраты на производство единицы дополнительной продукции.

Рассмотрим задачу нахождения функции издержек по данной функции предельных издержек.

П р и м е р **23.7.** Задана функция предельных издержек $MS=3x^2-40x+125$, $x\in[0, 30]$. Найти функцию издержек $S(x)$ и вычислить издержки при производстве 20 единиц продукции, если известно, что издержки для производства первой единицы продукции составляют 100 денежных единиц.

Р е ш е н и е. Функцию издержек находим интегрированием:

$$S(x)=\int_1^x MS\,dx+C.$$

При этом константа C определяется условием $S(1)=100$, так что $C=100$, так как интеграл обращается в нуль.

Получаем функцию издержек:

$$S(x)=x^3-20x^2+125x+100.$$

Подставляя $x=20$, находим

$$S(20)=2600.$$

Еще одним примером приложения определенного интеграла является дисконтирование. *Дисконтированием* называется определение начальной денежной суммы S по ее конечной величине S_t через время t при процентной ставке p. Задачи дисконтирования встречаются при определении экономической эффективности капитальных вложений.

Как было установлено ранее (см. § 14.5), при непрерывном начислении процентов конечная сумма рассчитывается по формуле $S_t = Se^{rt}$, где $r = \frac{p}{100}$. Отсюда $S = S_t e^{-rt}$. Если $S_t = f(t)$, то дисконтированная сумма к моменту времени t составит $S = f(t)e^{-rt}$.

Полная дисконтированная сумма S_d за время T выражается формулой

$$S_d = \int_0^T f(t)e^{-rt}dt. \tag{23.6}$$

Вопросы

1. Как выражается площадь плоской фигуры с помощью определенного интеграла?
2. Объем каких тел и как можно вычислять с помощью определенных интегралов?
3. Можно ли, зная функцию производительности труда, с помощью определенного интеграла выразить объем произведенной продукции?
4. Как выражается полная дисконтированная сумма при непрерывном начислении процентов?

Глава 24
НЕСОБСТВЕННЫЕ ИНТЕГРАЛЫ

24.1. НЕСОБСТВЕННЫЕ ИНТЕГРАЛЫ С БЕСКОНЕЧНЫМИ ПРЕДЕЛАМИ ИНТЕГРИРОВАНИЯ

Вводя понятие определенного интеграла, мы предполагали, что отрезок интегрирования конечен, а подынтегральная функция ограничена на этом отрезке. Теперь рассмотрим случаи, когда хотя бы одно из этих условий не выполняется.

Определение. Пусть функция $y=f(x)$ определена на бесконечном промежутке $[a, +\infty)$ и интегрируема на любом конечном отрезке $[a, b]$, $b > a$, т.е. существует для любого $b > a$ определенный интеграл

$$\Phi(b)=\int_a^b f(x)\,dx.$$

Тогда **несобственным интегралом** $\int\limits_a^{+\infty} f(x)\,dx$ от функции $f(x)$ на промежутке $[a, +\infty)$ называется предел

$$\lim_{b\to+\infty}\Phi(b)=\lim_{b\to+\infty}\int_a^b f(x)\,dx.$$

Следовательно, по определению

$$\int_a^{+\infty} f(x)\,dx=\lim_{b\to+\infty}\int_a^b f(x)\,dx. \tag{24.1}$$

Если предел, стоящий в правой части равенства (24.1), существует и конечен, то несобственный интеграл называется **сходящимся**, в противном случае – **расходящимся**. Если несобственный интеграл (24.1) сходится (т.е. является сходящимся), то функцию $f(x)$ называют **интегрируемой** на $[a, +\infty)$.

***Пример* 24.1.** Вычислить интегралы: а) $\int\limits_1^{+\infty}\frac{dx}{x^2}$; б) $\int\limits_1^{+\infty}\frac{dx}{\sqrt{x}}$.

Решение.

$$\text{а) } \int\limits_1^{+\infty}\frac{dx}{x^2}=\lim_{b\to+\infty}\int\limits_1^{b}x^{-2}dx=\lim_{b\to+\infty}\left(-\frac{1}{x}\right)\Bigg|_1^b=\lim_{b\to+\infty}\left(-\frac{1}{b}+1\right)=1,$$

т.е. интеграл сходится к 1.

$$\text{б) } \int\limits_1^{+\infty}\frac{dx}{\sqrt{x}}=\lim_{b\to+\infty}\int\limits_1^{b}x^{-\frac{1}{2}}dx=2\lim_{b\to+\infty}\left(\sqrt{b}-1\right)=\infty,$$

т.е. интеграл расходится.

***Пример* 24.2.** Установить, при каких значениях α интеграл $\int\limits_1^{+\infty}\frac{dx}{x^\alpha}$ сходится, а при каких расходится.

Решение. При $\alpha\neq 1$ имеем

$$\int\limits_1^{b}\frac{dx}{x^\alpha}=\frac{1}{1-\alpha}x^{1-\alpha}\Big|_1^b=\frac{1}{1-\alpha}\left(b^{1-\alpha}-1\right),$$

при $\alpha = 1$

$$\int\limits_1^{b}\frac{dx}{x}=\ln x\Big|_1^b=\ln b.$$

Поэтому при $\alpha\neq 1$

$$\int\limits_1^{+\infty}\frac{dx}{x^\alpha}=\lim_{b\to+\infty}\frac{1}{1-\alpha}\left(b^{1-\alpha}-1\right),$$

при $\alpha = 1$

$$\int\limits_1^{+\infty}\frac{dx}{x}=\lim_{b\to+\infty}\ln b.$$

Итак, можно сделать следующие выводы:

если $\alpha>1$, то $\int\limits_1^{+\infty}\frac{dx}{x^\alpha}=\frac{1}{\alpha-1}$ – интеграл сходится;

если $\alpha<1$, то $\int\limits_1^{+\infty}\frac{dx}{x^\alpha}=\infty$ – интеграл расходится;

если $\alpha=1$, то $\int\limits_1^{+\infty}\frac{dx}{x}=\infty$ – интеграл расходится.

По аналогии с несобственным интегралом, определяемым равенством (24.1), определяется *несобственный интеграл с бесконечным нижним пределом*, а именно:

$$\int_{-\infty}^{b} f(x)\,dx = \lim_{a\to-\infty}\int_{a}^{b} f(x)\,dx. \tag{24.2}$$

Наконец, можно рассмотреть *несобственный интеграл с бесконечными нижним и верхним пределами*:

$$\int_{-\infty}^{+\infty} f(x)\,dx.$$

Для этого возьмем произвольную точку c. Она разобьет числовую прямую на две полупрямые. Если существуют несобственные интегралы $\int_{-\infty}^{c} f(x)\,dx$ и $\int_{c}^{+\infty} f(x)\,dx$, то по определению существует и несобственный интеграл $\int_{-\infty}^{+\infty} f(x)\,dx$. В этом случае полагают

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{+\infty} f(x)\,dx. \tag{24.3}$$

Можно доказать, что правая часть равенства (24.3) не зависит от выбора точки c.

***Пример* 24.3.** Вычислить $\int_{-\infty}^{+\infty}\frac{dx}{1+x^2}$.

Решение: $\int_{-\infty}^{+\infty}\frac{dx}{1+x^2} = \int_{-\infty}^{0}\frac{dx}{1+x^2} + \int_{0}^{+\infty}\frac{dx}{1+x^2}$.

Вычислим каждый из интегралов в правой части последнего равенства:

$$\int_{-\infty}^{0}\frac{dx}{1+x^2} = \lim_{a\to-\infty}\int_{a}^{0}\frac{dx}{1+x^2} = \lim_{a\to-\infty}\operatorname{arctg} x\Big|_{a}^{0} = \lim_{a\to-\infty}(\operatorname{arctg}0-\operatorname{arctg}a) = \frac{\pi}{2},$$

$$\int_{0}^{+\infty}\frac{dx}{1+x^2} = \lim_{b\to+\infty}\int_{0}^{b}\frac{dx}{1+x^2} = \lim_{b\to+\infty}\operatorname{arctg} x\Big|_{0}^{b} = \lim_{b\to+\infty}(\operatorname{arctg}b-\operatorname{arctg}0) = \frac{\pi}{2}.$$

Следовательно,

$$\int_{-\infty}^{+\infty}\frac{dx}{1+x^2} = \frac{\pi}{2}+\frac{\pi}{2} = \pi.$$

Несобственные интегралы с бесконечными пределами, т.е. интегралы (24.1), (24.2) и (24.3), называются **несобственными интегралами первого рода**.

24.2. НЕСОБСТВЕННЫЕ ИНТЕГРАЛЫ ОТ НЕОГРАНИЧЕННЫХ ФУНКЦИЙ

Пусть функция $y = f(x)$ определена на промежутке $[a, b)$. Точку $x = b$ будем называть **особой**, если $f(x)$ не ограничена в любой окрестности этой точки, но ограничена и интегрируема на любом отрезке, заключенном в промежутке $[a, b)$.

Определение. Если $f(x)$ не ограничена на $[a, b)$, но является интегрируемой на любом отрезке $[a, b-\varepsilon]$, $0 < \varepsilon < b-a$, то **несобственным интегралом** $\int\limits_a^b f(x)dx$ от функции $f(x)$ на $[a, b]$ называется предел

$$\lim_{\varepsilon \to 0} \int\limits_a^{b-\varepsilon} f(x)dx. \tag{24.4}$$

Таким образом,

$$\int\limits_a^b f(x)dx = \lim_{\varepsilon \to 0} \int\limits_a^{b-\varepsilon} f(x)dx. \tag{24.5}$$

Если предел (24.4) существует и конечен, то интеграл (24.5) называется **сходящимся**, в противном случае он называется **расходящимся**.

Аналогично определяется несобственный интеграл в случае, когда особой точкой является л е в ы й конец промежутка:

$$\int\limits_a^b f(x)dx = \lim_{\varepsilon \to 0} \int\limits_{a+\varepsilon}^{b} f(x)dx. \tag{24.6}$$

Наконец, если c – единственная в н у т р е н н я я особая точка на $[a, b]$, то по определению

$$\int\limits_a^b f(x)dx = \int\limits_a^c f(x)dx + \int\limits_c^b f(x)dx. \tag{24.7}$$

Несобственные интегралы от неограниченных функций называются **несобственными интегралами второго рода**.

Пример **24.4.** Вычислить несобственный интеграл $\int\limits_0^1 \frac{dx}{\sqrt{1-x}}$.

Решение. По формуле (24.5) получаем

$$\int\limits_0^1 \frac{dx}{\sqrt{1-x}} = \lim_{\varepsilon\to 0} \int\limits_0^{1-\varepsilon} \frac{dx}{\sqrt{1-x}} = -2\lim_{\varepsilon\to 0}\sqrt{1-x}\Big|_0^{1-\varepsilon} = -2\lim_{\varepsilon\to 0}\left(\sqrt{\varepsilon}-1\right) = 2.$$

Пример **24.5.** Вычислить несобственный интеграл $\int\limits_0^1 \frac{dx}{\sqrt{x}}$.

Решение: $\int\limits_0^1 \frac{dx}{\sqrt{x}} = \lim\limits_{\varepsilon\to 0}\int\limits_\varepsilon^1 \frac{dx}{\sqrt{x}} = \lim\limits_{\varepsilon\to 0} 2\sqrt{x}\Big|_\varepsilon^1 = \lim\limits_{\varepsilon\to 0}\left(2-2\sqrt{\varepsilon}\right) = 2.$

Заметим, что несобственный интеграл $\int\limits_0^1 \frac{dx}{x^\alpha}$, где $\alpha > 0$, сходится при $0 < \alpha < 1$ (и равен $\frac{1}{1-\alpha}$) и расходится при $\alpha \geq 1$. (Проверьте это самостоятельно.)

24.3. ПРИЗНАКИ СХОДИМОСТИ НЕСОБСТВЕННЫХ ИНТЕГРАЛОВ

При вычислении и исследовании несобственных интегралов значительное место занимает исследование их сходимости. Во многих случаях бывает достаточно установить, сходится данный интеграл или расходится, и оценить его значение. Для исследования сходимости применяются, в частности, признаки сравнения, основанные на сопоставлении заданного интеграла с интегралом, сходимость которого известна.

Примем без доказательства следующее утверждение (**признак сравнения**).

Теорема **24.1.** Пусть функции $f(x)$ и $g(x)$ непрерывны на промежутке $[a, +\infty)$ и удовлетворяют на нем условию $0 \leq f(x) \leq g(x)$. Тогда:

1) из сходимости $\int\limits_a^{+\infty} g(x)\,dx$ следует сходимость $\int\limits_a^{+\infty} f(x)\,dx$;

2) из расходимости $\int\limits_a^{+\infty} f(x)\,dx$ следует расходимость $\int\limits_a^{+\infty} g(x)\,dx$.

***Пример* 24.6.** Исследовать сходимость интеграла $\int\limits_1^{+\infty} \frac{\sqrt{x^3}+1}{x^2}dx$.

Решение. Очевидно,

$$\frac{\sqrt{x^3}+1}{x^2} > \frac{\sqrt{x^3}}{x^2} = \frac{1}{\sqrt{x}}.$$

Но интеграл $\int\limits_1^{+\infty} \frac{dx}{\sqrt{x}}$ расходится (см. пример 24.1). Следовательно, расходится и данный интеграл.

***Пример* 24.7.** Исследовать сходимость интеграла $\int\limits_3^{+\infty} \frac{dx}{x\sqrt{x}\,(x-1)}$.

Решение. Сравним подынтегральную функцию $f(x) = \frac{1}{x\sqrt{x}\,(x-1)}$ с функцией $g(x) = \frac{1}{x\sqrt{x}}$ на $[3, +\infty)$. Очевидно, что

$$\frac{1}{x\sqrt{x}\,(x-1)} < \frac{1}{x\sqrt{x}}.$$

Но интеграл $\int\limits_3^{+\infty} \frac{dx}{x\sqrt{x}}$ сходится (см. пример 24.2 с учетом $\alpha = \frac{3}{2}$). Следовательно, сходится и данный интеграл.

Аналогичный **признак сравнения** имеет место и **для несобственных интегралов второго рода**: если функции $f(x)$ и $g(x)$ непрерывны на промежутке $[a, b)$ и для всех x в некоторой окрестности особой точки b выполняются условия $0 \le f(x) \le g(x)$, то:

1) из сходимости $\int\limits_a^b g(x)\,dx$ следует сходимость $\int\limits_a^b f(x)\,dx$;

2) из расходимости $\int\limits_a^b f(x)\,dx$ следует расходимость $\int\limits_a^b g(x)\,dx$.

***Пример* 24.8.** Исследовать сходимость интеграла $\int\limits_0^1 \frac{dx}{\sqrt{x+3x^2}}$.

Решение. Особой точкой здесь является $x = 0$. Сравним подынтегральную функцию с функцией $\frac{1}{\sqrt{x}}$:

$$\frac{1}{\sqrt{x+3x^2}} < \frac{1}{\sqrt{x}}.$$

Несобственный интеграл $\int\limits_0^1 \frac{dx}{\sqrt{x}}$ сходится (см. пример 24.5). Следовательно, данный интеграл сходится.

Приведем без доказательства еще один **признак сравнения**, также часто применяемый на практике.

***Теорема* 24.2.** Если $f(x)$ и $g(x)$ – неотрицательные функции и существует конечный предел $\lim\limits_{x\to+\infty} \frac{f(x)}{g(x)} = A \neq 0$, то несобственные интегралы $\int\limits_a^{+\infty} f(x)\,dx$ и $\int\limits_a^{+\infty} g(x)\,dx$ сходятся или расходятся одновременно.

***Пример* 24.9.** Несобственный интеграл $\int\limits_3^{+\infty} \frac{dx}{x^2-2x}$ сходится. Действительно, предел отношения функции $f(x) = \frac{1}{x^2-2x}$ к функции $g(x) = \frac{1}{x^2}$ конечен: $\lim\limits_{x\to+\infty} \frac{f(x)}{g(x)} = \lim\limits_{x\to+\infty} \frac{x^2}{x^2-2x} = 1$, а интеграл $\int\limits_3^{+\infty} \frac{dx}{x^2}$ сходится (см. пример 24.1).

Абсолютная и условная сходимость несобственных интегралов

Отметим важное свойство несобственных интегралов, отличающее их от обычных определенных интегралов.

Для определенного интеграла, как известно, справедливо следующее утверждение: если существует $\int\limits_a^b f(x)\,dx$, то существует $\int\limits_a^b |f(x)|\,dx$.

В случае же несобственных интегралов имеет место такое утверждение: если несобственный интеграл $\int\limits_a^{+\infty} |f(x)|\,dx$ сходится, то сходится и $\int\limits_a^{+\infty} f(x)\,dx$.

Примем это утверждение без доказательства. Обратное утверждение, вообще говоря, не верно: сходимость интеграла $\int\limits_a^{+\infty} f(x)\,dx$ не влечет сходимость интеграла $\int\limits_a^{+\infty} |f(x)|\,dx$.

Если несобственный интеграл $\int\limits_a^{+\infty} |f(x)|\,dx$ сходится, то говорят, что **интеграл** $\int\limits_a^{+\infty} f(x)\,dx$ **сходится абсолютно**.

Если же интеграл $\int\limits_a^{+\infty} f(x)\,dx$ сходится, а интеграл $\int\limits_a^{+\infty} |f(x)|\,dx$ расходится, то говорят, что **интеграл** $\int\limits_a^{+\infty} f(x)\,dx$ **сходится условно**.

Таким образом, из абсолютной сходимости интеграла следует его сходимость (в обычном смысле): абсолютно сходящийся интеграл сходится.

Пример **24.10.** Исследовать сходимость интеграла $\int\limits_1^{+\infty} \frac{\cos x}{x^2}\,dx$.

Решение. Очевидно, $\left|\frac{\cos x}{x^2}\right| \le \frac{1}{x^2}$. Мы знаем, что интеграл $\int\limits_1^{+\infty} \frac{dx}{x^2}$ сходится (см. пример 24.1). Следовательно, сходится $\int\limits_1^{+\infty} \left|\frac{\cos x}{x^2}\right| dx$, т.е. данный интеграл $\int\limits_1^{+\infty} \frac{\cos x}{x^2}\,dx$ сходится абсолютно. Отсюда следует, что он сходится.

Вопросы

1. Как определяется несобственный интеграл от функции на бесконечном полуинтервале вида $[a, +\infty)$? В каком случае интеграл называется сходящимся, а в каком – расходящимся?
2. Как определяется несобственный интеграл от функции на бесконечном полуинтервале вида $(-\infty, b]$?
3. Как определяется несобственный интеграл от функции на бесконечном интервале $(-\infty, +\infty)$?
4. Какие интегралы называются несобственными интегралами первого рода?

5. Как определяется несобственный интеграл от неограниченной функции в случае, когда особая точка является одним из концов отрезка интегрирования?
6. Как определяется несобственный интеграл от неограниченной функции, когда особая точка является внутренней точкой отрезка интегрирования?
7. Каковы признаки сравнения для несобственных интегралов первого и второго рода?
8. Всегда ли сходится абсолютно сходящийся несобственный интеграл?
9. Какой несобственный интеграл называется условно сходящимся?

Раздел VII
ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Глава 25
ЕВКЛИДОВО ПРОСТРАНСТВО. ПОНЯТИЕ ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ. ПРЕДЕЛ, НЕПРЕРЫВНОСТЬ

25.1. ЕВКЛИДОВО ПРОСТРАНСТВО

Определение. Множество всевозможных упорядоченных совокупностей n действительных чисел $(x_1, x_2, ..., x_n)$ называется **n-мерным координатным пространством** A^n.

При этом каждая упорядоченная совокупность $(x_1, x_2, ..., x_n)$ называется *точкой** этого пространства и обозначается одной буквой (например, M). Числа $x_1, x_2, ..., x_n$ называются *координатами* точки. Запись $M(x_1, x_2, ..., x_n)$ означает, что точка M имеет координаты $x_1, x_2, ..., x_n$.

Определение. Координатное пространство A^n называется **n-мерным евклидовым пространством $\mathbf{R}^n$**, если между любыми двумя точками $M(x_1, x_2, ..., x_n)$ и $M'(x'_1, x'_2, ..., x'_n)$ пространства A^n определено *расстояние* $\rho(M, M')$ по формуле

$$\rho(M, M') = \sqrt{(x'_1 - x_1)^2 + (x'_2 - x_2)^2 + ... + (x'_n - x_n)^2}. \quad (25.1)$$

Указанное расстояние удовлетворяет следующим условиям:

1) для любых M и M' выполняется равенство $\rho(M, M') = \rho(M', M)$;

* Не следует удивляться тому, что в начале книги мы назвали упорядоченный набор (упорядоченную совокупность) чисел $(x_1, x_2, ..., x_n)$ вектором, а здесь – точкой. Мы знаем, в частности, что пара (x, y) может рассматриваться и как пара координат точки на плоскости, и как пара координат вектора на этой плоскости.

2) для любых M и M' справедливо неравенство $\rho(M, M') \geq 0$, причем если $\rho(M, M') = 0$, то точки M и M' совпадают;

3) для любых M, M' и M'' выполняется неравенство $\rho(M, M'') \leq$ $\leq \rho(M, M') + \rho(M', M'')$.

Заметим, что мы уже дали определение n-мерного евклидова пространства $\mathbf{R}^n$ (см. § 1.6). Нетрудно убедиться в том, что расстояние, определенное равенством (25.1), есть норма вектора $\overline{MM'} = (x'_1 - x_1, x'_2 - x_2, ..., x'_n - x_n)$ и данное здесь определение пространства $\mathbf{R}^n$ по существу не отличается от определения, данного в главе 1.

25.2. МНОЖЕСТВА В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ

Рассмотрим простейшие множества точек, или области, в евклидовом пространстве.

1. Множество точек $M(x_1, x_2, ..., x_n)$, координаты которых удовлетворяют неравенству

$$\left(x_1 - x_1^0\right)^2 + \left(x_2 - x_2^0\right)^2 + ... + \left(x_n - x_n^0\right)^2 \leq r^2, \tag{25.2}$$

называется ***n*-мерным шаром** радиуса r с центром в точке $M_0(x_1^0, x_2^0, ..., x_n^0)$.

Неравенство (25.2) можно коротко записать в виде

$$\rho(M_0, M) \leq r. \tag{25.3}$$

Множество всех таких точек M, для которых выполняется неравенство $\rho(M_0, M) < r$, называется **открытым *n*-мерным шаром** радиуса r с центром в точке M_0. Множество всех таких точек M, для которых выполняется равенство $\rho(M_0, M) = r$, называется ***n*-мерной сферой**.

2. Множество точек $M(x_1, x_2, ..., x_n)$, координаты которых удовлетворяют неравенствам

$$\left|x_1 - x_1^0\right| \leq d_1,\ \left|x_2 - x_2^0\right| \leq d_2,\ ...,\ \left|x_n - x_n^0\right| \leq d_n, \tag{25.4}$$

называется ***n*-мерным параллелепипедом** с центром в точке $M_0(x_1^0, x_2^0, ..., x_n^0)$.

Если в соотношениях (25.4) исключить равенства:

$$\left|x_1 - x_1^0\right| < d_1,\ \left|x_2 - x_2^0\right| < d_2,\ ...,\ \left|x_n - x_n^0\right| < d_n,$$

то этим определится **открытый *n*-мерный параллелепипед**.

Перейдем к определению окрестностей точки. В пространстве $\mathbf{R}^n$ различают окрестности двух типов: *прямоугольные* и *сферические*.

Будем называть **ε-окрестностью** точки $M_0(x_1^0, x_2^0, ..., x_n^0)$ открытый n-мерный шар радиуса ε с центром в точке M_0. (Это так называемая сферическая окрестность.)

Прямоугольной окрестностью точки $M_0(x_1^0, x_2^0, ..., x_n^0)$ называется любой открытый n-мерный параллелепипед с центром в точке M_0.

В дальнейшем, говоря об окрестности точки, будем иметь в виду окрестность одного из двух упомянутых типов.

Точка M_0 называется **внутренней точкой** множества D, если она принадлежит множеству D вместе с некоторой своей окрестностью. Множество D называется **открытым**, если каждая его точка является внутренней.

Точка M_0 называется **граничной точкой** множества D, если каждая ее окрестность содержит как точки, принадлежащие множеству D, так и точки, не принадлежащие ему. Множество D называется **замкнутым**, если оно содержит все свои граничные точки.

Точка M_0 называется **предельной точкой** множества D, если в любой окрестности этой точки имеются точки множества D, отличные от M_0.

Последовательности точек в евклидовом пространстве

Если каждому натуральному числу n поставлена в соответствие точка M_n евклидова пространства $\mathbf{R}^n$, то множество точек M_1, M_2, ..., M_n, ... называется **последовательностью точек** евклидова пространства $\mathbf{R}^n$ и обозначается $\{M_n\}$.

Последовательность точек $\{M_n\}$ называется **сходящейся** к точке M_0, если $\rho(M_n, M_0)$ является величиной бесконечно малой:

$$\lim_{n\to\infty} \rho(M_n, M_0) = 0.$$

В этом случае точка M_0 называется **пределом** последовательности $\{M_n\}$.

Множество D точек евклидова пространства называется **ограниченным**, если оно содержится в некотором параллелепипеде.

На пределы последовательностей точек в евклидовом пространстве переносятся многие ***с в о й с т в а***, которые были установлены ранее для числовых последовательностей. Наиболее важными из них являются:

1) *единственность предела*;

2) *ограниченность сходящейся последовательности*.

25.3. ПОНЯТИЕ ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ

Определение. Будем говорить, что в области $D \subset \mathbf{R}^n$ задана **функция** $u = f(M)$ (или $u = f(x_1, x_2, ..., x_n)$), если каждой точке $M \in D$ по определенному правилу или закону ставится в соответствие одно определенное число u.

Координаты точки M (т.е. переменные $x_1, x_2, ..., x_n$) называются **независимыми переменными**, или **аргументами**, u – **зависимой переменной**, а символ f – **законом соответствия**. Множество D называется **областью определения** функции.

Область определения функции нескольких переменных (как и в случае функции одной переменной) либо задается заранее, либо является естественной областью определения, т.е. множеством всех таких точек, для которых формула функциональной зависимости f имеет смысл.

В случае когда число аргументов равно двум, функцию обычно обозначают в виде

$$z = f(x, y). \tag{25.5}$$

Область определения такой функции есть некоторое множество точек на плоскости Oxy. ε-*окрестность* точки $M_0(x_0, y_0)$ есть открытый круг с центром в этой точке. *Прямоугольная окрестность* точки $M_0(x_0, y_0)$ – открытый прямоугольник с центром в этой точке.

***Пример* 25.1.** Найти область определения функций

а) $z = \ln(x + y)$; б) $z = \sqrt{x^2 + y^2 - 4} + \dfrac{1}{\sqrt{9 - x^2 - y^2}}$.

Решение. а) Область определения задается неравенством $x + y > 0$, т.е. $y > -x$. Это множество всех точек плоскости, находящихся выше прямой $y = -x$.

б) Очевидно, должны одновременно выполняться неравенства $x^2 + y^2 - 4 \geq 0$ и $9 - x^2 - y^2 > 0$. Поэтому область определения есть множество всех точек плоскости, координаты которых удовлетворяют двойному неравенству:

$$4 \leq x^2 + y^2 < 9.$$

Эта область заключена между окружностью $x^2 + y^2 = 4$ и окружностью $x^2 + y^2 = 9$. При этом точки первой окружности принадлежат этой области, а точки второй – не принадлежат.

График функции двух аргументов

Известно, что график функции одного аргумента есть *линия на плоскости*. Графиком функции двух переменных $z = f(x, y)$ является *поверхность в трехмерном пространстве*, состоящая из точек $(x, y, f(x, y))$.

***Пример* 25.2.** Рассмотрим функцию $z = \sqrt{9 - x^2 - y^2}$. Областью определения этой функции является круг $x^2 + y^2 \le 9$. График есть полусфера радиуса 3 с центром в начале координат (рис. 25.1).

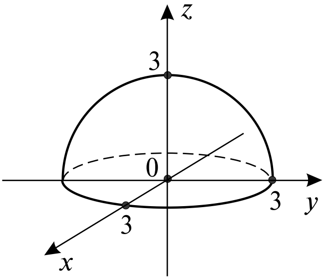


Рис. 25.1. График функции $z = \sqrt{9 - x^2 - y^2}$

***Пример* 25.3.** Рассмотрим функцию $z = x^2 + y^2$. Она определена на всей плоскости Oxy. График этой функции есть поверхность, называемая параболоидом вращения (рис. 25.2). Эта поверхность пересекает плоскость Oxz по параболе $z = x^2$, а плоскость Oyz – по параболе $z = y^2$.

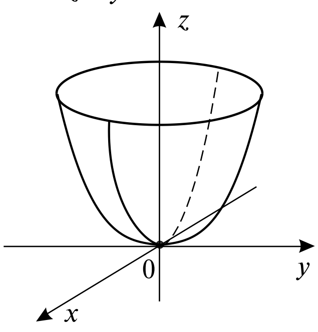


Рис. 25.2. График функции $z = x^2 + y^2$

Для функции n аргументов при $n \geq 3$ формально можно определить понятие графика (это так называемая *гиперповерхность* в $(n + 1)$-мерном пространстве), однако изобразить его на рисунке не удается. Следует заметить, что и для функции двух аргументов построение графика связано со значительными трудностями, а сам график не столь нагляден, как график функции одного аргумента. Поэтому для образного представления функции двух переменных используются линии уровня. **Линией уровня** функции (25.5) называют линию

$$f(x, y) = C,$$

где C – константа.

Например, для функции $z = x^2 - y^2$ любая линия уровня $x^2 - y^2 = C$ при $C \neq 0$ представляет собой гиперболу.

Аналогично для функции трех аргументов $w = f(x, y, z)$ определяется *поверхность уровня*:

$$f(x, y, z) = C.$$

Примеры функций нескольких переменных

Рассмотрим некоторые примеры часто встречающихся функций нескольких переменных.

1. Линейная функция – это функция вида

$$u = a_1x_1 + a_2x_2 + ... + a_nx_n + b, \tag{25.6}$$

где $a_1,\ a_2,\ ...,\ a_n,\ b$ – постоянные числа. Ее можно рассматривать как сумму n линейных функций, каждая из которых зависит от одного аргумента.

2. Функция вида

$$u = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 + 2a_{13}x_1x_3 + ... + a_{nn}x_n^2$$

или, что то же самое,

$$u = \sum_{i,j=1}^{n} a_{ij}x_ix_j, \tag{25.7}$$

где a_{ij} – постоянные числа, называется **квадратичной формой** от n переменных $x_1, x_2, ..., x_n$ (см. § 6.1).

3. В § 13.4 была определена функция полезности. Ее многомерный аналог есть функция $u = f(x_1, x_2, ..., x_n)$, выражающая полезность приобретения n товаров. Наиболее часто она встречается в виде:

- **логарифмической функции**

$$u = \sum_{i=1}^{n} a_i \ln(x_i - c_i), \quad a_i > 0,\ x_i > c_i \geq 0, \tag{25.8}$$

- **функции постоянной эластичности**

$$u = \sum_{i=1}^{n} \frac{a_i}{1-b_i}(x_i - c_i)^{1-b_i}, \quad a_i > 0,\ 0 < b_i < 1,\ x_i > c_i \geq 0. \tag{25.9}$$

4. Функция Кобба – Дугласа

$$z = b_0 x_1^{b_1} x_2^{b_2}. \tag{25.10}$$

Чаще для ее записи используются другие обозначения:

$$Q = AK^{\alpha}L^{\beta}. \tag{25.11}$$

При этом $\alpha,\ \beta \geq 0,\ \alpha + \beta \leq 1.$ В частности, при $\alpha + \beta = 1$ она имеет вид

$$Q = AK^{\alpha}L^{1-\alpha}. \tag{25.12}$$

Это *производственная функция*, выражающая объем выпуска продукции Q (в стоимостном или натуральном выражении) при затратах капитала K и трудовых ресурсов L. Здесь A – параметр производительности конкретно взятой технологии, α – доля капитала в доходе ($0 < \alpha < 1$).

В данном разделе, чтобы избежать громоздких обозначений, будем вести изложение материала в основном для функций двух переменных. При этом практически все понятия и утверждения, сформулированные для функции двух переменных, легко переносятся на случай, когда функция зависит от любого числа переменных.

25.4. ПРЕДЕЛ И НЕПРЕРЫВНОСТЬ

Рассмотрим функцию $z = f(x, y)$, определенную на множестве D. Пусть $M_0(x_0, y_0)$ – предельная точка множества D.

Определение. Число b называется **пределом** функции $f(x, y)$ при стремлении точки $M(x, y)$ к точке $M_0(x_0, y_0)$, если для любого числа $\varepsilon > 0$ существует такое число $\delta > 0$, что для всех точек $M(x, y)$, удовлетворяющих условию $\rho(M, M_0) < \delta$, выполняется неравенство

$$|f(x, y) - b| < \varepsilon.$$

Тот факт, что число b есть предел функции $f(x, y)$ при $M(x, y) \to M_0(x_0, y_0)$, записывается так:

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y) = b, \quad \text{или} \quad \lim_{M \to M_0} f(M) = b. \tag{25.13}$$

Определение. Функция $z = f(x, y)$ называется **непрерывной в точке** $M_0(x_0, y_0)$, если она определена в этой точке, имеет конечный предел при $M(x, y) \to M_0(x_0, y_0)$ и если выполняется равенство

$$\lim_{\substack{x \to x_0 \\ y \to y_0}} f(x, y) = f(x_0, y_0). \tag{25.14}$$

Назовем **полным приращением** функции разность $\Delta z = f(M) - f(M_0) = f(x, y) - f(x_0, y_0)$. Если обозначить $x - x_0 = \Delta x$, $y - y_0 = \Delta y$, то равенство (25.14) можно переписать так:

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \Delta z = 0. \tag{25.15}$$

Функция, непрерывная в каждой точке данной области D, называется **непрерывной в области** D.

Геометрический смысл непрерывности функции двух аргументов очевиден: график непрерывной функции $z = f(x, y)$ представляет собой сплошную поверхность, не имеющую разрывов.

Вопросы

1. Что такое n-мерное евклидово пространство?
2. Как определяется n-мерный шар в евклидовом пространстве?
3. Что такое открытый n-мерный шар? n-мерная сфера?
4. Как определяется n-мерный параллелепипед в евклидовом пространстве?
5. Что такое ε-окрестность точки в n-мерном евклидовом пространстве?
6. Какие точки называются внутренними точками множества в евклидовом пространстве?
7. Что такое замкнутое множество?
8. Может ли множество быть неоткрытым и незамкнутым одновременно?
9. Что называется линией уровня функции двух аргументов?

10. Что называется поверхностью уровня функции трех аргументов?
11. Что такое функция Кобба – Дугласа? Что она выражает?
12. Что такое полное приращение функции двух аргументов?
13. Как определяется понятие непрерывности функции двух аргументов?

Глава 26
ЧАСТНЫЕ ПРОИЗВОДНЫЕ И ИХ ЭКОНОМИЧЕСКИЙ СМЫСЛ. ПОЛНЫЙ ДИФФЕРЕНЦИАЛ

26.1. ЧАСТНЫЕ ПРИРАЩЕНИЯ И ЧАСТНЫЕ ПРОИЗВОДНЫЕ

Пусть функция $z=f(x,y)$ определена в некоторой окрестности точки $M_0(x_0, y_0)$. (Для краткости записи проводим рассуждения для функции двух переменных.) Придадим аргументу x приращение Δx (т.е. перейдем от значения x_0 к значению $x_0+\Delta x$) при фиксированном $y=y_0$, причем так, чтобы точка $M(x_0+\Delta x, y_0)$ принадлежала указанной окрестности. Тогда функция z изменится на величину

$$\Delta_x z = f(x_0+\Delta x, y_0) - f(x_0, y_0).$$

Эта разность называется **частным приращением** функции z по аргументу x. Аналогично определяется частное приращение по аргументу y:

$$\Delta_y z = f(x_0, y_0+\Delta y) - f(x_0, y_0).$$

Определение. **Частной производной** функции нескольких переменных по одной из этих переменных называется предел отношения частного приращения функции к приращению рассматриваемой независимой переменной при стремлении последнего к нулю (если этот предел существует).

Частные производные функции $z=f(x, y)$ в точке $M_0(x_0, y_0)$ обозначаются так:

$$z'_x,\ \frac{\partial z}{\partial x},\ f'_x(x_0, y_0) \qquad \text{(производная по } x\text{);}$$

$$z'_y,\ \frac{\partial z}{\partial y},\ f'_y(x_0, y_0) \qquad \text{(производная по } y\text{).}$$

Таким образом, по определению

$$z'_x = \lim_{\Delta x\to 0}\frac{\Delta_x z}{\Delta x} = \lim_{\Delta x\to 0}\frac{f(x_0+\Delta x, y_0) - f(x_0, y_0)}{\Delta x}, \qquad (26.1)$$

$$z'_y = \lim_{\Delta y \to 0} \frac{\Delta_y z}{\Delta y} = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}. \qquad (26.2)$$

Из определения частных производных следует, что для нахождения частной производной $f'_x(x, y)$ надо рассматривать функцию $z = f(x, y)$ как функцию о д н о г о аргумента x при постоянном y. Аналогично для нахождения $f'_y(x, y)$ константой следует считать x.

П р и м е р **26.1.** Найти частные производные функций:

а) $z = x^5y^2 + x^3y^4$; б) $z = y^x$.

Р е ш е н и е.

а) Считая $y = \text{const}$, находим $z'_x = 5x^4y^2 + 3x^2y^4$. Считая теперь $x = \text{const}$, находим $z'_y = 2x^5y + 4x^3y^3$.

б) Считая $y = \text{const}$, находим z'_x как производную показательной функции: $z'_x = y^x \ln y$. Считая теперь $x = \text{const}$, находим z'_y как производную степенной функции: $z'_y = xy^{x-1}$.

Частные производные второго и высших порядков

Частные производные $z'_x = f'_x(x, y)$ и $z'_y = f'_y(x, y)$ являются функциями от x и y, поэтому можно находить частные производные от них. Эти производные называются **частными производными второго порядка** от функции $z = f(x, y)$:

$$z''_{xx} = (z'_x)'_x, \quad z''_{xy} = (z'_x)'_y, \quad z''_{yx} = (z'_y)'_x, \quad z''_{yy} = (z'_y)'_y.$$

Аналогично определяются частные производные третьего и более высоких порядков, например: $z'''_{xxy} = (z''_{xx})'_y$.

П р и м е р **26.2.** Найти частные производные второго порядка от функции $z = x^5y^2 + x^3y^4$.

Р е ш е н и е. В примере 26.1 уже были найдены частные производные первого порядка:

$$z'_x = 5x^4y^2 + 3x^2y^4; \qquad z'_y = 2x^5y + 4x^3y^3.$$

Найдем теперь частные производные второго порядка:

$$z''_{xx} = 20x^3y^2 + 6xy^4; \qquad z''_{xy} = 10x^4y + 12x^2y^3;$$

$$z''_{yx} = 10x^4y + 12x^2y^3; \qquad z''_{yy} = 2x^5 + 12x^3y^2.$$

Частные производные z''_{xy} и z''_{yx} называются **смешанными частными производными**.

Пример **26.3.** Найти z''_{xy} и z''_{yx} от функции $z = x^2 e^y$.

Решение: $z'_x = 2xe^y$, $z''_{xy} = 2xe^y$; $z'_y = x^2 e^y$, $z''_{yx} = 2xe^y$.

Как видим, для функций, рассмотренных в примерах 26.2 и 26.3, смешанные производные совпадают: $z''_{xy} = z''_{yx}$. Это не случайно. Справедливо следующее утверждение.

Теорема **26.1.** Если частные производные второго порядка функции $z = f(x, y)$ непрерывны в точке (x_0, y_0), то в этой точке $z''_{xy} = z''_{yx}$.

Не будем доказывать эту теорему. Заметим только, что обычно функции, используемые в экономике, имеют непрерывные частные производные второго порядка.

Экономический смысл частных производных

Рассмотрим в качестве примера производственную функцию Кобба – Дугласа [см. формулу (25.12)]:

$$Q = AK^{\alpha}L^{1-\alpha}.$$

Найдем скорость изменения объема продукции Q при изменении одного из факторов: затрат капитала K или величины трудовых ресурсов L. Решение этой задачи дают частные производные функции Q:

$$Q'_K = A\alpha K^{\alpha-1}L^{1-\alpha}, \qquad Q'_L = A(1-\alpha)K^{\alpha}L^{-\alpha}.$$

Частная производная $Q'_K = A\alpha K^{\alpha-1}L^{1-\alpha}$ называется **предельной фондоотдачей**, а частная производная $Q'_L = A(1-\alpha)K^{\alpha}L^{-\alpha}$ – **предельной производительностью труда**.

Напомним, что в случае функции одной переменной $y = f(x)$ *эластичностью функции по аргументу* называется величина

$$E_x(y) = x\frac{y'}{y}$$

[см. формулу (20.2)].

Для функции нескольких переменных обычная производная заменяется частной производной. Для функции Кобба – Дугласа эластичность выпуска продукции по затратам капитала

$$E_K(Q) = K\,\frac{Q'_K}{Q} = \alpha.$$

Аналогично эластичность выпуска по труду $E_L(Q) = L\,\dfrac{Q'_L}{Q} = 1-\alpha.$

Итак, в функции Кобба – Дугласа показатели степеней α и $1-\alpha$ представляют собой соответственно *коэффициенты эластичности* по каждому из ее аргументов.

26.2. ПОЛНОЕ ПРИРАЩЕНИЕ И ПОЛНЫЙ ДИФФЕРЕНЦИАЛ

Полным приращением функции $z=f(x, y)$ в точке $M_0(x_0, y_0)$, соответствующим приращениям аргументов Δx и Δy, называется разность

$$\Delta z = f(x_0+\Delta x,\ y_0+\Delta y) - f(x_0, y_0). \tag{26.3}$$

Определение. Функция $z=f(x, y)$ называется **дифференцируемой** в точке $M_0(x_0, y_0)$, если ее полное приращение (26.3) в этой точке можно представить в виде

$$\Delta z = A_1\Delta x + A_2\Delta y + \alpha_1\Delta x + \alpha_2\Delta y, \tag{26.4}$$

где A_1, A_2 – не зависящие от $\Delta x, \Delta y$ константы, а α_1, α_2 – бесконечно малые при $\Delta x \to 0,\ \Delta y \to 0$.

Если хотя бы одно из чисел A_1, A_2 отлично от нуля, то сумма

$$L = A_1\Delta x + A_2\Delta y \tag{26.5}$$

есть главная линейная относительно $\Delta x,\ \Delta y$ часть приращения дифференцируемой функции. Эта главная часть приращения функции называется **полным дифференциалом**.

***Теорема* 26.2.** Если функция $z=f(x, y)$ дифференцируема в точке $M_0(x_0, y_0)$, то она имеет в этой точке частные производные по x и y.

Доказательство. Из равенства (26.4) следует, что $\Delta_x z = A_1\Delta x + \alpha_1\Delta x$, откуда

$$\frac{\Delta_x z}{\Delta x} = A_1 + \alpha_1. \tag{*}$$

Так как $\alpha_1 \to 0$ при $\Delta x \to 0$, то из (*) следует, что существует предел

$$\lim_{\Delta x\to 0}\frac{\Delta_x z}{\Delta x} = A_1,$$

т.е. z'_x существует (и $z'_x = A_1$).

Аналогично доказывается, что существует z'_y (и $z'_y = A_2$).

Таким образом, главная линейная часть приращения функции $z = f(x, y)$ имеет вид $z'_x \Delta x + z'_y \Delta y$.

Теперь мы можем сформулировать понятие полного дифференциала в следующем виде.

Определение. **Полным дифференциалом** dz функции $z = f(x, y)$ называется сумма произведений частных производных этой функции на приращения соответствующих независимых аргументов, т.е.

$$dz = z'_x \Delta x + z'_y \Delta y. \tag{26.6}$$

Рассмотрим, в частности, функцию $z = f(x, y) = x$. Очевидно, $dz = dx = 1 \cdot \Delta x$, т.е. $dx = \Delta x$. Аналогично, рассматривая $z = f(x, y) = y$, получаем $dy = \Delta y$. Поэтому формулу (26.6) можно записать в виде

$$dz = z'_x dx + z'_y dy, \tag{26.7}$$

или, что то же самое,

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy.$$

Заметим, что дифференциал, являясь главной частью приращения функции, применяется в приближенных вычислениях.

***Пример* 26.4.** Вычислить приближенно $\sqrt{(1{,}02)^3 + (1{,}97)^3}$.

Решение. Искомое число будем рассматривать как значение функции $z = f(x, y) = \sqrt{x^3 + y^3}$ при $x = x_0 + \Delta x$, $y = y_0 + \Delta y$, где $x_0 = 1$, $y_0 = 2$, $\Delta x = 0{,}02$, $\Delta y = -0{,}03$. Имеем:

$$f(1, 2) = \sqrt{1^3 + 2^3} = 3,$$

$$\Delta z \approx dz = \frac{3x^2 dx + 3y^2 dy}{2\sqrt{x^3 + y^3}},$$

$$\Delta z(1, 2) \approx \frac{3 \cdot 1 \cdot 0{,}02 + 3 \cdot 2^2 (-0{,}03)}{2 \cdot 3} = \frac{0{,}06 - 0{,}36}{6} = -0{,}05.$$

Следовательно, $\sqrt{(1{,}02)^3 + (1{,}97)^3} \approx 3 - 0{,}05 = 2{,}95.$

***Теорема* 26.3.** Если функция $z = f(x, y)$ дифференцируема в данной точке, то она непрерывна в этой точке.

Доказательство. Из условия дифференцируемости (26.4) следует, что

$$\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\Delta z=0,$$

а это и означает непрерывность функции [см. формулу (25.15)].

Напомним, что для функции о д н о г о аргумента $y=f(x)$ существование производной равносильно дифференцируемости. Однако для функции н е с к о л ь к и х аргументов аналогичное утверждение, вообще говоря, не верно. Из существования частных производных по всем аргументам не следует, что функция дифференцируема, и даже не следует, что она непрерывна. Можно показать (мы этого делать не будем), что функция

$$z=\begin{cases}\dfrac{xy}{x^2+y^2}, & \text{если } x^2+y^2\neq 0,\\[2mm] 0, & \text{если } x=0,\ y=0,\end{cases}$$

не является дифференцируемой (и не является непрерывной) в точке $O(0,\,0)$. Тем не менее в этой точке (и во всех других точках) указанная функция имеет частные производные по x и y.

Итак, существования частных производных, вообще говоря, недостаточно для дифференцируемости функции нескольких переменных.

Д о с т а т о ч н ы е условия дифференцируемости дает следующая теорема (приведем ее без доказательства).

***Т е о р е м а* 26.4.** Если функция $z=f(x,y)$ имеет частные производные $f'_x(x,y)$ и $f'_y(x,y)$ в некоторой окрестности точки M и эти производные н е п р е р ы в н ы в точке M, то данная функция дифференцируема в точке M.

Дифференциалы высших порядков

Пусть в области D задана функция $z=f(x,y)$, имеющая непрерывные частные производные первого порядка. Тогда, как мы уже знаем, дифференциалом dz называется следующее выражение:

$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy,$$

где dx, dy – дифференциалы независимых аргументов x и y, или, что то же самое, произвольные приращения этих аргументов.

Очевидно, dz также является некоторой функцией от x, y. Если функция z имеет непрерывные частные производные второго порядка, то дифференциал dz имеет непрерывные частные производные первого порядка и можно поставить вопрос о дифференциале от этого дифференциала dz, т.е. о $d(dz)$. Последний называется **дифференциалом второго порядка** (или **вторым дифференциалом**) и обозначается символом d^2z:

$$d^2z = d(dz).$$

При вычислении второго дифференциала дифференциалы независимых аргументов (т.е. их приращения) dx и dy рассматриваются как константы. Поэтому $d^2x = d^2y = 0$. Итак,

$$d^2z = d(dz) = d\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\right) = d\left(\frac{\partial z}{\partial x}\right)dx + d\left(\frac{\partial z}{\partial y}\right)dy.$$

Отсюда

$$d^2z = \left(\frac{\partial^2 z}{\partial x^2}dx + \frac{\partial^2 z}{\partial x \partial y}dy\right)dx + \left(\frac{\partial^2 z}{\partial y \partial x}dx + \frac{\partial^2 z}{\partial y^2}dy\right)dy,$$

и с учетом того, что $\dfrac{\partial^2 z}{\partial x \partial y} = \dfrac{\partial^2 z}{\partial y \partial x}$, получаем

$$d^2z = \frac{\partial^2 z}{\partial x^2}dx^2 + 2\frac{\partial^2 z}{\partial x \partial y}dx\,dy + \frac{\partial^2 z}{\partial y^2}dy^2.$$

Разумеется, можно было записать это выражение несколько иначе, учитывая, что мы по-разному можем обозначать одни и те же частные производные, например $\dfrac{\partial^2 z}{\partial x^2}$ можно обозначать и как $\dfrac{\partial^2 f}{\partial x^2}$, и как z''_{xx} и т.д.

Аналогично записываются дифференциалы второго порядка для функций трех и большего числа аргументов, т.е. для функций $w = f(x, y, z)$, $w = f(x_1, x_2, ..., x_n)$.

Дифференциалы более высоких порядков определяются по тому же правилу:

$$d^3z = d(d^2z), ..., d^nz = d(d^{n-1}z).$$

Мы здесь рассматривали функции независимых аргументов. Для случая, когда аргументы сами являются *функциями*, например $x = \varphi(s, t)$, $y = \psi(s, t)$, можно показать (подобно тому, как это делалось для функций одного аргумента), что форма дифференциала первого порядка и н в а р и а н т н а, а для дифференциалов второго и высших порядков — не инвариантна. Не будем останавливаться на этом подробно.

Замечание. Несмотря на то что для дифференциалов второго и более высоких порядков инвариантность формы вообще не имеет места, тем не менее в некоторых простых частных случаях форма дифференциала любого порядка может оставаться неизменной. В частности, в случае когда аргументы x и y функции $z = f(x, y)$ линейно зависят от независимого аргумента t: $x = a_1 t + a_2$, $y = b_1 t + b_2$, форма дифференциала любого порядка $d^n z$ остается неизменной. В этом легко убедиться.

26.3. ПРОИЗВОДНАЯ ПО НАПРАВЛЕНИЮ. ГРАДИЕНТ

Пусть функция $z = f(x, y)$ определена в некоторой окрестности точки $M_0(x_0, y_0)$. Рассмотрим некоторое направление, задаваемое единичным вектором $\bar{l} = (\cos\alpha, \cos\beta)$, где $\cos^2\alpha + \cos^2\beta = 1$ (рис. 26.1). На прямой, проходящей по этому направлению через точку M_0, возьмем точку $M(x_0 + \Delta x, y_0 + \Delta y)$. Обозначим через Δl длину отрезка M_0M. Очевидно,

$$\Delta l = \sqrt{\Delta x^2 + \Delta y^2}.$$

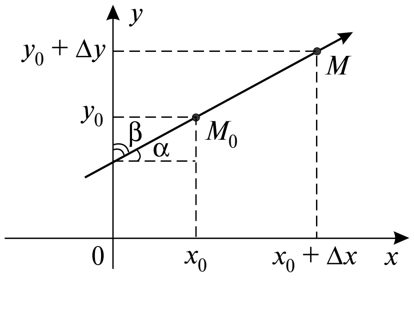


Рис. 26.1. Направление $\bar{l} = (\cos\alpha, \cos\beta)$

Рассмотрим приращение функции $f(x, y)$:

$$\Delta z = f(x_0 + \Delta x,\ y_0 + \Delta y) - f(x_0, y_0),$$

где Δx и Δy связаны соотношениями $\Delta x = \Delta l \cos\alpha$, $\Delta y = \Delta l \cos\beta$.

Определение. Предел отношения $\frac{\Delta z}{\Delta l}$ при $\Delta l \to 0$ называется **производной** функции $z = f(x, y)$ в точке $M_0(x_0, y_0)$ **по направлению** $\bar{l}$ и обозначается $\frac{\partial z}{\partial l}$ (или z'_l):

$$\frac{\partial z}{\partial l} = \lim_{\Delta l \to 0} \frac{\Delta z}{\Delta l}.$$

Производная $\frac{\partial z}{\partial l}$ характеризует скорость изменения функции в направлении $\bar{l}$.

Очевидно, что обычные частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ представляют собой производные по направлениям, параллельным осям Ox и Oy соответственно. Легко убедиться в том, что для дифференцируемой функции $z = f(x, y)$:

$$\frac{\partial z}{\partial l} = \frac{\partial z}{\partial x}\cos\alpha + \frac{\partial z}{\partial y}\cos\beta, \quad \text{или} \quad z'_l = z'_x \cos\alpha + z'_y \cos\beta. \tag{26.8}$$

Определение. **Градиентом** функции $z = f(x, y)$ в точке M называется вектор, координаты которого равны соответственно частным производным $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ в этой точке.

Градиент функции обозначается $\operatorname{grad} z$:

$$\operatorname{grad} z = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right), \quad \text{или} \quad \operatorname{grad} z = \left(z'_x,\ z'_y\right). \tag{26.9}$$

Сравнивая равенства (26.8) и (26.9), видим, что производная по направлению $\bar{l}$ есть скалярное произведение векторов $\bar{l}$ и $\operatorname{grad} z$:

$$\frac{\partial z}{\partial l} = \bar{l} \operatorname{grad} z.$$

Известно, что скалярное произведение двух векторов имеет максимальное значение, когда угол между ними равен нулю. Следовательно, производная функции $z = f(x, y)$ по направлению $\bar{l}$ принимает максимальное значение, когда направления векторов $\bar{l}$ и $\operatorname{grad} z$ совпадают.

Таким образом, градиент функции характеризует направление, по которому функция изменяется наиболее быстро.

Рассмотрим **геометрический смысл градиента**. Линия уровня функции $z = f(x, y)$, проходящая через точку (x_0, y_0), задается уравнением $f(x, y) = C$, где $C = f(x_0, y_0)$. При определенных условиях это уравнение можно разрешить относительно y, т.е. выразить y в виде $y = g(x)$. (Если это невозможно, то, разрешив уравнение относительно x, $x = h(y)$, можно повторить все рассуждения для этого случая.) Известно, что угловой коэффициент касательной есть $g'(x)$, т.е. направляющий вектор касательной имеет координаты $(1, g'(x))$, или, что то же самое, $\left(1, \frac{dg}{dx}\right)$. Следовательно, вектор (dx, dy) также есть направляющий вектор касательной.

Беря дифференциал от левой и правой частей уравнения, задающего линию уровня, получаем

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy = 0,$$

т.е. скалярное произведение градиента и направляющего вектора касательной равно нулю, следовательно, эти векторы перпендикулярны.

Понятие градиента функции обобщается на случай любого числа переменных. В частности, для функции $u = f(x, y, z)$:

$$\operatorname{grad} u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}\right).$$

В случае функции трех аргументов все сказанное выше остается в силе, только вместо линии уровня будет фигурировать поверхность уровня, а вместо касательной к линии уровня – касательная плоскость к поверхности уровня, т.е. плоскость

$$f'_x(x_0, y_0, z_0)(x - x_0) + f'_y(x_0, y_0, z_0)(y - y_0) + f'_z(x_0, y_0, z_0)(z - z_0) = 0.$$

***Пример* 26.5.** Найти градиент функции $z = x^2 + \frac{y^2}{4}$ в точке $M_0(2, 6)$ и его модуль.

Решение: $\operatorname{grad} z = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = \left(2x, \frac{y}{2}\right)$. При $x = 2$, $y = 6$ получаем:

$$\operatorname{grad} z\big|_{(2,6)} = (4, 3); \quad |\operatorname{grad} z| = \sqrt{4^2 + 3^2} = 5.$$

Пример 26.6. Найти градиент функции $u = x^2 + \frac{y^2}{2} - z^2$ в точке $M_0(1, 1, 1)$ и его модуль.

Решение: $\operatorname{grad} u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}\right) = (2x, y, -2z)$;

$\operatorname{grad} u\big|_{(1,1,1)} = (2, 1, -2)$; $|\operatorname{grad} u| = \sqrt{2^2 + 1^2 + (-2)^2} = 3.$

26.4. ФОРМУЛА ТЕЙЛОРА

Пусть функция $z = f(x, y)$ в окрестности некоторой точки $M_0(x_0, y_0)$ имеет непрерывные производные всех порядков – до $(n + 1)$-го включительно. Придадим x_0 и y_0 некоторые приращения Δx и Δy соответственно, причем так, чтобы прямолинейный отрезок, соединяющий точки (x_0, y_0) и $(x_0 + \Delta x, y_0 + \Delta y)$, целиком принадлежал рассматриваемой окрестности точки (x_0, y_0).

Покажем, что в этом случае имеет место следующее равенство:

$$\Delta f(x_0, y_0) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = df(x_0, y_0) +$$
$$+ \frac{1}{2!} d^2 f(x_0, y_0) + \ldots + \frac{1}{n!} d^n f(x_0, y_0) + \frac{1}{(n+1)!} d^{n+1} f(\xi, \eta), \qquad (26.10)$$

где ξ – точка, расположенная между x_0 и $x_0 + \Delta x$, а η – между y_0 и $y_0 + \Delta y$ ($\xi = x_0 + \theta \Delta x$, $\eta = y_0 + \theta \Delta y$, $0 < \theta < 1$).

Для доказательства сделаем замену:

$$x = x_0 + t\Delta x, \quad y = y_0 + t\Delta y, \quad t \in [0, 1]. \qquad (26.11)$$

Подставив эти значения x и y в функцию $f(x, y)$, получим сложную функцию от одного аргумента t:

$$F(t) = f(x_0 + t\Delta x, \ y_0 + t\Delta y).$$

Формулы (26.11) геометрически выражают прямолинейный отрезок, соединяющий точки $M_0(x_0, y_0)$ и $M_1(x_1, y_1)$, при этом точка $M_0(x_0, y_0)$ соответствует значению $t = 0$, а точка $M_1(x_1, y_1)$ – значению $t = 1$.

Теперь приращение $\Delta f(x_0, y_0) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ можно заменить равным ему приращением $\Delta F(0) = F(1) - F(0)$. Но функция $F(t)$ является функцией от одной переменной и имеет непрерывные производные до (n + 1)-го порядка включительно. Следовательно, она может быть разложена по формуле Тейлора. Запишем это разложение в форме (17.13″):

$$\Delta F(t_0)=dF(t_0)+\frac{1}{2!}d^2F(t_0)+\ldots+\frac{1}{n!}d^nF(t_0)+\frac{1}{(n+1)!}d^{n+1}F(t_0+\theta\Delta t),$$

$$0<\theta<1.$$

Отсюда получаем

$$\Delta F(0)=F(1)-F(0)=dF(0)+\frac{1}{2!}d^2F(0)+\ldots+$$

$$+\frac{1}{n!}d^nF(0)+\frac{1}{(n+1)!}d^{n+1}F(\theta),\qquad 0<\theta<1. \tag{26.12}$$

Заметим при этом, что дифференциал dt, входящий в различных степенях в правую часть (26.12) (т.е. содержащийся в выражениях $dF(0) = F'(0)dt$, $d^2F(0) = F''(0)dt^2$, ...), равен приращению $\Delta t = 1 - 0 = 1$.

Теперь с учетом проведенной (линейной) замены (26.11), а также принимая во внимание сделанное ранее замечание об инвариантности дифференциала любого порядка по отношению к линейной замене переменных (см. § 26.2), получаем

$$dF(0)=f'_x(x_0,y_0)dx+f'_y(x_0,y_0)dy=df(x_0,y_0),$$

$$d^2F(0)=f''_{xx}(x_0,y_0)dx^2+2f''_{xy}(x_0,y_0)dx\,dy+f''_{yy}(x_0,y_0)dy^2=$$
$$=d^2f(x_0,y_0),$$

$$\ldots$$

$$d^nF(0)=d^nf(x_0,y_0).$$

Наконец, для дифференциала (n + 1)-го порядка

$$d^{n+1}F(\theta)=d^{n+1}f(x_0+\theta\Delta x,\,y_0+\theta\Delta y).$$

Отметим, что дифференциалы dx и dy (как и для независимых аргументов) равны приращениям Δx и Δy соответственно. Действительно, учитывая (26.11) и то, что $dt = 1$, имеем:

$$dx = x'_t dt = (x_0 + t\Delta x)'_t dt = \Delta x dt = \Delta x,$$
$$dy = y'_t dt = (y_0 + t\Delta y)'_t dt = \Delta y dt = \Delta y.$$

Теперь подставим выражения для $dF(0)$, $d^2F(0)$, ... в (26.12):

$$\Delta f(x_0, y_0) = f(x_0 + \Delta x,\ y_0 + \Delta y) - f(x_0, y_0) = df(x_0, y_0) +$$
$$+\frac{1}{2!}d^2 f(x_0, y_0) + \ldots + \frac{1}{n!}d^n f(x_0, y_0) +$$
$$+\frac{1}{(n+1)!}d^{n+1} f(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y), \qquad 0 < \theta < 1. \qquad (26.13)$$

Мы получили для функции $z = f(x, y)$ **формулу Тейлора в дифференциальной форме**. В развернутом виде она (даже для рассматриваемого случая функции двух аргументов) выглядит значительно сложнее.

Напишем формулу (26.13) с учетом того, что $dx = \Delta x$, $dy = \Delta y$, в развернутом виде, ограничившись только двумя членами разложения (т.е. при $n = 2$):

$$f(x, y) = f(x_0, y_0) + f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0) +$$
$$+\frac{1}{2!}\Big[f''_{xx}(x_0, y_0)(x - x_0)^2 + 2f''_{xy}(x_0, y_0)(x - x_0)(y - y_0) +$$
$$+ f''_{yy}(x_0, y_0)(y - y_0)^2\Big] +$$
$$+\frac{1}{3!}\Big[f'''_{xxx}(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y)(x - x_0)^3 +$$
$$+ 3f'''_{xxy}(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y)(x - x_0)^2(y - y_0) +$$
$$+ 3f'''_{xyy}(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y)(x - x_0)(y - y_0)^2 +$$
$$+ f'''_{yyy}(x_0 + \theta\Delta x,\ y_0 + \theta\Delta y)(y - y_0)^3\Big]. \qquad (26.14)$$

Это формула Тейлора для $z = f(x, y)$ при $n = 2$. Как видим, выглядит она в развернутом виде громоздко, хотя функция зависит только от двух переменных, и мы взяли всего два члена разложения.

Вопросы

1. Что называется частным приращением функции? Чем отличается частное приращение от полного?
2. Что называется частной производной функции нескольких аргументов по одному из аргументов?
3. Отличается ли принципиально процесс нахождения частной производной от процесса дифференцирования функции одного аргумента?
4. Что такое смешанные частные производные?
5. Каким свойством обладают непрерывные смешанные частные производные?
6. Каков экономический смысл частных производных функции Кобба – Дугласа?
7. Что есть эластичность выпуска продукции по труду для функции Кобба – Дугласа? А что есть эластичность выпуска продукции по затратам капитала для той же функции?
8. Что такое дифференцируемая функция двух аргументов?
9. Что называется полным дифференциалом функции двух аргументов?
10. Достаточно ли существования частных производных по обоим аргументам для того, чтобы функция двух аргументов была дифференцируемой?
11. Можно ли утверждать, что функция двух аргументов, имеющая в данной точке частные производные по обоим аргументам, непрерывна в этой точке?
12. На чем основано применение полного дифференциала в приближенных вычислениях?
13. Как определяется производная по данному направлению? Что характеризует производная по направлению и какой величиной она является – скалярной или векторной?
14. Что такое градиент функции двух аргументов? Скалярной или векторной величиной является градиент?
15. В каком случае производная по направлению принимает наибольшее значение?

16. В чем заключается геометрический смысл градиента?
17. Как выглядит формула Тейлора для функции двух аргументов в дифференциальной форме?

Глава 27
ЭКСТРЕМУМЫ. УСЛОВНЫЕ ЭКСТРЕМУМЫ

27.1. ЛОКАЛЬНЫЙ ЭКСТРЕМУМ ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Как уже отмечалось, проводим рассуждения для функции двух аргументов.

Пусть функция $z = f(x, y)$ определена в некоторой окрестности точки $M_0(x_0, y_0)$.

Определение. Точка $M_0(x_0, y_0)$ называется **точкой локального максимума (минимума)** функции $z = f(x, y)$, если существует такая окрестность точки M_0, что для всех точек $M(x, y)$ из этой окрестности выполняется неравенство $f(x_0, y_0) > f(x, y)$ (соответственно $f(x_0, y_0) < f(x, y)$).

Если $M_0(x_0, y_0)$ – точка локального максимума (минимума) функции $f(x, y)$, то значение $f(x_0, y_0)$ называется **локальным максимумом (минимумом)** функции. Общий термин для локального максимума и минимума – **локальный экстремум**.

Необходимое условие экстремума

***Теорема* 27.1.** Если функция $z = f(x, y)$ имеет частные производные в точке локального экстремума $M_0(x_0, y_0)$, то

$$f'_x(x_0, y_0) = f'_y(x_0, y_0) = 0. \tag{27.1}$$

Доказательство. Зафиксируем $y = y_0$. Получим функцию одной переменной $f(x, y_0)$. Ее производная совпадает с частной производной $f'_x(x, y_0)$, а сама функция имеет локальный экстремум в точке x_0. Согласно теореме Ферма (см. § 17.1) $f'_x(x_0, y_0) = 0$. Аналогично, фиксируя $x = x_0$ и рассматривая $f(x_0, y)$, докажем, что $f'_y(x_0, y_0) = 0$.

Следует отметить, что условие (27.1) не является достаточным условием экстремума. Рассмотрим, например, функцию $z = xy$. Ее частные производные равны нулю в точке $O(0, 0)$, однако в этой точке функция не имеет экстремума. Действительно,

$f(0, 0) = 0$, но в любой окрестности точки O есть как положительные, так и отрицательные значения функции.

Точки, в которых выполняются необходимые условия экстремума (т.е. частные производные z'_x и z'_y равны нулю), называются **критическими**, или **стационарными**. Стационарные точки функции $f(x, y)$ можно найти, решив систему уравнений

$$\begin{cases} f'_x(x, y) = 0, \\ f'_y(x, y) = 0. \end{cases} \tag{27.2}$$

***Пример* 27.1.** Найти стационарные точки функции $z = x^3 + 8y^3 - 6xy + 1$.

Решение: $z'_x = 3x^2 - 6y$, $z'_y = 24y^2 - 6x$. Получаем систему

$$\begin{cases} 3x^2 - 6y = 0, \\ 24y^2 - 6x = 0, \end{cases} \quad \text{или} \quad \begin{cases} x^2 - 2y = 0, \\ 4y^2 - x = 0. \end{cases}$$

Из первого уравнения находим $2y = x^2$, $4y^2 = x^4$. Подставляя во второе уравнение, получаем $x^4 - x = 0$, т.е. $x(x^3 - 1) = 0$. Это уравнение имеет два действительных корня: $x_1 = 0$, $x_2 = 1$. Из первого уравнения системы находим $y_1 = 0$, $y_2 = \frac{1}{2}$. Следовательно, имеется две стационарные точки: $(0, 0)$ и $\left(1, \frac{1}{2}\right)$.

Достаточные условия экстремума

***Теорема* 27.2.** Пусть функция $z = f(x, y)$ имеет непрерывные частные производные второго порядка в некоторой окрестности стационарной точки $M_0(x_0, y_0)$, пусть $f''_{xx}(x_0, y_0) = A$, $f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0) = B$, $f''_{yy}(x_0, y_0) = C$, $D = AC - B^2$. Тогда: 1) если $D > 0$, то функция имеет в точке (x_0, y_0) локальный экстремум, причем если $A < 0$ – локальный максимум, а если $A > 0$ – локальный минимум; 2) если $D < 0$, то в точке (x_0, y_0) нет экстремума.

Доказательство. Рассмотрим разность $\Delta f = f(x, y) - f(x_0, y_0)$. Воспользуемся формулой Тейлора (26.13), ограничившись $n = 1$ (т.е. разложение будет содержать только первый член и остаточный член R_1):

$$\Delta f = \Delta f(x_0, y_0) = f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y +$$
$$+ \frac{1}{2!}\left[f''_{xx}(\xi, \eta)\Delta x^2 + 2f''_{xy}(\xi, \eta)\Delta x\Delta y + f''_{yy}(\xi, \eta)\Delta y^2 \right]. \tag{27.3}$$

(Здесь $\xi = x_0 + \theta\Delta x$, $\eta = y_0 + \theta\Delta y$, $0 < \theta < 1$.)

Так как точка $M_0(x_0, y_0)$ – стационарная, то первые члены разложения обращаются в нуль, и мы получаем более простое выражение для приращения функции в точке (x_0, y_0):

$$\Delta f = \frac{1}{2!}\left[f''_{xx}(\xi, \eta)\Delta x^2 + 2f''_{xy}(\xi, \eta)\Delta x\Delta y + f''_{yy}(\xi, \eta)\Delta y^2 \right]. \tag{27.4}$$

В соответствии с нашими обозначениями

$$f''_{xx}(x_0, y_0) = A, \qquad f''_{xy}(x_0, y_0) = B, \qquad f''_{yy}(x_0, y_0) = C.$$

Так как вторые производные непрерывны, то

$$f''_{xx}(\xi, \eta) = f''_{xx}(x_0 + \theta\Delta x, y_0 + \theta\Delta y) = A + \alpha_{11},$$

$$f''_{xy}(\xi, \eta) = B + \alpha_{12}, \qquad f''_{yy}(\xi, \eta) = C + \alpha_{22},$$

где $\alpha_{11}, \alpha_{12}, \alpha_{22}$ – бесконечно малые при $\Delta x \to 0$, $\Delta y \to 0$.

Теперь можно переписать Δf в виде

$$\Delta f = \frac{1}{2}\left[A\Delta x^2 + 2B\Delta x\Delta y + C\Delta y^2 + \alpha_{11}\Delta x^2 + 2\alpha_{12}\Delta x\Delta y + \alpha_{22}\Delta y^2 \right].$$

Нас интересует знак разности Δf. Мы видим, что знак Δf зависит от знака выражения $D = AC - B^2$. Обозначим расстояние между точками $M_0(x_0, y_0)$ и $M(x, y)$ через r. Очевидно, $r = \sqrt{\Delta x^2 + \Delta y^2}$. Теперь $\Delta x = r\cos\varphi$, $\Delta y = r\sin\varphi$ (где φ – угол между отрезком M_0M и Ox). Еще раз перепишем выражение Δf:

$$\Delta f = \frac{r^2}{2}\Big[A\cos^2\varphi + 2B\cos\varphi\sin\varphi + C\sin^2\varphi +$$

$$+ \alpha_{11}\cos^2\varphi + 2\alpha_{12}\cos\varphi\sin\varphi + \alpha_{22}\sin^2\varphi \Big]. \tag{27.5}$$

1. Пусть $\boldsymbol{AC - B^2 > 0}$. В этом случае $AC > 0$, следовательно, $A \neq 0$, и **первый** трехчлен в скобках в выражении (27.5) можно преобразовать следующим образом:

$$A\cos^2\varphi + 2B\cos\varphi\sin\varphi + C\sin^2\varphi =$$

$$= \frac{1}{A}\left[(A\cos\varphi + B\sin\varphi)^2 + (AC - B^2)\sin^2\varphi \right]. \tag{27.6}$$

Отсюда ясно, что выражение в квадратных скобках (при нашем предположении $AC - B^2 > 0$) всегда положительно. Следовательно, упомянутый трехчлен при всех значениях φ отличен от нуля и имеет тот же знак, что и коэффициент A. Этот трехчлен есть функция аргумента φ, непрерывная на отрезке $[0, 2\pi]$. Эта функция, согласно второй теореме Вейерштрасса (см. § 15.2), достигает на $[0, 2\pi]$ своего наименьшего значения. Это наименьшее значение отлично от нуля. Следовательно, модуль этого квадратного трехчлена имеет положительное наименьшее значение m:

$$|A\cos^2\varphi + 2B\cos\varphi\sin\varphi + C\sin^2\varphi| \geq m > 0.$$

Рассмотрим теперь в т о р о й трехчлен в скобках в правой части равенства (27.5). Очевидно,

$$|\alpha_{11}\cos^2\varphi + 2\alpha_{12}\cos\varphi\sin\varphi + \alpha_{22}\sin^2\varphi| \leq |\alpha_{11}| + 2|\alpha_{12}| + |\alpha_{22}|.$$

Так как $\alpha_{11}, \alpha_{12}, \alpha_{22}$ – бесконечно малые при $\Delta x \to 0$, $\Delta y \to 0$, то при достаточно малых Δx и Δy будет выполняться неравенство

$$|\alpha_{11}| + 2|\alpha_{12}| + |\alpha_{22}| < m.$$

Поэтому выражение в скобках в правой части равенства (27.5) будет сохранять тот же знак, что и первый из трехчленов, т.е. знак A. Следовательно, и левая часть $\Delta f = f(x, y) - f(x_0, y_0)$ сохраняет знак A.

Итак, если $\boldsymbol{A > 0}$, то и $\Delta f > 0$, т.е. в точке (x_0, y_0) функция имеет м и н и м у м; в случае же $\boldsymbol{A < 0}$ будет $\Delta f < 0$, т.е. имеется м а к с и м у м.

2. Пусть теперь $\boldsymbol{AC - B^2 < 0}$. Рассмотрим отдельно случаи, когда $A \neq 0$ и $A = 0$.

1) $\boldsymbol{A \neq 0}$. Снова воспользуемся преобразованием (27.6). Убедимся, что в этом случае в сколь угодно малой близости от рассматриваемой точки $M_0(x_0, y_0)$ разность Δf может быть и положительной, и отрицательной, т.е. в точке $M_0(x_0, y_0)$ экстремума нет.

Пусть $\varphi = \varphi_1 = 0$. Тогда в правой части равенства (27.6) выражение в квадратных скобках будет положительно (и равно A^2).

Если же $\varphi = \varphi_2$ определить из условия $A\cos\varphi + B\sin\varphi = 0$ (т.е. $\varphi_2 = -\operatorname{arctg}\dfrac{A}{B}$), то упомянутое выражение будет отрицательным (и равно $(AC - B^2)\sin^2\varphi_2$).

Как уже отмечалось, второй трехчлен в правой части равенства (27.5) при достаточно малом r на знак Δf не влияет.

Очевидно, мы можем сколь угодно близко к $M_0(x_0, y_0)$ взять точку $M_1(x_1, y_1)$ так, чтобы отрезок M_0M_1 образовывал с Ox угол $\varphi = \varphi_1 = 0$. Для нее $\Delta f > 0$. Точно так же мы можем сколь угодно близко к $M_0(x_0, y_0)$ взять точку $M_2(x_2, y_2)$ так, чтобы отрезок M_0M_2 образовывал с Ox угол $\varphi = \varphi_2$. Для этой точки будет $\Delta f < 0$.

Итак, в рассматриваемом случае $AC - B^2 < 0$, $A \neq 0$ в любой близости от рассматриваемой точки (x_0, y_0) разность Δf может быть как положительной, так и отрицательной. Следовательно, в этой точке экстремума нет.

2) $\boldsymbol{A = 0}$. В этом случае

$$A\cos^2\varphi + 2B\cos\varphi\sin\varphi + C\sin^2\varphi = 2B\cos\varphi\sin\varphi + C\sin^2\varphi =$$
$$= \sin\varphi(2B\cos\varphi + C\sin\varphi).$$

Очевидно, $B \neq 0$ (в противном случае $AC - B^2 = 0$).

Выберем такой угол $\tilde{\varphi}$, чтобы

$$|C\sin\tilde{\varphi}| < |2B\cos\tilde{\varphi}|.$$

Тогда при $\varphi = \tilde{\varphi}$ и $\varphi = -\tilde{\varphi}$ трехчлен (27.6) будет иметь противоположные знаки. Поэтому (повторяя проведенные рассуждения) убеждаемся, что экстремума в точке $M_0(x_0, y_0)$ нет.

Теорема доказана.

В случае когда $AC - B^2 = 0$, вопрос об экстремуме остается открытым и для его решения требуется дополнительное исследование (например, с привлечением высших производных).

Заметим, что

$$D = \begin{vmatrix} f''_{xx}(x_0, y_0) & f''_{xy}(x_0, y_0) \\ f''_{yx}(x_0, y_0) & f''_{yy}(x_0, y_0) \end{vmatrix}.$$

Пример **27.2.** Исследовать на экстремум функцию

$$z = 3x^2 - x^3 + 3y^2 + 4y.$$

Решение: $z'_x = 6x - 3x^2 = 3(2x - x^2)$, $z'_y = 6y + 4 = 2(3y + 2)$. Получаем систему

$$\begin{cases} 2x - x^2 = 0, \\ 3y + 2 = 0. \end{cases}$$

Решая систему, находим две стационарные точки: $M_1\left(0, -\dfrac{2}{3}\right)$ и $M_2\left(2, -\dfrac{2}{3}\right)$.

Найдем частные производные второго порядка:

$$z''_{xx} = 6 - 6x, \quad z''_{xy} = 0, \quad z''_{yy} = 6.$$

Вычислим A, B, C и D для каждой стационарной точки:

• для точки $M_1\left(0, -\dfrac{2}{3}\right)$

$A_1 = 6$, $B_1 = 0$, $C_1 = 6$, $D_1 = 6 \cdot 6 - 0 = 36 > 0$ – экстремум есть;

$A_1 = 6 > 0$, следовательно, минимум;

$$z_{\min} = f\left(0, -\frac{2}{3}\right) = -\frac{4}{3};$$

• для точки $M_2\left(2, -\frac{2}{3}\right)$

$A_2 = -6,\ B_2 = 0,\ C_2 = 6,\ D_2 = -36 < 0$ – экстремума нет.

Пример **27.3.** Исследовать на экстремум функцию $z = x^3 + 8y^3 - 6xy + 1$.

Решение. Первые частные производные и стационарные точки (0, 0) и $\left(1, \frac{1}{2}\right)$ этой функции были найдены в примере 27.1. Так как $z''_{xx} = 6x,\ z''_{xy} = -6,\ z''_{yy} = 48y,$ то в точке (0, 0) будет $A = 0,$ $B = -6,\ C = 0,\ D = -36 < 0$, следовательно, экстремума нет. В точке $\left(1, \frac{1}{2}\right)$ имеем $A = 6,\ B = -6,\ C = 24,\ D = 108 > 0$ – экстремум есть; $A > 0$, следовательно, минимум;

$$z_{\min} = f\left(1, \frac{1}{2}\right) = 0.$$

27.2. НАИБОЛЬШЕЕ И НАИМЕНЬШЕЕ ЗНАЧЕНИЯ ФУНКЦИИ В ЗАМКНУТОЙ ОБЛАСТИ

Наибольшее и наименьшее значения (т.е. глобальный максимум и минимум) функции, непрерывной на некотором замкнутом множестве, могут достигаться или *в точках экстремума*, или *на границе множества*.

Пример **27.4.** Найти наибольшее и наименьшее значения функции $z = x^2 + y^2$ в круге радиуса 2 с центром в точке (0, 1).

Решение. Очевидно, граница области имеет уравнение $x^2 + (y-1)^2 = 4$.

Найдем частные производные: $z'_x = 2x,\ z'_y = 2y$. Приравнивая их к нулю, находим единственную стационарную точку $O(0, 0)$.

Исследуем функцию на границе области. Подставляя из уравнения границы $x^2 = 4 - (y-1)^2$ в функцию $z = x^2 + y^2,$ получаем функцию одной переменной $z = 4 - (y-1)^2 + y^2,$ т.е. $z = 2y + 3$. Очевидно, $y \in [-1, 3]$. Эта функция не имеет стационарных точек, поэтому ее наибольшее и наименьшее значения могут достигаться только на концах отрезка $[-1, 3]$.

Значение функции $z = x^2 + y^2$ в стационарной точке (0, 0) равно 0. Значение функции $z = 2y + 3$ при $y = -1$ равно 1, а при $y = 3$ равно 9. Сравнивая эти три значения, находим $z_{\text{наиб}} = f(0, 3) = 9$, $z_{\text{наим}} = f(0, 0) = 0$.

27.3. УСЛОВНЫЙ ЭКСТРЕМУМ

Рассмотрим задачу нахождения экстремумов функции н е с к о л ь к и х аргументов при наличии *дополнительных условий*, связывающих значения аргументов. Такие экстремумы называются **условными**.

Например, пусть необходимо найти экстремумы функции

$$z = x^2 y, \tag{*}$$

если ее аргументы удовлетворяют условию

$$2x + y - 1 = 0. \tag{**}$$

В этом случае экстремумы ищутся не на всей плоскости Oxy, а только на прямой $2x + y - 1 = 0$. Подставим в (*) выражение $y = -2x + 1$ из условия (**), и задача об условном экстремуме функции (*) сведется к задаче нахождения безусловного экстремума функции $z = x^2(-2x + 1) = x^2 - 2x^3$. Итак, $z' = 2x - 6x^2 = 2x(1 - 3x)$. Таким образом, исходная функция имеет минимум при $x = 0$ и максимум при $x = \frac{1}{3}$, т.е. функция $z = x^2 y$ при наличии связи (**) имеет условный минимум $z = 0$ в точке $(0, 1)$ и условный максимум $z = \frac{1}{27}$ в точке $\left(\frac{1}{3}, \frac{1}{3}\right)$.

Определение. Функция $u = f(M) = f(x_1, x_2, \ldots, x_n)$ имеет **условный максимум (минимум)** в точке $M_0(x_1^0, x_2^0, \ldots, x_n^0)$, если существует такая окрестность точки M_0, что для всех точек $M(x_1, x_2, \ldots, x_n)$ этой окрестности, удовлетворяющих m уравнениям $(m < n)$:

$$\begin{cases} g_1(x_1, \ldots, x_n) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ g_m(x_1, \ldots, x_n) = 0, \end{cases} \tag{27.7}$$

выполняется неравенство $f(M_0) > f(M)$ (соответственно $f(M_0) < f(M)$).

Уравнения (27.7) называются **уравнениями связи**.

Задача нахождения условного экстремума сводится к исследованию на обычный экстремум функции

$$L(x_1, ..., x_n, \ \lambda_1, ..., \lambda_m) =$$

$$= f(x_1, ..., x_n) + \lambda_1 g_1(x_1, ..., x_n) + ... + \lambda_m g_m(x_1, ..., x_n).$$

Функция L называется **функцией Лагранжа**, а числа $\lambda_1, ..., \lambda_m$ – **множителями Лагранжа**.

Необходимые условия условного экстремума выражаются системой $m + n$ уравнений

$$\begin{cases} \dfrac{\partial L(M)}{\partial x_i} = 0, \quad i = 1, 2, ..., n, \\ g_k(M) = 0, \quad k = 1, 2, ..., m, \end{cases} \tag{27.8}$$

из которой могут быть найдены неизвестные $x_1, ..., x_n$; $\lambda_1, ..., \lambda_m$, где $x_1, ..., x_n$ – координаты точки, в которой возможен условный экстремум.

Достаточные условия условного экстремума связаны с изучением второго дифференциала функции Лагранжа d^2L: если в точке возможного экстремума M_0 выполняется неравенство $d^2L < 0$, то в этой точке имеет место *условный максимум*, если же $d^2L > 0$, то *условный минимум*.

В случае функции двух переменных $z = f(x, y)$ при уравнении связи $g(x, y) = 0$ функция Лагранжа имеет вид

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y).$$

Система (27.8) состоит из трех уравнений:

$$\frac{\partial L}{\partial x} = 0, \qquad \frac{\partial L}{\partial y} = 0, \qquad g(x, y) = 0.$$

Пусть (x_0, y_0, λ_0) – любое из решений этой системы, $M_0(x_0, y_0)$ – точка возможного экстремума и

$$\Delta = -\begin{vmatrix} 0 & g'_x(M_0) & g'_y(M_0) \\ g'_x(M_0) & L''_{xx}(M_0, \lambda_0) & L''_{xy}(M_0, \lambda_0) \\ g'_y(M_0) & L''_{yx}(M_0, \lambda_0) & L''_{yy}(M_0, \lambda_0) \end{vmatrix}.$$

Тогда если $\Delta < 0$, то функция $z = f(x, y)$ имеет в точке $M_0(x_0, y_0)$ *условный максимум*, если $\Delta > 0$, то *условный минимум*.

***Пример* 27.5.** Найти условный экстремум функции $z = 2x + y$ при $x^2 + y^2 = 1$.

Решение. Составим функцию Лагранжа:

$L(x, y, \lambda) = 2x + y + \lambda(x^2 + y^2 - 1)$.

Имеем $\dfrac{\partial L}{\partial x} = 2 + 2\lambda x$, $\dfrac{\partial L}{\partial y} = 1 + 2\lambda y$.

Система уравнений (27.8) имеет вид

$$\begin{cases} 2 + 2\lambda x = 0, \\ 1 + 2\lambda y = 0, \\ x^2 + y^2 - 1 = 0. \end{cases}$$

Находим решения этой системы: $x_1 = -\dfrac{2}{\sqrt{5}}$, $y_1 = -\dfrac{1}{\sqrt{5}}$, $\lambda_1 = \dfrac{\sqrt{5}}{2}$; $x_2 = \dfrac{2}{\sqrt{5}}$, $y_2 = \dfrac{1}{\sqrt{5}}$, $\lambda_2 = -\dfrac{\sqrt{5}}{2}$. Имеем:

$$g(x, y) = x^2 + y^2 - 1, \qquad g'_x = 2x, \qquad g'_y = 2y,$$

$$g'_x\left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right) = -\frac{4}{\sqrt{5}}, \qquad g'_y\left(-\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right) = -\frac{2}{\sqrt{5}},$$

$$L''_{xx} = \sqrt{5},\ L''_{xy} = 0,\ L''_{yy} = \sqrt{5} \quad \text{при} \quad \lambda = \frac{\sqrt{5}}{2}.$$

Следовательно,

$$\Delta = -\begin{vmatrix} 0 & -\frac{4}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ -\frac{4}{\sqrt{5}} & \sqrt{5} & 0 \\ -\frac{2}{\sqrt{5}} & 0 & \sqrt{5} \end{vmatrix} = 4\sqrt{5} > 0,$$

т.е. функция имеет условный минимум в точке $M_1\left(-\dfrac{2}{\sqrt{5}}, -\dfrac{1}{\sqrt{5}}\right)$, $z_{\min} = -\sqrt{5}$.

Аналогично для точки $M_2\left(\dfrac{2}{\sqrt{5}}, \dfrac{1}{\sqrt{5}}\right)$:

$$\Delta = -\begin{vmatrix} 0 & \frac{4}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{4}{\sqrt{5}} & -\sqrt{5} & 0 \\ \frac{2}{\sqrt{5}} & 0 & -\sqrt{5} \end{vmatrix} = -4\sqrt{5} < 0,$$

т.е. в точке $M_2\left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ функция имеет условный максимум, $z_{\max} = \sqrt{5}$.

27.4. МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Метод наименьших квадратов применяется в теории ошибок. Он относится к так называемым методам аппроксимации, т.е. приближенного выражения каких-либо математических объектов через другие, более простые.

На практике мы часто сталкиваемся с необходимостью «сгладить» зависимости, выявленные в результате наблюдений. Обычно задача формулируется так: имеются данные наблюдений в n точках $M_1, M_2, ..., M_n$ некоторой величины u и получены соответствующие значения этой величины $u_1, u_2, ..., u_n$; надо подобрать такую функцию $u = f(M)$, чтобы она наиболее точно выражала общую зависимость измеряемой величины от параметров точек измерений $M_1, M_2, ..., M_n$.

Формулы, аналитически представляющие опытные данные (или результаты измерений), называются **эмпирическими**.

Рассмотрим для простоты случай, когда точки M_i, в которых проводятся измерения, имеют одну координату x_i, т.е. зависимость между переменными x и y представлена в виде набора $x_1, x_2, ..., x_n$ и соответствующих значений $y_1, y_2, ..., y_n$. Эти пары значений изображаются на координатной плоскости точками $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. Ломаная, которая соединяет эти точки, называется **экспериментальной кривой** (рис. 27.1).

Необходимо найти аналитическое представление зависимости между x и y в виде формулы $y = f(x)$. Вид функции $y = f(x)$ опре-

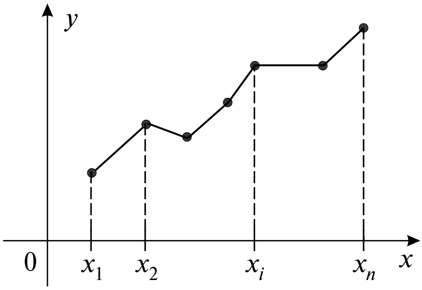


Рис. 27.1. Экспериментальная кривая

деляется экономическими или иными соображениями. Обычно в качестве таких функций используются следующие:

- $y = ax + b$ — *линейная*;
- $y = ax^2 + bx + c$ — *параболическая*;
- $y = \dfrac{a}{x} + b$ — *гиперболическая*;
- $y = ae^{bx}$ — *экспоненциальная*.

(Используются также *логарифмическая*, *степенная* и другие функции.)

Задача нахождения эмпирических формул обычно решается в два этапа.

На первом этапе определяют *общий вид* зависимости $y = f(x)$, т.е. необходимо решить, является ли она линейной, квадратичной, экспоненциальной или еще какой-нибудь.

Предположим, что результаты измерений (экспериментальные данные) нанесены на координатную сетку (рис. 27.2).

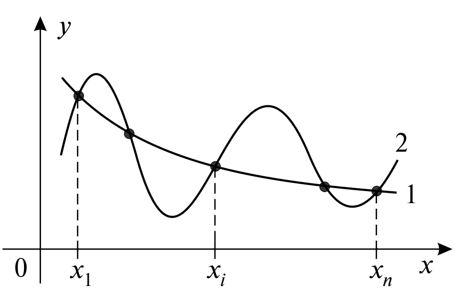


Рис. 27.2. Определение общего вида зависимости $y = f(x)$

Очевидно, существует много различных кривых, проходящих через эти точки. В случае, изображенном на рис. 27.2, кривая *1* для исследователя предпочтительнее кривой *2*. Следует подчеркнуть, что первый этап – этап подбора вида эмпирической функции – весьма важен. Мы видим, что кривая *2* на рис. 27.2, хотя и проходит через соответствующие точки, но не дает удовлетворительного представления зависимости между x и y.

На практике для проверки правильности выбора функции $y = f(x)$ проводятся дополнительные исследования, т.е. осуществляются дополнительные измерения величин x, y и на координатную плоскость наносятся дополнительные точки. Если они оказываются на достаточно близком расстоянии от выбранной кривой, то считают, что вид кривой установлен, т.е. установлен вид функции $y = f(x)$. После выбора вида функции переходят ко второму этапу.

На втором этапе определяют *параметры* выбранной эмпирической функции $y = f(x)$. В указанных выше функциях (см. с. 328) параметрами являются неизвестные числа a, b и c. Параметры следует выбрать такими, чтобы значения эмпирической функции менее всего отклонялись бы в точках $x_1, x_2, ..., x_n$ от измеренных значений.

Метод наименьших квадратов (предложенный К. Гауссом) заключается в том, чтобы минимизировать сумму квадратов отклонений «теоретических» значений $f(x_i)$, найденных по эмпирической формуле $y = f(x)$, от соответствующих экспериментальных значений y_i. Иначе говоря, величина

$$S = \sum_{i=1}^{n} \delta_i^2 = \sum_{i=1}^{n} (f(x_i) - y_i)^2$$

должна быть минимальной (рис. 27.3).

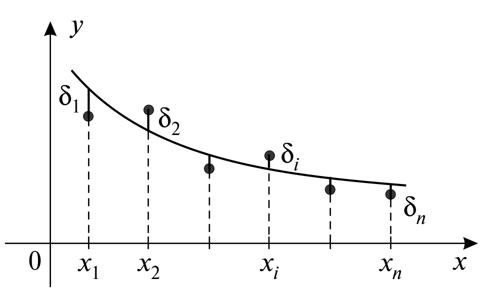


Рис. 27.3. Иллюстрация метода наименьших квадратов

Проиллюстрируем общий метод наименьших квадратов на примере линейной функции. Итак, пусть в качестве $y=f(x)$ взята функция $y=ax+b$ и необходимо найти такие значения неизвестных параметров a и b, при которых функция

$$S=\sum_{i=1}^{n}(ax_i+b-y_i)^2$$

принимает наименьшее значение. Здесь x_i и y_i – постоянные, найденные экспериментально, а функция S есть функция параметров a и b: $S=S(a, b)$.

Итак, найдем критические точки функции $S(a, b)$, а затем исследуем их. Для нахождения критических точек надо решить систему

$$\begin{cases} S'_a(a,b)=0, \\ S'_b(a,b)=0, \end{cases}$$

или, что то же самое,

$$\begin{cases} \sum\limits_{i=1}^{n}2(ax_i+b-y_i)\,x_i=0, \\ \sum\limits_{i=1}^{n}2(ax_i+b-y_i) \quad =0. \end{cases}$$

После очевидных элементарных преобразований получим эквивалентную ей систему, называемую **системой нормальных уравнений**,

$$\begin{cases} \left(\sum\limits_{i=1}^{n}x_i^2\right)a+\left(\sum\limits_{i=1}^{n}x_i\right)b=\sum\limits_{i=1}^{n}x_iy_i, \\ \left(\sum\limits_{i=1}^{n}x_i\right)a+ \qquad nb=\sum\limits_{i=1}^{n}y_i. \end{cases} \tag{27.9}$$

Эта система линейна относительно неизвестных a и b. Определитель этой системы отличен от нуля:

$$d=\begin{vmatrix} \sum\limits_{i=1}^{n}x_i^2 & \sum\limits_{i=1}^{n}x_i \\ \sum\limits_{i=1}^{n}x_i & n \end{vmatrix}=n\sum_{i=1}^{n}x_i^2-\left(\sum_{i=1}^{n}x_i\right)^2\neq 0. \tag{27.10}$$

(Можно даже доказать, что этот определитель положителен.)

Поэтому система имеет единственное решение, которое можно найти по правилу Крамера:

$$a^* = \frac{1}{d}\begin{vmatrix} \sum x_i y_i & \sum x_i \\ \sum y_i & n \end{vmatrix}, \qquad b^* = \frac{1}{d}\begin{vmatrix} \sum x_i^2 & \sum x_i y_i \\ \sum x_i & \sum y_i \end{vmatrix}. \qquad (27.11)$$

Итак, мы нашли единственную критическую точку (a^*, b^*). Убедимся в том, что в ней достигается минимум функции $S(a, b)$. Для этого вычислим вторые частные производные:

$$S''_{aa} = 2\sum_{i=1}^{n} x_i^2 = A, \quad S''_{ab} = 2\sum_{i=1}^{n} x_i = B, \quad S''_{bb} = 2n = C.$$

Имеем

$$D = AC - B^2 = 4n\sum_{i=1}^{n} x_i^2 - 4\left(\sum_{i=1}^{n} x_i\right)^2 = 4d.$$

Ранее мы отметили, что $d > 0$. Следовательно, $D > 0$, поэтому, согласно достаточному условию экстремума, в рассматриваемой точке (a^*, b^*) экстремум есть. Так как $A = 2\sum_{i=1}^{n} x_i^2 > 0$, то этот экстремум – минимум. Из всего сказанного заключаем, что функция $S = S(a, b)$ имеет единственную точку минимума (a^*, b^*), определяемую из системы нормальных уравнений. Следует заметить, что в этой точке имеется не только локальный, но и глобальный минимум, т.е. наименьшее значение функции.

***Пример* 27.6.** Получены следующие данные о стоимости основных фондов x (тыс. усл. ед.) и прибыли предприятия y (тыс. усл. ед.):

x_i	110	132	154	176	198	220	242
y_i	40	43,2	52,8	67,2	64	78,4	96

Предполагая, что между переменными x и y существует линейная зависимость, найти эмпирическую формулу, используя метод наименьших квадратов.

Решение. Для определения неизвестных параметров a^* и b^* эмпирической формулы $y = a^*x + b^*$ применим формулы (27.11). Нам понадобится предварительно вычислить суммы $\sum_{i=1}^{n} x_i$, $\sum_{i=1}^{n} y_i$,

$\sum_{i=1}^{n} x_i y_i, \sum_{i=1}^{n} x_i^2$ (здесь $n = 7$). Для удобства сведем вычисления в таблицу:

i	x_i	y_i	x_i^2	$x_i y_i$
1	110	40	12 100	4400
2	132	43,2	17 424	5702,4
3	154	52,8	23 716	8131,2
4	176	67,2	30 976	11 827,2
5	198	64	39 204	12 672
6	220	78,4	48 400	17 248
7	242	96	58 564	23 232
Σ	1232	441,6	230 384	83 212,8

Система нормальных уравнений (27.9) имеет вид

$$\begin{cases} 230384a + 1232b = 83212{,}8, \\ 1232a + 7b = 441{,}6. \end{cases}$$

Находим:

$$d = \begin{vmatrix} 230384 & 1232 \\ 1232 & 7 \end{vmatrix} = 230384 \cdot 7 - 1232^2 = 94864,$$

$$a^* = \frac{1}{94864} \begin{vmatrix} 83212{,}8 & 1232 \\ 441{,}6 & 7 \end{vmatrix} = \frac{38438{,}4}{94864} = 0{,}405,$$

$$b^* = \frac{1}{94864} \begin{vmatrix} 230384 & 83212{,}8 \\ 1232 & 441{,}6 \end{vmatrix} = \frac{-780595{,}2}{94864} = -8{,}229.$$

Таким образом, искомая зависимость имеет вид

$$y = 0{,}405x - 8{,}229.$$

Вопросы

1. Что такое локальный максимум (минимум) функции двух переменных?
2. Если $f'_x(x_0, y_0) = 0$, то можно ли утверждать, что (x_0, y_0) – точка экстремума для $f(x, y)$?
3. Что такое критическая (стационарная) точка для функции двух переменных?

4. В чем заключается достаточное условие экстремума для функции двух переменных?
5. Что такое условный экстремум функции n переменных?
6. Что такое функция Лагранжа?
7. Что такое эмпирические формулы? Какая линия называется экспериментальной кривой?
8. Из скольких этапов состоит обычно решение задачи нахождения эмпирических формул? Каковы эти этапы?
9. В чем заключается метод наименьших квадратов?

Глава 28
ЗАДАЧИ ОПТИМИЗАЦИИ

28.1. ОСНОВНЫЕ ПОНЯТИЯ

Для упрощения изложения отождествим точки евклидова пространства $\mathbf{R}^n$ и векторы их координат, т.е. точку $M(x_1, x_2, ..., x_n)$ будем записывать как $\overline{x}$, где $\overline{x} = (x_1, x_2, ..., x_n)$. При этом на точки пространства $\mathbf{R}^n$ переносятся все операции над векторами и их свойства. Пусть $\overline{x}$ и $\overline{y}$ – две точки евклидова пространства.

Определение. **Прямой**, проходящей через точки $\overline{x}$ и $\overline{y}$ евклидова пространства $\mathbf{R}^n$, называется множество точек

$$\overline{x} + t(\overline{y} - \overline{x}) \tag{28.1}$$

этого пространства, где $t \in (-\infty, \infty)$. **Отрезком** $\overline{xy}$, соединяющим эти точки, называется множество (28.1), где $t \in [0, 1]$.

Очевидно, сами точки $\overline{x}$ и $\overline{y}$ получаются из (28.1) при $t = 0$ и $t = 1$ соответственно.

Определение. Множество D точек евклидова пространства называется **выпуклым**, если вместе с его любыми двумя точками $\overline{x}$ и $\overline{y}$ все точки отрезка $\overline{xy}$ также принадлежат этому множеству.

Примеры выпуклых и невыпуклых плоских множеств показаны на рис. 28.1.

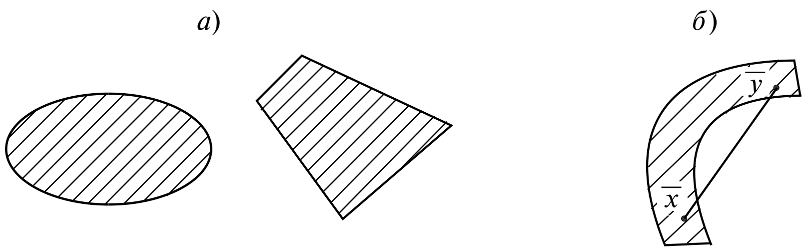


Рис. 28.1. Множества: выпуклые (*а*) и невыпуклое (*б*)

Определение. Функция $f(\overline{x})$, заданная на выпуклом множестве $D \subset \mathbf{R}^n$, называется **выпуклой** на D, если для любых двух точек $\overline{x}$ и $\overline{y}$ из D и любого $\alpha \in [0, 1]$ выполняется неравенство

$$f(\alpha\overline{x} + (1-\alpha)\overline{y}) \le \alpha f(\overline{x}) + (1-\alpha) f(\overline{y}). \qquad (28.2\text{а})$$

Функция $f(\overline{x})$ называется **вогнутой** на выпуклом множестве D, если для любых двух точек $\overline{x}$ и $\overline{y}$ из D и любого $\alpha \in [0, 1]$ выполняется неравенство

$$f(\alpha\overline{x} + (1-\alpha)\overline{y}) \ge \alpha f(\overline{x}) + (1-\alpha) f(\overline{y}). \qquad (28.2\text{б})$$

Если неравенства (28.2) заменить на строгие неравенства, то получим определение **строго выпуклой** и **строго вогнутой** функций соответственно.

Заметим, что график выпуклой функции одной переменной обращен выпуклостью вниз, а график вогнутой функции – выпуклостью вверх.

Имеет место следующее важное утверждение (приведем его без доказательства).

***Теорема* 28.1.** Если функция $f(\overline{x})$ дифференцируема и строго вогнута (строго выпукла) на выпуклом множестве D, то она имеет локальный экстремум лишь в одной точке этого множества.

28.2. НАИБОЛЬШЕЕ ЗНАЧЕНИЕ ВОГНУТОЙ ФУНКЦИИ. УСЛОВИЯ КУНА – ТАККЕРА

Мы знаем, что линейной функцией n переменных называется функция $l(\overline{x}) = a_1x_1 + a_2x_2 + \ldots + a_nx_n + b$, где $a_1, a_2, \ldots, a_n, b$ – константы. Нетрудно убедиться, что линейная функция является одновременно выпуклой и вогнутой в $\mathbf{R}^n$.

Неравенство вида $l(\overline{x}) \ge 0$, где $l(\overline{x})$ – линейная функция, также называется **линейным**.

***Теорема* 28.2.** Пусть выпуклое множество $D \subset \mathbf{R}^n$ задано системой линейных неравенств

$$\begin{cases} l_1(\overline{x}) \ge 0, \\ \ldots\ldots\ldots\ldots \\ l_m(\overline{x}) \ge 0, \end{cases} \qquad (28.3)$$

D' – некоторое выпуклое подмножество в D; $f(\bar{x})$ – функция, вогнутая на D' и дифференцируемая в точке $\bar{x}^0 \in D'$. Тогда:

1) если для некоторых чисел $\lambda_1, ..., \lambda_m$ выполняются условия:

$$\left.\begin{array}{l} \text{а) } \bar{x}^0 \text{ – критическая точка функции} \\ \quad L(\bar{x}) = f(\bar{x}) + \lambda_1 l_1(\bar{x}) + ... + \lambda_m l_m(\bar{x}); \\ \text{б) } \lambda_i l_i(\bar{x}^0) = 0 \text{ и } \lambda_i \geq 0,\ i = 1, ..., m, \end{array}\right\} \textbf{(условия Куна – Таккера)}$$

то $f(\bar{x}^0)$ – н а и б о л ь ш е е значение $f(\bar{x})$ на D;

2) напротив, если $D' = D$ и $f(\bar{x}^0)$ – наибольшее значение $f(\bar{x})$ на D, то с у щ е с т в у ю т числа $\lambda_1, ..., \lambda_m$, для которых выполняются условия «а» и «б».

Функция $L(\bar{x})$ – это уже знакомая нам функция Лагранжа (см. § 27.3), а числа $\lambda_1, ..., \lambda_m$ – множители Лагранжа.

Теорему 28.2 также принимаем без доказательства.

Пример **28.1.** Найти точку наибольшего значения (глобального максимума) функции $u = \ln x_1 + \ln x_2 + \ln x_3$ при условии $x_1 + 4x_2 + 9x_3 \leq 108$.

Р е ш е н и е. Заметим прежде всего, что функция u вогнута. Действительно, логарифмическая функция одной переменной вогнута, а сумма вогнутых функций, как легко проверить, есть вогнутая функция.

Составим функцию Лагранжа:

$L(\bar{x}) = \ln x_1 + \ln x_2 + \ln x_3 + \lambda(108 - x_1 - 4x_2 - 9x_3)$.

Условия Куна – Таккера имеют вид

$$\begin{cases} \dfrac{1}{x_1} - \lambda = 0, \\ \dfrac{1}{x_2} - 4\lambda = 0, \\ \dfrac{1}{x_3} - 9\lambda = 0, \\ \lambda(108 - x_1 - 4x_2 - 9x_3) = 0, \\ \lambda \geq 0. \end{cases}$$

Из первых трех условий находим: $x_1 = \dfrac{1}{\lambda}$, $x_2 = \dfrac{1}{4\lambda}$, $x_3 = \dfrac{1}{9\lambda}$. При этом, очевидно, $\lambda \neq 0$. Подставляя найденные значения x_1, x_2 и x_3 в четвертое уравнение, получаем

$$108 - x_1 - 4x_2 - 9x_3 = 108 - \frac{3}{\lambda} = 0.$$

Отсюда $\lambda=\frac{1}{36}$, $x_1=36$, $x_2=9$, $x_3=4$. Как уже отмечалось, функция u вогнута, поэтому точка $\overline{x}^0 = (36, 9, 4)$ есть точка глобального максимума.

Максимизация прибыли

Пусть $F(K, L)$ – производственная функция (где K и L – соответственно затраты капитала и трудовых ресурсов), P – цена продукции. Функция прибыли Π вычисляется обычно по формуле

$$\Pi(K, L) = P\cdot F(K, L) - WL - RK, \tag{28.4}$$

где W и R – соответственно цены на труд и капитальные затраты, W и R – положительные числа.

Точка (K_0, L_0) называется **оптимальным планом** (см. § 10.1), если в ней функция прибыли (28.4) принимает м а к с и м а л ь н о е значение.

Рассмотрим задачу: найти предельную норму замещения производственной функции F

$$\mu=-\frac{F'_L}{F'_K}$$

при оптимальном плане.

В точке локального максимума первые частные производные функции прибыли $\Pi(K, L)$ равны нулю. Система (27.2) в данном случае имеет вид

$$\begin{cases} P\cdot F'_K(K_0, L_0)-R=0,\\ P\cdot F'_L(K_0, L_0)-W=0. \end{cases}$$

Отсюда $\mu=-\frac{W}{R}$.

Рассмотрим теперь задачу максимизации функции прибыли.

***П р и м е р* 28.2.** Найти оптимальный план и максимум функции прибыли (28.4), если производственная функция имеет вид $F(K, L) = 3K^{1/3}L^{1/3}$.

Р е ш е н и е. Функция прибыли в данном случае имеет вид

$$\Pi(K, L) = 3P\cdot K^{1/3}L^{1/3} - WL - RK.$$

Вычисляем первые частные производные по K и L и приравниваем их к нулю:

$$\begin{cases} P\cdot K^{-2/3}L^{1/3}-R=0,\\ P\cdot K^{1/3}L^{-2/3}-W=0. \end{cases}$$

Отсюда находим координаты оптимального плана:

$$K_0 = \frac{P^3}{R^2 W}, \qquad L_0 = \frac{P^3}{R W^2}.$$

Подставляя эти значения в функцию прибыли, получим

$$\Pi_{\max} = \frac{P^3}{RW}.$$

Оптимизация спроса

Рассмотрим задачу оптимизации функции полезности при ограничениях на доход потребителя.

Пример **28.3.** Найти величины спроса x и y на две разновидности товара при ценах на них соответственно p и q, если доход потребителя равен M, функция полезности имеет вид $U(x, y) = x^{\frac{p}{p+q+1}} y^{\frac{q}{p+q+1}}$ и потребитель стремится максимизировать функцию полезности.

Решение. Из условия следует, что потребитель может покупать только такие наборы (x, y), стоимость которых не превышает его дохода, т.е.

$$px + qy \le M, \quad x \ge 0, \quad y \ge 0. \tag{28.5}$$

Ограничения (28.5) задают на плоскости замкнутую область в виде треугольника (рис. 28.2). Надо найти точку максимума функции $U(x, y)$. Вычислим частные производные функции полезности:

$$U'_x(x, y) = \frac{p}{p+q+1} x^{-\frac{q+1}{p+q+1}} y^{\frac{q}{p+q+1}};$$

$$U'_y(x, y) = \frac{q}{p+q+1} x^{\frac{p}{p+q+1}} y^{-\frac{p+1}{p+q+1}}.$$

Видим, что внутри области критических точек нет. Стало быть, максимум может достигаться только на границе. На прямых $x = 0$, $y = 0$ функция полезности равна нулю: $U(0, y) = U(x, 0) = 0$, поэтому надо искать точку максимума на прямой $px + qy = M$. Отсюда

$$y = \frac{M - px}{q}. \tag{*}$$

Подставив это выражение y в $U(x, y)$, получим функцию одной переменной x:

$$U\left(x, \frac{M-px}{q}\right) = f(x) = q^{-\frac{q}{p+q+1}} x^{\frac{p}{p+q+1}} (M-px)^{\frac{q}{p+q+1}}.$$

Вычислим $f'(x)$:

$$f'(x) = q^{-\frac{q}{p+q+1}} \left[\frac{p}{p+q+1} x^{-\frac{q+1}{p+q+1}} (M-px)^{\frac{q}{p+q+1}} + \right.$$

$$\left. + \frac{q}{p+q+1} x^{\frac{p}{p+q+1}} (M-px)^{-\frac{p+1}{p+q+1}} (-p) \right].$$

Приравнивая $f'(x)$ к нулю, после преобразований получаем

$$M - px - qx = 0,$$

откуда $x = \dfrac{M}{p+q}$ и с учетом (*) $y = \dfrac{M}{p+q}$.

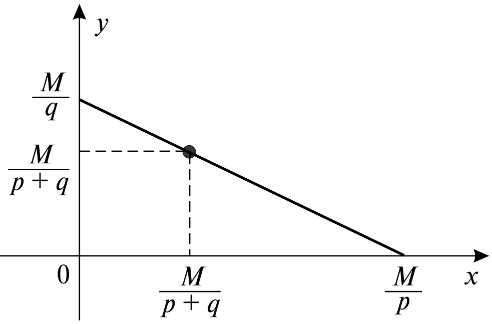


Рис. 28.2. Максимум функции полезности

Заметим, что эту задачу можно было бы решать, выписав функцию Лагранжа и условия Куна – Таккера, однако для такого простого случая в этом не было необходимости.

Вопросы

1. Что такое выпуклое множество в евклидовом пространстве?
2. Какая функция, заданная на выпуклом множестве, называется выпуклой (вогнутой)?

3. Сколько локальных экстремумов имеет строго выпуклая функция на выпуклом множестве?
4. Может ли функция быть одновременно выпуклой и вогнутой?
5. Что такое условия Куна – Таккера?
6. Что такое функция прибыли? Как она вычисляется?
7. Что называется оптимальным планом?

Раздел VIII
ДИФФЕРЕНЦИАЛЬНЫЕ И РАЗНОСТНЫЕ УРАВНЕНИЯ

Глава 29
ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ПЕРВОГО ПОРЯДКА И ИХ ПРИМЕНЕНИЕ В НЕПРЕРЫВНЫХ МОДЕЛЯХ ЭКОНОМИКИ

29.1. ОСНОВНЫЕ ПОНЯТИЯ

Определение. **Дифференциальным уравнением** называется уравнение

$$F(x, y, y', \ldots, y^{(n)}) = 0, \tag{29.1}$$

которое связывает неизвестную функцию y, ее независимый аргумент x и ее производные $y', \ldots, y^{(n)}$. **Порядком** дифференциального уравнения называется наибольший порядок производной, входящей в уравнение.

Например, $y' - \frac{2}{x}y = \mathrm{e}^x x^2$ есть дифференциальное уравнение первого порядка, а $y'' + 4y = 0$ – дифференциальное уравнение второго порядка.

Определение. **Решением дифференциального уравнения** называется функция $y = \varphi(x)$, которая при подстановке в уравнение (29.1) превращает его в тождество.

В этой главе будем рассматривать дифференциальные уравнения *первого* порядка, т.е. уравнения вида

$$F(x, y, y') = 0. \tag{29.2}$$

В случае когда из уравнения (29.2) можно выразить y', оно имеет вид

$$y' = f(x, y). \tag{29.3}$$

Уравнение (29.3) называется **уравнением первого порядка**, **разрешенным относительно производной**.

В теории дифференциальных уравнений основной задачей является вопрос о *существовании* и *единственности решения*. Приведем без доказательства теорему, которая дает ответ на этот вопрос.

***Теорема* 29.1 (теорема Коши).** Пусть дано дифференциальное уравнение (29.3) и пусть функция $f(x, y)$ и ее частная производная $f'_y(x, y)$ непрерывны в некоторой области D плоскости Oxy. Тогда в некоторой окрестности любой внутренней точки $M(x_0, y_0) \in D$ существует единственное решение уравнения (29.3), удовлетворяющее условию $y = y_0$ при $x = x_0$.

График решения дифференциального уравнения называется **интегральной кривой**. В области D содержится бесконечное множество интегральных кривых. Теорема Коши утверждает, что при соблюдении определенных условий через каждую внутреннюю точку области D проходит только одна интегральная кривая.

Условия, которые задают значение функции y в фиксированной точке x_0, называются **начальными условиями** (или **условиями Коши**) и записываются в форме $y(x_0) = y_0$ или в форме

$$y\big|_{x=x_0} = y_0. \tag{29.4}$$

Задача нахождения решения уравнения (29.3), удовлетворяющего условию (29.4), называется **задачей Коши**.

Рисунок 29.1 иллюстрирует теорему 29.1. Вся область D заполнена интегральными кривыми, при этом они не могут ни пересекаться, ни касаться друг друга.

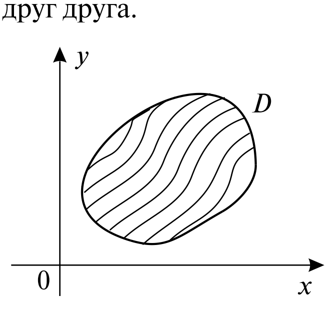


Рис. 29.1. Интегральные кривые

Теорема 29.1 позволяет описать множество решений дифференциального уравнения в виде общего решения.

Определение. **Общим решением** дифференциального уравнения (29.2) называется функция

$$y = \varphi(x, C), \tag{29.5}$$

зависящая от x и произвольной постоянной C, если выполняются следующие условия:

1) для любого значения константы C функция (29.5) является решением дифференциального уравнения (29.2);

2) каково бы ни было начальное условие (29.4), существует такое значение $C = C_0$, что функция $y = \varphi(x, C_0)$ удовлетворяет данному начальному условию.

Общее решение, записанное в неявном виде $\Phi(x, y, C) = 0$, называется **общим интегралом**.

Определение. Если в общем решении (29.5) зафиксирована константа $C = C_0$, то (29.5) называется **частным решением**.

Частное решение, представленное в неявном виде, называется **частным интегралом**.

Решить дифференциальное уравнение – значит найти его общее решение или общий интеграл.

29.2. ВИДЫ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ПЕРВОГО ПОРЯДКА И МЕТОДЫ ИХ РЕШЕНИЯ

Уравнения с разделяющимися переменными

Дифференциальное уравнение первого порядка называется **уравнением с разделяющимися переменными**, если оно может быть представлено в виде

$$\frac{dy}{dx} = f(x)g(y). \tag{29.6}$$

Метод решения такого вида уравнений носит название «**разделение переменных**». Умножим обе части уравнения (29.6) на dx и поделим на $g(y)$, полагая $g(y) \neq 0$:

$$\frac{dy}{g(y)} = f(x)dx. \tag{29.7}$$

Это **уравнение с разделенными переменными**. Поскольку дифференциалы равны, то равны и неопределенные интегралы (точнее, отличаются на константу), поэтому

$$\int \frac{dy}{g(y)} = \int f(x)dx + C,$$

где C – произвольная постоянная.

Пример **29.1.** Решить уравнение $xy' - y = 0$.

Решение. Разделим переменные. Для этого представим уравнение в виде

$$x\frac{dy}{dx} = y,$$

умножим обе части на dx и разделим на xy:

$$\frac{dy}{y} = \frac{dx}{x}.$$

Интегрируем обе части этого уравнения:

$$\ln|y| = \ln|x| + C_1.$$

Представим произвольную постоянную C_1 в виде $C_1 = \ln|C|$, тогда общий интеграл будет иметь вид

$$\ln|y| = \ln|x| + \ln|C|.$$

Отсюда получаем общее решение $y = \pm Cx$, или, заменяя $\pm C$ на C:

$$y = Cx.$$

Пример **29.2.** Решить дифференциальное уравнение $y' = -\frac{2xy^2}{x^2-1}$ и найти частное решение, удовлетворяющее начальному условию $y(0) = 1$.

Решение. Разделим переменные:

$$\frac{dy}{dx} = -\frac{2xy^2}{x^2-1},$$

$$\frac{dy}{y^2} = -\frac{2x\,dx}{x^2-1}.$$

Интегрируя, получаем

$$\frac{1}{y} = \ln|x^2-1| + C,$$

отсюда общий интеграл

$$y(\ln|x^2-1| + C) = 1.$$

Подставим начальные условия $y(0) = 1$: $1 \cdot (0 + C) = 1$. Получаем $C = 1$. Отсюда

$$y = \frac{1}{\ln|x^2 - 1| + 1}.$$

Однородные дифференциальные уравнения первого порядка

Функция $f(x, y)$ называется **однородной функцией *n*-го измерения**, если при любом λ

$$f(x, y) = \lambda^n f(x, y).$$

Например, функция $f(x, y) = xy - y^2$ является однородной функцией второго измерения, так как $\lambda x \lambda y - (\lambda y)^2 = \lambda^2(xy - y^2)$, а функция $f(x, y) = \frac{y}{x}$ – нулевого измерения, так как $\frac{\lambda y}{\lambda x} = \frac{y}{x}$.

Дифференциальное уравнение

$$y' = f(x, y)$$

называется **однородным**, если $f(x, y)$ – однородная функция нулевого измерения.

Данное уравнение можно решить следующим образом. Преобразуем его правую часть, положив $\lambda = \frac{1}{x}$: $f(x, y) = f(\lambda x, \lambda y) = f\left(1, \frac{y}{x}\right)$. Уравнение приобретает вид

$$y' = f\left(1, \frac{y}{x}\right). \qquad (29.8)$$

Сделаем подстановку $u = \frac{y}{x}$, т.е. $y = ux$. Тогда $y' = u'x + u$. Подставляя это выражение производной в уравнение (29.8), получаем

$$u + x\frac{du}{dx} = f(1, u).$$

Это уравнение с разделяющимися переменными:

$$x\frac{du}{dx} = f(1, u) - u, \quad \text{или} \quad \frac{du}{f(1, u) - u} = \frac{dx}{x}.$$

Отсюда

$$\int \frac{du}{f(1, u) - u} = \int \frac{dx}{x} + C.$$

После нахождения функции $u = u(x)$ необходимо вернуться к функции $y = ux$.

Пример **29.3.** Решить уравнение $y' = \dfrac{x+y}{x-y}$.

Решение. Убедимся, что уравнение однородное: $\dfrac{\lambda x + \lambda y}{\lambda x - \lambda y} = \dfrac{x+y}{x-y}$. Делаем замену $\dfrac{y}{x} = u$. Подставив u в уравнение, получим, с учетом того, что $f(1,u) - u = \dfrac{1+u}{1-u} - u = \dfrac{1+u^2}{1-u}$,

$$\int \frac{1-u}{1+u^2}du = \int \frac{dx}{x} + C,$$

отсюда

$$\operatorname{arctg} u - \frac{1}{2}\ln(1+u^2) = \ln|x| + C.$$

Возвращаясь к функции y, получим общий интеграл:

$$\operatorname{arctg}\frac{y}{x} - \frac{1}{2}\ln\left(1 + \left(\frac{y}{x}\right)^2\right) = \ln|x| + C.$$

Линейные дифференциальные уравнения первого порядка

Линейным дифференциальным уравнением первого порядка называется уравнение вида

$$y' + p(x)y = q(x). \tag{29.9}$$

Это уравнение называется линейным потому, что неизвестная функция y и ее производная y' входят в это уравнение линейно, т.е. в первой степени, не перемножаясь между собой.

Одним из методов решения линейного уравнения (29.9) является **метод Бернулли**, который заключается в следующем. Будем искать решение уравнения (29.9) в виде $y = u(x)\,v(x) = uv$. Одну из этих функций можно взять произвольной, другая определится на основании уравнения (29.9). Сделав замену $y = uv$, получим

$$u'v + uv' + puv = q,$$

или

$$u'v + u(v' + pv) = q. \tag{29.10}$$

Выберем функцию $v = v(x)$ такой, чтобы

$$v' + pv = 0, \quad \text{т.е.} \quad \frac{dv}{dx} + pv = 0. \tag{29.11}$$

Разделяя переменные, находим

$$\frac{dv}{v} = -pdx.$$

Интегрируя, получаем

$$\ln|v| = -\int pdx + \ln|C_1|, \quad \text{или} \quad v = Ce^{-\int pdx},$$

где $C = \pm C_1$.

Так как нам достаточно какого-нибудь о д н о г о отличного от нуля решения уравнения (29.11), то за функцию $v = v(x)$ возьмем

$$v(x) = e^{-\int pdx},$$

где $\int pdx$ – какая-нибудь первообразная.

Подставим найденное значение $v(x)$ в уравнение (29.10):

$$u'v(x) = q(x), \quad \text{или} \quad v(x)\frac{du}{dx} = q(x),$$

отсюда

$$\frac{du}{dx} = \frac{q(x)}{v(x)}.$$

Находим общее решение для $u = u(x)$:

$$u = \int \frac{q(x)}{v(x)}dx + C.$$

Подставляя u и v в формулу $y = uv$, окончательно находим

$$y = v(x)\left[\int \frac{q(x)}{v(x)}dx + C\right].$$

З а м е ч а н и е. Можно доказать, что общее решение уравнения (29.9) есть сумма любого его ч а с т н о г о решения и о б щ е г о решения так называемого сопровождающего его однородного уравнения $y' + p(x)y = 0$.

П р и м е р 29.4. Решить уравнение $y' - \frac{2}{x}y = 2x^3$.

Р е ш е н и е. Сделав замену $y = uv$, получим

$$u'v + uv' - \frac{2uv}{x} = 2x^3,$$

$$u'v + u\left(v' - \frac{2v}{x}\right) = 2x^3. \qquad (*)$$

Приравниваем к нулю выражение в скобках:

$$v' - \frac{2v}{x} = 0, \quad \text{откуда} \quad \frac{dv}{v} = \frac{2dx}{x}.$$

Находим функцию v:

$$\ln v = \ln x^2, \quad v = x^2.$$

Подставив $v = x^2$ в (*), находим u:

$$u'x^2 = 2x^3, \quad \frac{du}{dx} = 2x, \quad u = x^2 + C.$$

Отсюда

$$y = (x^2 + C)x^2.$$

Уравнение Бернулли

Уравнением Бернулли называется уравнение вида

$$y' + p(x)y = q(x)y^n,$$

где $n \neq 0$, $n \neq 1$.

Уравнение Бернулли также можно решать методом Бернулли.

***Пример* 29.5.** Решить уравнение $y' + 2y = y^2 e^x$.

Решение. Сделаем замену $y = uv$:

$$u'v + uv' + 2uv = (uv)^2 e^x,$$

$$u'v + u(v' + 2v) = (uv)^2 e^x. \qquad (**)$$

Приравниваем к нулю выражение в скобках:

$$v' + 2v = 0.$$

Разделяем переменные:

$$\frac{dv}{v} = -2dx.$$

Находим v:

$$\ln v = -2x,$$

$$v = e^{-2x}.$$

Подставляем v в уравнение (**):

$$u'e^{-2x} = u^2 e^{-4x} e^x,$$

отсюда

$$u' = u^2 e^{-x}, \quad \text{т.е.} \quad \frac{du}{dx} = u^2 e^{-x}.$$

Разделяем переменные:

$$\frac{du}{u^2} = \mathrm{e}^{-x}dx.$$

Интегрируем:

$$-\frac{1}{u} = -\mathrm{e}^{-x} + C.$$

Находим u (заменяя C на $-C$):

$$u = \frac{1}{\mathrm{e}^{-x} + C}.$$

Следовательно, общим решением будет

$$y = \frac{\mathrm{e}^{-2x}}{\mathrm{e}^{-x} + C}, \quad \text{или} \quad y = \frac{1}{\mathrm{e}^{x} + C\mathrm{e}^{2x}}.$$

29.3. ПРИМЕНЕНИЕ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ В НЕПРЕРЫВНЫХ МОДЕЛЯХ ЭКОНОМИКИ

Рассмотрим некоторые примеры применения дифференциальных уравнений в динамических задачах экономики. Независимой переменной здесь является время t. Время в экономической динамике может рассматриваться и как непрерывное, и как дискретное. Мы рассматриваем непрерывное время, так как в этом случае можно использовать аппарат дифференциального исчисления и дифференциальные уравнения.

Начнем с примеров применения самых простых дифференциальных уравнений первого порядка – *уравнений с разделяющимися переменными*.

Рассмотрим уравнение вида

$$y' = g(y). \tag{29.12}$$

Очевидно, это частный случай дифференциального уравнения с разделяющимися переменными. Такие уравнения часто встречаются в задачах экономической динамики (иногда их называют *автономными уравнениями*, но в теории дифференциальных уравнений этот термин не является общеупотребительным).

Если y^* – корень уравнения $g(y) = 0$ ($y^* = \mathrm{const}$), то $y = y^*$ является решением уравнения (29.12). Такое решение называется **стационарным**.

Модель естественного роста выпуска

Обозначим через $y(t)$ интенсивность выпуска продукции. Предполагается, что продукция продается по фиксированной цене p и что рынок ненасыщаем, т.е. вся производимая продукция полностью реализуется. Назовем чистыми инвестициями разность $I = I(t)$ между общим объемом инвестиций и амортизационными затратами. Чтобы увеличить интенсивность выпуска $y(t)$, необходимо, чтобы чистые инвестиции I были больше нуля. Из предположения о ненасыщаемости рынка следует, что в результате расширения производства будет получен прирост дохода, часть которого опять будет использована для расширения выпуска продукции. Это приведет к росту интенсивности выпуска.

Предполагается, что между I и ростом интенсивности выпуска y' прямая пропорциональная зависимость, т.е. имеет место так называемый **принцип акселерации**:

$$I = my', \tag{29.13}$$

где m – норма акселерации (m = const). Пусть a – норма чистых инвестиций, т.е. часть дохода py, полученного от реализации продукции, которая тратится на чистые инвестиции, $0 < a < 1$. Тогда

$$I = apy. \tag{29.14}$$

Подставляя выражение I из (29.14) в (29.13), получаем $y' = \dfrac{ap}{m}y$. Обозначим $\dfrac{ap}{m} = k$, тогда

$$y' = ky. \tag{29.15}$$

Уравнение (29.15) – это уравнение с разделяющимися переменными. Разделяем переменные:

$$\frac{dy}{y} = kdt.$$

Интегрируя, находим общее решение:

$$\ln|y| = kt + \ln|C|,$$

$$y = Ce^{kt}. \tag{29.16}$$

Пусть в начальный момент времени $t = t_0$ зафиксирован объем выпуска y_0: $y(t_0) = y_0$,

$$y_0 = Ce^{kt_0}.$$

Тогда можно найти константу C:

$$C = y_0 e^{-kt_0},$$

следовательно,

$$y = y_0 e^{k(t-t_0)}. \tag{29.17}$$

Уравнение (29.17) называется **уравнением естественного роста**. Этим уравнением описываются также демографические процессы, процессы радиоактивного распада, размножения бактерий.

Теперь рассмотрим примеры применения линейных дифференциальных уравнений первого порядка.

Динамическая модель Кейнса

Рассмотрим простейшую балансовую модель. Пусть $Y(t)$ – национальный доход, $E(t)$ – государственные расходы, $S(t)$ – потребление, $I(t)$ – инвестиции. Все эти величины являются функциями времени t.

Составим **уравнения баланса**. Прежде всего, сумма всех *расходов* должна быть равна *национальному доходу*:

$$Y(t) = S(t) + I(t) + E(t).$$

Общее потребление $S(t)$ состоит из *внутреннего потребления* некоторой части национального дохода плюс *конечное потребление*. Первое из этих слагаемых имеет вид $a(t)Y(t)$, где $a(t)$ – коэффициент склонности к потреблению ($0 < a(t) < 1$); второе слагаемое обозначим через $b(t)$:

$$S(t) = a(t)Y(t) + b(t).$$

Наконец, размер инвестиций характеризуется произведением *нормы акселерации* $m = m(t)$ на *предельный национальный доход*:

$$I(t) = m(t)Y'(t).$$

Получаем систему

$$\begin{cases} Y(t) = S(t) + I(t) + E(t), \\ S(t) = a(t)Y(t) + b(t), \\ I(t) = m(t)Y'(t). \end{cases} \tag{29.18}$$

Все функции, входящие в уравнения (29.18), положительны.

Предполагается, что функции $a(t)$, $b(t)$, $m(t)$ и $E(t)$ заданы – они являются характеристиками функционирования данного государства.

Требуется найти динамику национального дохода, т.е. найти Y как функцию времени t.

Подставим выражения для $S(t)$ из второго уравнения и $I(t)$ из третьего уравнения системы (29.18) в первое уравнение:

$$Y(t) = a(t)Y(t) + b(t) + m(t)Y'(t) + E(t).$$

Выражаем $Y'(t)$:

$$Y'(t) = \frac{1-a(t)}{m(t)}Y(t) - \frac{b(t)+E(t)}{m(t)},$$

или

$$Y'(t) - \frac{1-a(t)}{m(t)}Y(t) = -\frac{b(t)+E(t)}{m(t)}. \qquad (29.19)$$

Это линейное дифференциальное уравнение:

$$Y' + p(t)Y = q(t),$$

где $p(t) = -\dfrac{1-a(t)}{m(t)}$, $q(t) = -\dfrac{b(t)+E(t)}{m(t)}$.

Нам уже известен метод нахождения общего решения линейного уравнения первого порядка (см. § 29.2). Однако его реализация в применении к уравнению (29.19) была бы весьма громоздкой. Рассмотрим частный случай, когда основные параметры a, b и m постоянны. Тогда уравнение (29.19) упрощается:

$$Y' - \frac{1-a}{m}Y = -\frac{b+E}{m}. \qquad (29.20)$$

Это линейное дифференциальное уравнение с постоянными коэффициентами.

Как уже отмечалось (см. с. 347), *общее решение* неоднородного уравнения есть сумма какого-нибудь его *частного решения* и *общего решения* сопровождающего однородного уравнения. В качестве частного решения $\widetilde{Y}$ уравнения (29.20) возьмем решение, получаемое при $Y' = 0$, т.е.

$$\widetilde{Y} = \frac{b+E}{1-a}.$$

Это решение называется **равновесным**.

Так как $E > 0$, $0 < a < 1$, то $\widetilde{Y} > 0$.

Общее решение однородного уравнения $Y' - \dfrac{1-a}{m}Y = 0$ имеет вид $Y_0 = Ce^{\alpha t}$, где $\alpha = \dfrac{1-a}{m}$. (Очевидно, что $\alpha > 0$.) Поэтому общее решение уравнения (29.20) имеет вид

$$Y(t)=\frac{b+E}{1-a}+C\exp\left(\frac{1-a}{m}t\right),$$

или

$$Y(t)=\frac{b+E}{1-a}+Ce^{\alpha t},\quad \text{где}\quad \alpha=\frac{1-a}{m}.$$

Если в начальный момент $Y_0<\tilde{Y}$, то $C=Y_0-\tilde{Y}<0$ и национальный доход со временем падает при фиксированных параметрах a, b, m и E. Если же $Y_0>\tilde{Y}$, то $C>0$ и национальный доход растет (рис. 29.2).

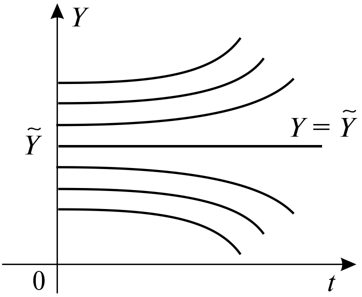


Рис. 29.2. Интегральные кривые уравнения (29.20)

Уравнение Самуэльсона

Уравнением Самуэльсона называется уравнение

$$p'=k[D(p)-S(p)], \tag{29.21}$$

где $D(p)$ и $S(p)$ – соответственно величины спроса и предложения при цене p, $k>0$. Уравнение (29.21) моделирует связь между изменением цены p и неудовлетворенным спросом $D(p)-S(p)$.

Рассмотрим простой случай, когда спрос и предложение задаются линейными функциями:

$$D(p)=a-bp,\qquad S(p)=m+np,$$

здесь a, b, m, n – некоторые положительные числа. При этом, очевидно, $a>m$, так как при нулевой цене спрос превышает предложение. В этом случае уравнение (29.21) имеет вид

$$p'=k(a-m)-k(n+b)p. \tag{29.22}$$

Уравнение (29.21), а следовательно, и (29.22), является линейным дифференциальным уравнением. Найдем *решение* однородного уравнения, соответствующего уравнению (29.22):

$$p'=-k(n+b)p,$$

или

$$\frac{dp}{dt} = -k(n+b)\,p.$$

Разделяем переменные и интегрируем:

$$\frac{dp}{p} = -k(n+b)dt,$$

$$\ln|p| = -k(n+b)t + \ln|C|,$$

$$p(t) = Ce^{-k(n+b)t}. \tag{29.23}$$

Как и в предыдущем случае, в качестве *частного решения* уравнения (29.22) можно использовать равновесное решение $p(t) = \tilde{p} = \text{const}$, где $\tilde{p}$ – корень уравнения $D(p) - S(p) = 0$ (в этом случае $p' = 0$). Из (29.22) находим

$$\tilde{p} = \frac{a-m}{n+b}.$$

Получаем общее решение уравнения (29.22):

$$p(t) = \frac{a-m}{n+b} + Ce^{-k(n+b)t}. \tag{29.24}$$

Вопросы

1. Что называется дифференциальным уравнением?
2. Что такое порядок дифференциального уравнения?
3. Что называется решением дифференциального уравнения?
4. Как формулируется задача Коши для дифференциального уравнения первого порядка?
5. Что такое общее решение дифференциального уравнения первого порядка?
6. Что значит решить дифференциальное уравнение?
7. Какое уравнение называется дифференциальным уравнением с разделяющимися переменными?
8. Какая функция называется однородной функцией n-го измерения?
9. Какое дифференциальное уравнение первого порядка называется однородным?
10. Какое дифференциальное уравнение первого порядка называется линейным?

11. В чем заключается метод Бернулли решения дифференциального уравнения? Для каких дифференциальных уравнений первого порядка он обычно применяется?
12. Что называется уравнением естественного роста? Какие процессы описывает это уравнение?
13. Что такое динамическая модель Кейнса?
14. Как выглядит уравнение Самуэльсона? Каков смысл входящих в него величин?

Глава 30
ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ВТОРОГО И ВЫСШИХ ПОРЯДКОВ

30.1. ОСНОВНЫЕ ПОНЯТИЯ

Дифференциальное уравнение второго порядка имеет вид

$$F(x, y, y', y'') = 0 \tag{30.1}$$

или

$$y'' = f(x, y, y'). \tag{30.2}$$

Условия

$$y(x_0) = y_0, \quad y'(x_0) = y'_0 \tag{30.3}$$

называются **начальными условиями**.

Определение. Функция $y = \varphi(x, C_1, C_2)$ называется **общим решением** уравнения (30.1), если она является решением уравнения (30.1) при любых значениях C_1 и C_2 и если при любых начальных условиях (30.3) существуют единственные значения постоянных $C_1 = C_1^0$, $C_2 = C_2^0$, такие, что функция $y = \varphi(x, C_1^0, C_2^0)$ удовлетворяет данным начальным условиям.

Определение. Любая функция $y = \varphi(x, C_1^0, C_2^0)$, получающаяся из общего решения $y = \varphi(x, C_1, C_2)$ уравнения (30.1) при определенных значениях постоянных $C_1 = C_1^0$, $C_2 = C_2^0$, называется **частным решением**.

В некоторых случаях решение дифференциального уравнения второго порядка может быть сведено к последовательному решению двух дифференциальных уравнений первого порядка.

30.2. ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ, ДОПУСКАЮЩИЕ ПОНИЖЕНИЕ ПОРЯДКА

1. *Уравнение не содержит в явном виде искомой функции* y, т.е. имеет вид $\boldsymbol{F(x, y', y'') = 0}$. В этом случае достаточно сделать замену $y' = z$. Тогда $y'' = z'$ и уравнение принимает вид

$$F(x, z, z') = 0,$$

т.е. является уравнением первого порядка относительно z.

Найдем его общее решение

$$z = \varphi(x, C_1).$$

Сделаем обратную замену:

$$y' = \varphi(x, C_1).$$

Отсюда

$$y = \int \varphi(x, C_1)\,dx + C_2.$$

Пример **30.1.** Решить уравнение $xy'' + y' = 0$.

Решение. Положим $z = y'$. Тогда $y'' = z'$ и исходное уравнение принимает вид

$$xz' + z = 0, \quad \text{или} \quad x\frac{dz}{dx} + z = 0,$$

откуда

$$\frac{dz}{z} = -\frac{dx}{x}.$$

Интегрируя, получаем

$$z = \frac{C_1}{x}, \quad \text{или} \quad y' = \frac{C_1}{x}.$$

Решая последнее уравнение, получаем

$$y = C_1 \ln|x| + C_2.$$

2. *Уравнение не содержит в явном виде аргумента* x, т.е. имеет вид $\boldsymbol{F(y, y', y'') = 0}$. В этом случае порядок уравнения можно понизить, если положить $y' = z = z(y)$. Тогда

$$y'' = z'_x = z'_y\, y'_x = z z'_y = z\frac{dz}{dy}.$$

Пример **30.2.** Решить уравнение $y'' - (y')^2 = 0$.

Решение. Заменой $y' = z = z(y)$ сводим данное уравнение к уравнению первого порядка:

$$z\frac{dz}{dy} - z^2 = 0, \quad \text{или} \quad z\left(\frac{dz}{dy} - z\right) = 0.$$

Первое решение этого уравнения $z = 0$, или $y = C$, где $C = \text{const}$. Далее получаем

$$\frac{dz}{dy} - z = 0.$$

Разделяя переменные и интегрируя, получаем

$$z = C_1 e^y.$$

Сделаем обратную замену:

$$\frac{dy}{dx} = C_1 e^{y}.$$

Здесь также переменные разделяются:

$$e^{-y} dy = C_1 dx,$$

$$-e^{-y} = C_1 x + C_2.$$

Так как C_1 и C_2 – произвольные постоянные, можно написать (взяв $-C_1$ вместо C_1 и $-C_2$ вместо C_2):

$$e^{-y} = C_1 x + C_2,$$

$$y = -\ln(C_1 x + C_2).$$

Очевидно, это решение включает в себя и решение $y = C$, полученное выше.

3. Уравнение имеет вид $\boldsymbol{y'' = f(y)}$. Это частный случай уравнения, рассмотренного в п. 2. Поэтому оно решается заменой $y' = z = z(y)$, $y'' = z\dfrac{dz}{dy}$. В результате такой замены данное уравнение преобразуется в уравнение первого порядка:

$$z\frac{dz}{dy} = f(y).$$

Отсюда

$$zdz = f(y)dy.$$

Интегрируя, получаем

$$\frac{z^2}{2} = \int f(y)dy + C_1.$$

Отсюда

$$z = \pm\sqrt{2\left(C_1 + \int f(y)dy\right)},$$

т.е.

$$\frac{dy}{dx} = \pm\sqrt{2\left(C_1 + \int f(y)dy\right)}, \quad \text{или} \quad \frac{dy}{\sqrt{2\left(C_1 + \int f(y)dy\right)}} = \pm\, dx.$$

Взяв интегралы от левой и правой частей последнего равенства, получим общий интеграл.

Пример **30.3.** Найти частное решение $2y^3y''=1,\ \ y\left(\frac{1}{2}\right)=1,$ $y'\left(\frac{1}{2}\right)=1.$

Решение. Делая замену $y''=z\frac{dz}{dy},$ находим

$$2y^3z\frac{dz}{dy}=1.$$

Отсюда

$$2z\frac{dz}{dy}=\frac{1}{y^3},\qquad 2zdz=\frac{dy}{y^3}.$$

Интегрируя, получаем

$$z^2=-\frac{1}{2y^2}+C_1,\qquad z=\pm\sqrt{-\frac{1}{2y^2}+C_1}\,.$$

Полагая здесь $x=\frac{1}{2}$ и учитывая, что при этом значении x будет $y'=z=1$, видим, что перед радикалом надо взять знак «плюс», и находим $C_1=\frac{3}{2}.$ Таким образом,

$$z=\sqrt{\frac{3}{2}-\frac{1}{2y^2}},$$

или, что то же самое,

$$\frac{dy}{dx}=\sqrt{\frac{3}{2}-\frac{1}{2y^2}},\qquad \text{или}\qquad \frac{dy}{dx}=\sqrt{\frac{3y^2-1}{2y^2}}.$$

Отсюда

$$\frac{\sqrt{2}\;ydy}{\sqrt{3y^2-1}}=dx.$$

Интегрируем. Интеграл от левой части берется заменой $t=3y^2-1,\ \ dt=6ydy$:

$$\frac{\sqrt{2}}{3}\sqrt{3y^2-1}=x+C_2.$$

При $x=\frac{1}{2}$ находим $\frac{\sqrt{2}}{3}\cdot\sqrt{2}=\frac{1}{2}+C_2$, или $\frac{2}{3}=\frac{1}{2}+C_2$, откуда $C_2=\frac{1}{6}.$

Получаем

$$\frac{\sqrt{2}}{3}\sqrt{3y^2-1}=x+\frac{1}{6}.$$

Возводя в квадрат и проводя очевидные преобразования, получаем окончательно

$$2y^2 = 3x^2 + x + \frac{3}{4}.$$

30.3. ЛИНЕЙНЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ *n*-го ПОРЯДКА

Линейным дифференциальным уравнением *n*-го порядка называется уравнение вида

$$a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \ldots + a_{n-1}(x)y' + a_n(x)y = F(x). \tag{30.4}$$

Это уравнение называется линейным, потому что неизвестная функция y и ее производные $y', y'', \ldots, y^{(n)}$ входят в него линейно, т.е. в первой степени, не перемножаясь между собой. Здесь $a_0(x)$, $a_1(x), \ldots, a_{n-1}(x), a_n(x), F(x)$ – заданные функции от x (в частности, они могут быть постоянными), причем $a_0(x) \neq 0$ для всех значений x из той области, в которой мы рассматриваем уравнение (30.4) (в противном случае порядок уравнения не был бы равен n). Поэтому можно разделить обе части уравнения на $a_0(x)$ и привести его к виду

$$y^{(n)} + p_1(x)y^{(n-1)} + \ldots + p_{n-1}(x)y' + p_n(x)y = f(x), \tag{30.5}$$

где $p_1(x) = \dfrac{a_1(x)}{a_0(x)}, \ \ldots, \ p_{n-1}(x) = \dfrac{a_{n-1}(x)}{a_0(x)}, \ p_n(x) = \dfrac{a_n(x)}{a_0(x)}, \ f(x) = \dfrac{F(x)}{a_0(x)}.$

В дальнейшем линейное дифференциальное уравнение будем записывать в форме (30.5). Функция $f(x)$ в уравнении (30.5) называется **свободным членом**. Если $f(x)$ тождественно равно нулю, то уравнение (30.5) называется **однородным**; в этом случае оно, очевидно, имеет вид

$$y^{(n)} + p_1(x)y^{(n-1)} + \ldots + p_{n-1}(x)y' + p_n(x)y = 0. \tag{30.6}$$

В противном случае уравнение (30.5) называется **неоднородным**. Функция $f(x)$ в (30.5) называется **правой частью** (или **свободным членом**) уравнения (30.5).

Далее будем вести изложение теории и проводить доказательства, как правило, для уравнений второго порядка, так как здесь можно изучить все основные интересующие нас закономерности. Итак, речь пойдет в основном об уравнениях

$$y'' + p(x)y' + q(x)y = f(x). \tag{30.7}$$

Прежде всего установим некоторые основные свойства линейных однородных уравнений.

Структура общего решения однородного линейного дифференциального уравнения

Рассмотрим уравнение вида

$$y'' + p(x)y' + q(x)y = 0. \tag{30.8}$$

В дальнейшем увидим, что для того чтобы уметь решать уравнение (30.7), в котором $f(x) \neq 0$, надо также уметь решать уравнение (30.8). Рассмотрим два простых **свойства решений уравнения** (30.8).

1. Если y_0 – решение уравнения (30.8), а C – постоянная, то произведение Cy_0 также является решением этого уравнения.

Доказательство. Подставим $y = Cy_0$ в уравнение (30.8). Так как

$$y' = Cy_0', \qquad y'' = Cy_0'',$$

то левая часть уравнения в результате подстановки выглядит так:

$$Cy_0'' + p(x)Cy_0' + q(x)Cy_0,$$

или, что то же самое,

$$C(y_0'' + p(x)y_0' + q(x)y_0).$$

Так как y_0 является решением дифференциального уравнения (30.8), то выражение, стоящее в скобках, тождественно равно нулю. Таким образом, уравнение (30.8) обратилось в тождество. Утверждение доказано.

2. Если y_1 и y_2 – решения дифференциального уравнения (30.8), то их сумма $y_1 + y_2$ также есть решение этого уравнения.

Доказательство. Так как

$$y' = y_1' + y_2', \qquad y'' = y_1'' + y_2''$$

и так как y_1 и y_2 – решения уравнения (30.8), то справедливы тождественные равенства

$$y_1'' + p(x)y_1' + q(x)y_1 = 0, \qquad y_2'' + p(x)y_2' + q(x)y_2 = 0. \tag{*}$$

Подставляя в уравнение (30.8) сумму $y_1 + y_2$ и принимая во внимание тождества (*), получаем

$$(y_1 + y_2)'' + p(x)(y_1 + y_2)' + q(x)(y_1 + y_2) =$$
$$= (y_1'' + p(x)y_1' + q(x)y_1) + (y_2'' + p(x)y_2' + q(x)y_2) = 0 + 0 = 0.$$

Итак, уравнение (30.8) обратилось в тождество. Утверждение доказано.

О п р е д е л е н и е. Две функции y_1 и y_2 называются **линейно независимыми**, если тождественное равенство

$$k_1 y_1 + k_2 y_2 = 0 \qquad (**)$$

имеет единственное возможное решение

$$k_1 = k_2 = 0.$$

Если же существует ненулевое решение (**), то функции y_1 и y_2 называются **линейно зависимыми**.

Очевидно, функции y_1 и y_2 линейно независимы тогда и только тогда, когда их отношение не является константой:

$$\frac{y_1}{y_2} \neq \text{const}.$$

О п р е д е л е н и е. Если $y_1 = y_1(x)$, $y_2 = y_2(x)$, то определитель

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1'$$

называется **определителем Вронского** данных функций.

***Л е м м а* 30.1.** Если функции $y_1 = y_1(x)$ и $y_2 = y_2(x)$ л и н е й н о з а в и с и м ы на отрезке $[a, b]$, то определитель Вронского, составленный из них, тождественно равен нулю на этом отрезке; если же функции л и н е й н о н е з а в и с и м ы на $[a, b]$, то определитель Вронского отличен от нуля на $[a, b]$.

Д о к а з а т е л ь с т в о. Пусть функции y_1 и y_2 линейно зависимы на отрезке y_1 и y_2. Тогда эти функции пропорциональны на $[a, b]$, т.е. $y_2 = ky_1$. Следовательно, определитель $W(x)$ содержит пропорциональные столбцы, поэтому он равен нулю на отрезке $[a, b]$:

$$W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = \begin{vmatrix} y_1 & ky_1 \\ y_1' & ky_1' \end{vmatrix} = 0.$$

Первая часть леммы доказана.

Доказательство второй части леммы проведем от противного. Пусть функции $y_1 = y_1(x)$ и $y_2 = y_2(x)$ линейно независимы на отрезке $[a, b]$; предположим, что определитель $W(x)$ тождественно равен нулю на этом отрезке. Тогда его столбцы необходимо пропорциональны: $y_2 = ky_1$, $y_2' = ky_1'$. А это значит, что функции y_1 и y_2 пропорциональны, а потому линейно зависимы, что противоречит условию леммы. Доказательство закончено.

Лемма **30.2.** Если определитель Вронского $W(x)$, составленный для решений y_1 и y_2 однородного линейного уравнения (30.8), не равен нулю при каком-нибудь значении $x = x_0$ на отрезке $[a, b]$ и коэффициенты уравнения непрерывны на этом отрезке, то $W(x)$ не обращается в нуль ни при каком значении x на этом отрезке.

Доказательство. Так как $y_1 = y_1(x)$ и $y_2 = y_2(x)$ – решения уравнения (30.8), то справедливы тождества

$$y_1'' + p(x)y_1' + q(x)y_1 = 0, \qquad y_2'' + p(x)y_2' + q(x)y_2 = 0.$$

Умножим второе равенство на y_1, а первое – на y_2 и вычтем из второго равенства первое:

$$(y_1 y_2'' - y_1'' y_2) + p(x)(y_1 y_2' - y_1' y_2) = 0. \tag{30.9}$$

Разность, стоящая во второй скобке, есть определитель Вронского $W(x)$. Действительно, $W(x) = y_1 y_2' - y_1' y_2$. Продифференцируем $W(x)$:

$$W'(x) = (y_1 y_2' - y_1' y_2)' =$$
$$= y_1' y_2' + y_1 y_2'' - y_1'' y_2 - y_1' y_2' = y_1 y_2'' - y_1'' y_2.$$

Как видим, разность, стоящая в первой скобке уравнения (30.9), есть производная от определителя Вронского, поэтому уравнение (30.9) может быть представлено в виде

$$W' + p(x)W = 0, \tag{30.10}$$

т.е. является дифференциальным уравнением с разделяющимися переменными. Найдем решение этого уравнения, удовлетворяющее условию $W(x_0) = W_0$ в предположении, что $W_0 \neq 0$. Разделяя переменные, получаем

$$\frac{dW}{W} = -p(x)dx.$$

Интегрируя, находим

$$\ln|W| = -\int_{x_0}^{x} p(x)dx + \ln|C|.$$

Отсюда

$$\ln\left|\frac{W}{C}\right| = -\int_{x_0}^{x} p(x)dx,$$

и мы получаем общее решение уравнения (30.10):

$$W = C\exp\left(-\int_{x_0}^{x} p(x)dx\right). \qquad (30.11)$$

(Напомним, что $\exp a$ означает e^a при любом a.)

Формула (30.11) называется **формулой Лиувилля**.

Определим теперь константу C так, чтобы выполнялось начальное условие $W(x_0) = W_0$. Подставим $x = x_0$ в левую и правую части равенства (30.11). Получаем (с учетом того, что $\int_{x_0}^{x_0} p(x)dx = 0$):

$$W_0 = C.$$

Подставляем найденное значение $C = W_0$ в равенство (30.11). Итак, решение уравнения (30.10), удовлетворяющее начальным условиям $W(x_0) = W_0$, имеет вид

$$W = W_0\exp\left(-\int_{x_0}^{x} p(x)dx\right).$$

Тогда (так как показательная функция не обращается в нуль ни при каком значении аргумента и так как по условию $W_0 \neq 0$) из последнего равенства следует, что $W \neq 0$ ни при каком значении x. Доказательство закончено.

Из доказанных ранее свойств 1 и 2 решений однородного линейного дифференциального уравнения (30.8) следует, что *линейная комбинация решений* $y_1(x)$ и $y_2(x)$ уравнения (30.8), т.е.

$$y = C_1 y_1(x) + C_2 y_2(x),$$

где C_1 и C_2 – константы, также есть решение уравнения (30.8).

Сформулируем и докажем теорему, описывающую **структуру общего решения однородного линейного дифференциального уравнения**.

***Теорема* 30.1.** Если $y_1(x)$ и $y_2(x)$ – линейно независимые на $[a, b]$ частные решения однородного линейного дифференциального уравнения (30.8), то общее решение этого уравнения имеет вид

$$y = C_1 y_1(x) + C_2 y_2(x), \qquad (30.12)$$

где C_1 и C_2 – произвольные константы.

Доказательство. Ранее было установлено, что линейная комбинация (30.12) есть *решение* уравнения (30.8); нужно доказать, что она представляет собой *общее решение*, т.е. надо показать, что для любых начальных условий существуют такие значения констант C_1 и C_2, при которых эта линейная комбинация есть решение, удовлетворяющее этим начальным условиям.

Возьмем любое число $x_0 \in [a, b]$ и любые числа y_0, y_0' и составим начальные условия:

$$y(x_0) = y_0, \qquad y'(x_0) = y_0'.$$

Выполнение этих условий для функции (30.12) означает, что

$$\begin{cases} C_1 y_1(x_0) + C_2 y_2(x_0) = y_0, \\ C_1 y_1'(x_0) + C_2 y_2'(x_0) = y_0'. \end{cases}$$

Мы получили линейную относительно неизвестных констант C_1 и C_2 систему двух уравнений. Определителем этой системы является определитель Вронского $W(x_0)$, и так как y_1 и y_2 линейно независимы, то этот определитель не равен нулю. Следовательно, система имеет единственное решение при любых значениях правых частей y_0 и y_0':

$$C_1 = C_1^0, \qquad C_2 = C_2^0.$$

Подставляя эти значения в решение (30.12), получим частное решение, удовлетворяющее заданным начальным условиям. Так как начальные условия заданы произвольно, можно утверждать, что решение (30.12) является общим решением уравнения (30.8). Теорема доказана.

Замечание. Если отбросить условие линейной зависимости функций y_1 и y_2, то функция (30.12), хотя и останется *решением* дифференциального уравнения (30.8), уже не будет его *общим решением*.

Действительно, пусть, например,

$$\frac{y_1}{y_2} = 5.$$

Тогда $y_1 = 5y_2$ и (30.12) имеет вид

$$y = C_1 y_1 + C_2 y_2 = 5C_1 y_2 + C_2 y_2 = (5C_1 + C_2) y_2 = C y_2,$$

где положено $5C_1 + C_2 = C$. Иначе говоря, в случае когда функции y_1 и y_2 линейно зависимы, число произвольных постоянных, входящих в (30.12), за счет введения новых обозначений может быть снижено до одной, а функция, содержащая одну произвольную константу, не может быть общим решением дифференциального уравнения (30.8).

Смысл теоремы 30.1 заключается в том, что она сводит задачу нахождения общего решения дифференциального уравнения (30.8) к более простой задаче нахождения двух линейно независимых частных решений этого уравнения. Этой последней задачей мы и займемся теперь, однако ограничимся самым простым случаем, когда коэффициенты уравнения постоянны: $p(x) = p = \mathrm{const}$, $q(x) = q = \mathrm{const}$.

Однородные линейные дифференциальные уравнения с постоянными коэффициентами

Рассмотрим уравнение вида

$$y'' + py' + qy = 0, \tag{30.13}$$

где p и q – некоторые постоянные.

Займемся сначала нахождением общего решения уравнения (30.13). В соответствии с теоремой 30.1 для этого надо найти два линейно независимых частных решения этого уравнения.

Решение уравнения (30.13) будем искать в виде

$$y = \mathrm{e}^{kx}. \tag{30.14}$$

Так как

$$y' = k\mathrm{e}^{kx}, \qquad y'' = k^2\mathrm{e}^{kx},$$

то, подставляя (30.14) в левую часть (30.13), получим

$$k^2\mathrm{e}^{kx} + pk\mathrm{e}^{kx} + q\mathrm{e}^{kx} = 0, \qquad \text{или} \qquad \mathrm{e}^{kx}(k^2 + pk + q) = 0.$$

Итак, для того чтобы (30.14) было решением уравнения (30.13), надо, чтобы k было корнем квадратного уравнения

$$k^2 + pk + q = 0. \tag{30.15}$$

Это уравнение называется *характеристическим* для дифференциального уравнения (30.13).

Возможны три случая.

1. Корни уравнения (30.15) действительные и различные: $k_1 = a$, $k_2 = b$, $a \neq b$. Тогда (30.13) имеет два решения:

$$y_1 = \mathrm{e}^{ax}, \quad y_2 = \mathrm{e}^{bx}.$$

Эти решения линейно независимы, так как

$$\frac{\mathrm{e}^{ax}}{\mathrm{e}^{bx}} \neq \mathrm{const}.$$

Общее решение уравнения (30.13) таково:

$$y = C_1 e^{ax} + C_2 e^{bx}.$$

***Пример* 30.4.** Решить уравнение $y'' - 5y' + 4y = 0$.

Решение. Составляем характеристическое уравнение:

$$k^2 - 5k + 4 = 0.$$

Находим его корни: $k_1 = 1$, $k_2 = 4$. Получаем общее решение данного уравнения:

$$y = C_1 e^{x} + C_2 e^{4x}.$$

2. Корни уравнения (30.15) действительные и совпадающие: $k_1 = k_2 = a$ $\left(a = -\dfrac{p}{2}\right)$. Одним из частных решений уравнения (30.13) будет

$$y_1 = e^{ax}.$$

Однако второе решение y_2, такое, чтобы y_1 и y_2 были линейно независимы, мы найти пока не можем.

В этом случае оказывается, что, наряду с решением

$$y_1 = e^{ax},$$

у уравнения (30.13) имеется решение

$$y_2 = x e^{ax}.$$

Убедимся в этом. Для того чтобы характеристическое уравнение $k^2 + pk + q = 0$ имело равные корни, а они выражаются в виде $k_{1,2} = -\dfrac{p}{2} \pm \sqrt{\dfrac{p^2}{4} - q}$, надо, чтобы дискриминант был равен нулю: $\dfrac{p^2}{4} - q = 0$. Иначе говоря, условием совпадения корней является равенство

$$\frac{p^2}{4} = q.$$

Тогда $k_1 = k_2 = -\dfrac{p}{2}$, и одним из решений уравнения (30.13) будет функция

$$y_1 = e^{-\frac{p}{2}x}.$$

Убедимся в том, что решением уравнения (30.13) будет также

$$y_2 = x\mathrm{e}^{-\frac{p}{2}x}.$$

Находим производные y_2' и y_2'':

$$y_2' = \mathrm{e}^{-\frac{p}{2}x} - \frac{p}{2}x\mathrm{e}^{-\frac{p}{2}x}, \qquad y_2'' = -p\mathrm{e}^{-\frac{p}{2}x} + \frac{p^2}{4}x\mathrm{e}^{-\frac{p}{2}x}.$$

Подставим y_2 и ее производные в левую часть уравнения (30.13):

$$-p\mathrm{e}^{-\frac{p}{2}x} + \frac{p^2}{4}x\mathrm{e}^{-\frac{p}{2}x} + p\left(\mathrm{e}^{-\frac{p}{2}x} - \frac{p}{2}x\mathrm{e}^{-\frac{p}{2}x}\right) + qx\mathrm{e}^{-\frac{p}{2}x}.$$

Сделав очевидные преобразования, получаем

$$\left(q - \frac{p^2}{4}\right)x\mathrm{e}^{-\frac{p}{2}x}.$$

Это выражение равно нулю (так как $q - \frac{p^2}{4} = 0$), и утверждение доказано.

Итак, $y_1 = \mathrm{e}^{-\frac{p}{2}x}$ и $y_2 = x\mathrm{e}^{-\frac{p}{2}x}$ – два решения уравнения (30.13) в случае, когда корни характеристического уравнения совпадают. Линейная независимость y_1 и y_2 очевидна. Поэтому общее решение дифференциального уравнения (30.13) имеет вид

$$y = C_1\mathrm{e}^{-\frac{p}{2}x} + C_2x\mathrm{e}^{-\frac{p}{2}x}.$$

Итак, отметим еще раз, что справедливо ***утверждение***: если характеристическое уравнение имеет совпадающие корни $k_1 = k_2 = a$, то наряду с функцией

$$y_1 = \mathrm{e}^{ax}$$

решением дифференциального уравнения (30.13) будет также функция

$$y_2 = x\mathrm{e}^{ax}.$$

Тогда *общим решением* уравнения (30.13) будет функция

$$y = C_1\mathrm{e}^{ax} + C_2x\mathrm{e}^{ax}. \tag{30.16}$$

Следует заметить, что в случае, когда характеристическое уравнение имеет два различных корня $k_1 = a$, $k_2 = b \neq a$, функция $y = x\mathrm{e}^{ax}$ не будет решением дифференциального уравнения (30.13).

Пример **30.5.** Решить уравнение $y'' - 4y' + 4y = 0$.

Решение. Характеристическое уравнение $k^2 - 4k + 4 = 0$ имеет два совпадающих корня $k_1 = k_2 = 2$. Поэтому общее решение данного уравнения таково:

$$y = C_1 e^{2x} + C_2 x e^{2x}.$$

3. Корни характеристического уравнения комплексные сопряженные: $k_{1,2} = \alpha \pm i\beta$, где i – мнимая единица, $i^2 = -1$. В этом случае можно доказать, что *общим решением* уравнения (30.13) будет

$$y = e^{\alpha x}(C_1 \cos\beta x + C_2 \sin\beta x).$$

(Примем это утверждение без доказательства.)

Пример **30.6.** Решить уравнение $y'' - 4y' + 13y = 0$.

Решение. Характеристическое уравнение $k^2 - 4k + 13 = 0$ имеет корни $k_{1,2} = 2 \pm 3i$. Общее решение данного уравнения

$$y = e^{2x}(C_1 \cos 3x + C_2 \sin 3x).$$

30.4. СТРУКТУРА ОБЩЕГО РЕШЕНИЯ НЕОДНОРОДНОГО ЛИНЕЙНОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ

Перейдем теперь к неоднородному линейному дифференциальному уравнению

$$y'' + p(x)y' + q(x)y = f(x). \qquad (30.17)$$

Вместе с ним рассмотрим однородное линейное уравнение с той же левой частью, т.е. уравнение

$$y'' + p(x)y' + q(x)y = 0. \qquad (30.18)$$

Уравнение (30.18) называют **соответствующим**, или **сопровождающим**, уравнением для дифференциального уравнения (30.17).

Теорема **30.2.** Если $\tilde{y}$ – какое-нибудь частное решение дифференциального уравнения (30.17), а y_0 – общее решение его сопровождающего уравнения (30.18), то их сумма

$$y = y_0 + \tilde{y} \qquad (30.19)$$

является общим решением дифференциального уравнения (30.17).

(Иначе говоря, *общее решение* неоднородного линейного дифференциального уравнения есть сумма его *частного решения* и *общего решения* соответствующего однородного уравнения.)

Доказательство. 1. Пусть $\tilde{y}$ – частное решение неоднородного уравнения (30.17), а $y_0 = C_1 y_1 + C_2 y_2$ – общее решение сопровождающего однородного уравнения. Убедимся сначала, что функция

$$y = y_0 + \tilde{y}$$

является *решением* уравнения (30.17). Подставим функцию (30.19) в левую часть уравнения (30.17):

$$y_0'' + \tilde{y}'' + p(x)(y_0' + \tilde{y}') + q(x)(y_0 + \tilde{y}).$$

Перегруппировав члены, получаем

$$[y_0'' + p(x)y_0' + q(x)y_0] + [\tilde{y}'' + p(x)\tilde{y}' + q(x)\tilde{y}].$$

Так как y_0 – решение уравнения (30.18), то выражение в первых квадратных скобках равно нулю. Так как $\tilde{y}$ есть решение уравнения (30.17), то выражение, стоящее во вторых квадратных скобках, равно $f(x)$. Итак, подставив (30.19) в уравнение (30.17), получаем тождество $f(x) = f(x)$.

Следовательно, функция (30.19) действительно является *решением* дифференциального уравнения (30.17).

2. Теперь надо убедиться в том, что функция (30.19) является *общим решением* неоднородного уравнения (30.17). Пусть y – любое решение неоднородного уравнения (30.17), а $\tilde{y}$ – решение этого же уравнения (30.17).

Рассмотрим разность $y - \tilde{y}$. Покажем, что эта разность является решением однородного уравнения (30.18), для чего подставим ее в левую часть уравнения (30.18) и сгруппируем соответствующие слагаемые:

$$(y - \tilde{y})'' + p(x)(y - \tilde{y})' + q(x)(y - \tilde{y}) =$$
$$= [y'' + p(x)y' + q(x)y] - [\tilde{y}'' + p(x)\tilde{y}' + q(x)\tilde{y}] = f(x) - f(x) = 0.$$

Следовательно, эта разность есть частное решение однородного уравнения (30.18), и это решение можно записать в виде

$$y - \tilde{y} = C_1^0 y_1 + C_2^0 y_2,$$

где C_1^0 и C_2^0 – соответствующие значения констант C_1 и C_2 в формуле общего решения однородного уравнения.

Мы доказали, что любое решение уравнения (30.17) можно получить по формуле (30.19) путем соответствующего подбора констант C_1 и C_2. Следовательно, функция (30.19) является общим решением неоднородного линейного дифференциального уравнения (30.17). Доказательство закончено.

В процессе доказательства теоремы 30.2 были доказаны следующие **свойства решений линейных дифференциальных уравнений**:

1. Если $\tilde{y}$ – решение неоднородного дифференциального уравнения (30.17), а y_0 – решение сопровождающего однородного уравнения (30.18), то их с у м м а $y = \tilde{y} + y_0$ есть решение неоднородного уравнения (30.17).

2. Если y_1 и y_2 – два решения неоднородного дифференциального уравнения (30.17), то их р а з н о с т ь $y = y_1 - y_2$ есть решение сопровождающего однородного уравнения (30.18).

Доказанная теорема 30.2 указывает **способ нахождения общего решения неоднородного уравнения** (30.17): надо найти общее решение сопровождающего однородного уравнения (30.18) и какое-нибудь частное решение уравнения (30.17).

Находить общее решение однородного уравнения (30.18) мы умеем, правда, только для случая, когда коэффициенты постоянны: $p(x) = p = \text{const}$, $q(x) = q = \text{const}$. **Задача нахождения частного решения неоднородного уравнения** (30.17) в общем случае весьма сложна. Рассмотрим ее лишь для простых случаев, когда коэффициенты уравнения (30.17) постоянны, а правая часть $f(x)$ имеет специальный вид. Итак, перейдем теперь к линейному неоднородному уравнению

$$y'' + py' + qy = f(x), \tag{30.20}$$

где $p, q = \text{const}$.

В дальнейшем будем применять символы $P_n(x)$ и $Q_n(x)$ для обозначения многочленов степени n:

$$P_n(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n,$$

$$Q_n(x) = b_0x^n + b_1x^{n-1} + \ldots + b_{n-1}x + b_n.$$

Рассмотрим три различных частных вида функции $f(x)$.

А. Пусть $f(x) = P_n(x)$. Правая часть уравнения (30.20) есть многочлен степени n. Так как производная от многочлена есть многочлен, можно попытаться найти частное решение $\tilde{y}$ уравнения (30.20) также в виде многочлена, коэффициенты которого пока неизвестны. Тем не менее для их нахождения существует уже знакомый нам *метод неопределенных коэффициентов*.

***Пример* 30.7.** Найти общее решение уравнения

$$y'' - 3y' + 2y = 2x^3 - 7x^2 - 4x + 10.$$

Р е ш е н и е. Мы видим, что если подставить многочлен третьей степени

$$y = ax^3 + bx^2 + cx + d$$

в левую часть данного уравнения, то в этой левой части также возникнет многочлен третьей степени. Постараемся подобрать коэффициенты a, b, c, d таким образом, чтобы уравнение превратилось в тождество. Дифференцируем многочлен и подставляем в левую часть:

$$y' = 3ax^2 + 2bx + c, \qquad y'' = 6ax + 2b,$$

$$6ax + 2b - 3(3ax^2 + 2bx + c) + 2(ax^3 + bx^2 + cx + d) = \\ = 2x^3 - 7x^2 - 4x + 10,$$

$$2ax^3 + (-9a + 2b)x^2 + (6a - 6b - 2c)x + 2b - 3c + 2d = \\ = 2x^3 - 7x^2 - 4x + 10,$$

$$\begin{array}{c|rcrcrcl} x^3 & 2a & & & & & = & 2, \\ x^2 & -9a & + & 2b & & & = & -7, \\ x & 6a & - & 6b & + & 2c & = & -4, \\ x^0 & & & 2b & - & 3c + 2d & = & 10. \end{array}$$

Решая полученную систему, находим

$$a = 1, \quad b = 1, \quad c = -2, \quad d = 1.$$

Получаем частное решение данного дифференциального уравнения:

$$\tilde{y} = x^3 + x^2 - 2x + 1.$$

Сопровождающее уравнение имеет вид

$$y'' - 3y' + 2y = 0.$$

Его характеристическое уравнение

$$k^2 - 3k + 2 = 0$$

имеет корни $k_1 = 2$, $k_2 = 1$, и поэтому общее решение сопровождающего уравнения есть

$$y_0 = C_1 e^{2x} + C_2 e^x.$$

Согласно теореме 30.2 общее решение данного уравнения есть

$$y = C_1 e^{2x} + C_2 e^{x} + x^3 + x^2 - 2x + 1.$$

В связи с этим примером может возникнуть предположение, что частное решение $\tilde{y}$ дифференциального уравнения

$$y'' + py' + qy = P_n(x)$$

надо искать в виде некоторого многочлена $Q_n(x)$ той же степени, для чего надо только подобрать его коэффициенты. Однако это предположение ошибочно.

Рассмотрим уравнение

$$y'' - y' = 4x^3 - 9x^2 - 6x - 2.$$

Попытаемся найти $\tilde{y}$ в виде многочлена той же, третьей, степени:

$$\tilde{y} = ax^3 + bx^2 + cx + d.$$

Дифференцируя $\tilde{y}$ и подставляя в левую часть уравнения, получаем

$$6ax + 2b - (3ax^2 + 2bx + c) = 4x^3 - 9x^2 - 6x - 2,$$

$$-3ax^2 + (6a - 2b)x + 2b - c = 4x^3 - 9x^2 - 6x - 2.$$

Тождественное равенство этих многочленов, степени которых различны (слева – многочлен второй степени, а справа – третьей), невозможно ни при каких a, b, c и d.

Наша попытка была неудачной, потому что мы не учли, что при дифференцировании степень многочлена снижается на единицу, а в левой части данного уравнения отсутствует член, содержащий неизвестную функцию y (присутствуют только члены, содержащие производные этой функции). Поэтому для того, чтобы левая часть данного уравнения стала многочленом той же (третьей) степени, надо в качестве y взять многочлен степени на единицу большей, т.е. многочлен четвертой степени. Однако при этом свободный член этого многочлена не будет учтен, так как при дифференцировании обратится в нуль. Стало быть, надо взять многочлен вида

$$Q_4(x) = ax^4 + bx^3 + cx^2 + dx$$

или вида

$$xQ_3(x) = x(ax^3 + bx^2 + cx + d).$$

(Напомним, что мы ищем лишь частное решение.)

С учетом высказанных здесь соображений будем искать решение данного уравнения

$$y'' - y' = 4x^3 - 9x^2 - 6x - 2$$

в виде $\tilde{y} = xQ_3(x) = x(ax^3 + bx^2 + cx + d) = ax^4 + bx^3 + cx^2 + dx$. Дифференцируем $\tilde{y}$ и подставляем $\tilde{y}'$ и $\tilde{y}''$ в левую часть данного уравнения:

$$\tilde{y}' = 4ax^3 + 3bx^2 + 2cx + d, \qquad \tilde{y}'' = 12ax^2 + 6bx + 2c,$$

$$12ax^2 + 6bx + 2c - (4ax^3 + 3bx^2 + 2cx + d) = 4x^3 - 9x^2 - 6x - 2,$$

$$-4ax^3 + (12a - 3b)x^2 + (6b - 2c)x + 2c - d = 4x^3 - 9x^2 - 6x - 2,$$

$$\begin{array}{l|lllll}
x^3 & -4a & & & = & 4, \\
x^2 & 12a - 3b & & & = & -9, \\
x & & 6b - 2c & & = & -6, \\
x^0 & & & 2c - d & = & -2.
\end{array}$$

Решая полученную систему, находим

$$a = -1, \quad b = -1, \quad c = 0, \quad d = 2.$$

Отсюда $\tilde{y} = -x^4 - x^3 + 2x$.

Найдем теперь y_0. Составим для этого характеристическое уравнение, найдем его корни и воспользуемся ими:

$$k^2 - k = 0,$$

$$k(k - 1) = 0.$$

Получаем общее решение сопровождающего уравнения:

$$y_0 = C_1 + C_2 e^x.$$

Общее решение данного уравнения:

$$y = C_1 + C_2 e^x - x^4 - x^3 + 2x.$$

Итак, при $q = 0$ и правой части $P_n(x)$ частное решение $\tilde{y}$ мы нашли в виде многочлена степени $n + 1$:

$$\tilde{y} = xQ_n(x).$$

Рассуждая аналогично, в случае когда не только q, но и $p = 0$, решение $\tilde{y}$ можно найти в виде

$$\tilde{y} = x^2 Q_n(x).$$

(Впрочем, в этом случае достаточно дважды проинтегрировать правую часть.)

Подведем первый итог и укажем способы нахождения частного решения дифференциального уравнения (30.20) в случае, когда правая часть – многочлен.

Если $\boldsymbol{f(x) = P_n(x)}$, где $P_n(x)$ – многочлен n-й степени, то частное решение $\tilde{y}$ надо искать в виде:

- $\tilde{y} = Q_n(x)$, если $q \neq 0$;
- $\tilde{y} = xQ_n(x)$, если $q = 0,\ p \neq 0$;
- $\tilde{y} = x^2Q_n(x)$, если $q = 0,\ p = 0$.

Эти правила сохраняются и в тех случаях, когда мы имеем дело с дифференциальными уравнениями более высокого порядка.

Пример **30.8.** Решить уравнение $y'' + 5y' = 10x + 12$.

Решение. Составляем характеристическое уравнение: $k^2 + 5k = 0$. Его корни $k_1 = 0$, $k_2 = -5$. Так как $q = 0$, то частное решение будем искать в виде

$$\tilde{y} = x(Ax + B) = Ax^2 + Bx.$$

Подставим его в исходное уравнение:

$$2A + 5(2Ax + B) = 10x + 12.$$

Приравняем коэффициенты при одинаковых степенях x:

$$\begin{cases} 10A = 10, \\ 2A + 5B = 12. \end{cases}$$

Отсюда $A = 1$, $B = 2$, $\tilde{y} = x^2 + 2x$. Общее решение сопровождающего уравнения $y_0 = C_1 + C_2e^{-5x}$. Поэтому общее решение данного уравнения:

$$y = C_1 + C_2e^{-5x} + x^2 + 2x.$$

Пример **30.9.** Найти общее решение дифференциального уравнения

$$y''' - 2y'' - y' + 2y = 2x^3 - 3x^2 - 12x + 8.$$

Решение. Здесь коэффициент при y в левой части отличен от нуля, поэтому ищем $\tilde{y}$ в виде многочлена той же степени, что и многочлен в правой части, т.е. в виде $Q_3(x)$:

$$\tilde{y} = ax^3 + bx^2 + cx + d.$$

Дифференцируем:

$$\tilde{y}' = 3ax^2 + 2bx + c,\quad \tilde{y}'' = 6ax + 2b,\quad \tilde{y}''' = 6a.$$

Подставляем $\widetilde{y}$ и ее производные:

$$6a - 2(6ax + 2b) - (3ax^2 + 2bx + c) + 2(ax^3 + bx^2 + cx + d) = \\ = 2x^3 - 3x^2 - 12x + 8,$$

$$2ax^3 + (-3a + 2b)x^2 + (-12a - 2b + 2c)x + 6a - 4b - c + 2d = \\ = 2x^3 - 3x^2 - 12x + 8,$$

$$\begin{array}{c|rcl} x^3 & 2a & = & 2, \\ x^2 & -3a + 2b & = & -3, \\ x & -12a - 2b + 2c & = & -12, \\ x^0 & 6a - 4b - c + 2d & = & 8. \end{array}$$

Находим коэффициенты: $a = 1, \quad b = 0, \quad c = 0, \quad d = 1.$ Итак,

$$\widetilde{y} = x^3 + 1.$$

Решаем сопровождающее уравнение:

$$y''' - 2y'' - y' + 2y = 0,$$

$$k^3 - 2k^2 - k + 2 = 0,$$

$$k_1 = 2,\ k_2 = 1,\ k_3 = -1.$$

Находим общее решение сопровождающего уравнения:

$$y_0 = C_1 e^{2x} + C_2 e^x + C_3 e^{-x}.$$

Окончательно получаем

$$y = C_1 e^{2x} + C_2 e^x + C_3 e^{-x} + x^3 + 1.$$

Рассмотрим еще один пример с дифференциальным уравнением, порядок которого выше второго.

***Пример* 30.10.** Найти общее решение уравнения

$$y^{(4)} + y''' = 24x + 36.$$

Решение. Ищем $\widetilde{y}$ в виде $x^3 Q_1(x)$:

$$\widetilde{y} = x^3(ax + b) = ax^4 + bx^3,$$

$$\widetilde{y}' = 4ax^3 + 3bx^2,\ \widetilde{y}'' = 12ax^2 + 6bx,\ \widetilde{y}''' = 24ax + 6b,\ \widetilde{y}^{(4)} = 24a.$$

Подставляя, получаем

$$24a + 24ax + 6b = 24x + 36,$$

$$\begin{cases} 24a = 24, \\ 24a + 6b = 36. \end{cases}$$

Отсюда $a = 1,\ b = 2;\ \widetilde{y} = x^4 + 2x^3.$

Решаем характеристическое уравнение:

$$k^4 + k^3 = 0,$$

$$k^3(k + 1) = 0,$$

$$k_1 = k_2 = k_3 = 0, \quad k_4 = -1.$$

Получаем

$$y_0 = C_1 + C_2 x + C_3 x^2 + C_4 e^{-x}.$$

Находим общее решение данного уравнения:

$$y = C_1 + C_2 x + C_3 x^2 + C_4 e^{-x} + x^4 + 2x^3.$$

Перейдем теперь к рассмотрению уравнений с более сложной правой частью.

B. Пусть $f(x) = e^{\alpha x} P_n(x)$. Рассмотренный ранее случай $f(x) = P_n(x)$ получается отсюда при $\alpha = 0$. (Формально говоря, его можно было и не рассматривать отдельно.) Итак, нужно решить уравнение

$$y'' + py' + qy = e^{\alpha x} P_n(x). \tag{30.21}$$

Постараемся свести эту задачу к предыдущей, т.е. к случаю, когда в правой части уравнения многочлен. Применим подстановку

$$y = e^{\alpha x} z,$$

где z – новая неизвестная функция $z = z(x)$. Находим производные:

$$y' = \alpha e^{\alpha x} z + e^{\alpha x} z',$$

$$y'' = \alpha^2 e^{\alpha x} z + \alpha e^{\alpha x} z' + \alpha e^{\alpha x} z' + e^{\alpha x} z'' = \alpha^2 e^{\alpha x} z + 2\alpha e^{\alpha x} z' + e^{\alpha x} z''.$$

Подставляем y, y' и y'' в левую часть уравнения (30.21) и выполняем очевидные преобразования:

$$\alpha^2 e^{\alpha x} z + 2\alpha e^{\alpha x} z' + e^{\alpha x} z'' + p(\alpha e^{\alpha x} z + e^{\alpha x} z') + q e^{\alpha x} z = e^{\alpha x} P_n(x),$$

$$e^{\alpha x} z'' + (2\alpha e^{\alpha x} + p e^{\alpha x}) z' + (\alpha^2 e^{\alpha x} + p\alpha e^{\alpha x} + q e^{\alpha x}) z = e^{\alpha x} P_n(x).$$

Сокращаем на $e^{\alpha x}$:

$$z'' + (2\alpha + p) z' + (\alpha^2 + p\alpha + q) z = P_n(x).$$

Получили уравнение уже рассмотренного вида

$$z'' + \bar{p} z' + \bar{q} z = P_n(x),$$

где $\bar{p} = 2\alpha + p$, $\bar{q} = \alpha^2 + p\alpha + q$. Следовательно, можно применить прежнее правило (см. с. 375):

- $\tilde{z} = Q_n(x)$, если $\bar{q} = \alpha^2 + p\alpha + q \neq 0$;
- $\tilde{z} = xQ_n(x)$, если $\bar{q} = \alpha^2 + p\alpha + q = 0$, $\bar{p} = 2\alpha + p \neq 0$;
- $\tilde{z} = x^2 Q_n(x)$, если $\bar{q} = \alpha^2 + p\alpha + q = 0$, $\bar{p} = 2\alpha + p = 0$.

Заметим, что условие $\overline{q} \neq 0$, т.е. $\alpha^2 + p\alpha + q \neq 0$, означает, что число α не является корнем характеристического уравнения $k^2 + pk + q = 0$. Поэтому если α – не корень характеристического уравнения, то ищем решение в виде $\tilde{z} = Q_n(x)$ или (с учетом того, что $y = e^{\alpha x}z$) в виде $\tilde{y} = e^{\alpha x}Q_n(x)$.

Прежде чем рассмотреть остальные условия, заметим, что решение характеристического уравнения $k^2 + pk + q = 0$ имеет вид

$$k_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}.$$

Если при этом $\frac{p^2}{4} - q \neq 0$, т.е. $\frac{p^2}{4} \neq q$, то оба корня характеристического уравнения различны и, очевидно, ни один из них не равен $-\frac{p}{2}$: $\alpha \neq -\frac{p}{2}$, т.е. $2\alpha + p \neq 0$.

Итак, если α – однократный корень характеристического уравнения, то $2\alpha + p \neq 0$, т.е. $\overline{q} = 0$, $\overline{p} \neq 0$. Если же α – корень характеристического уравнения и $2\alpha + p = 0$, то $\frac{p^2}{4} - q = 0$, т.е. α – двукратный корень характеристического уравнения. Суммируем все сказанное (и еще раз напомним, что $y = e^{\alpha x}z$).

Если $\boldsymbol{f(x) = e^{\alpha x}P_n(x)}$, то $\tilde{y}$ надо искать в виде:

- $\tilde{y} = e^{\alpha x}Q_n(x)$, если α – не корень характеристического уравнения;
- $\tilde{y} = xe^{\alpha x}Q_n(x)$, если α – однократный корень характеристического уравнения;
- $\tilde{y} = x^2e^{\alpha x}Q_n(x)$, если α – двукратный корень характеристического уравнения.

***Пример* 30.11.** Решить уравнение $y'' - 3y' + 2y = 2e^{3x}$.

Решение. Составляем характеристическое уравнение: $k^2 - 3k + 2 = 0$. Его корни $k_1 = 1$, $k_2 = 2$.

Очевидно, $\alpha = 3$ не является корнем характеристического уравнения. Поэтому $\tilde{y} = Ce^{3x}$, $\tilde{y}' = 3Ce^{3x}$, $\tilde{y}'' = 9Ce^{3x}$. Подставим все это в исходное уравнение:

$$9Ce^{3x} - 3 \cdot 3Ce^{3x} + 2Ce^{3x} = 2e^{3x}.$$

Получаем $C = 1$. Следовательно,

$$\tilde{y} = e^{3x}.$$

Очевидно, общее решение сопровождающего однородного уравнения есть

$$y_0 = C_1 e^x + C_2 e^{2x}.$$

Окончательно получаем

$$y = C_1 e^x + C_2 e^{2x} + e^{3x}.$$

Пример **30.12.** Решить уравнение $y'' + 2y' = (3x + 7)e^x$.

Решение. Здесь $\alpha = 1$. Составим и решим характеристическое уравнение:

$$k^2 + 2k = 0,$$

$$k_1 = 0, \quad k_2 = -2.$$

Видим, что α не является корнем характеристического уравнения. Поэтому $\tilde{y} = Q_1(x)e^x = (ax + b)e^x$. Дифференцируем $\tilde{y}$ и подставляем в правую часть уравнения:

$$\tilde{y}' = ae^x + (ax + b)e^x = (ax + a + b)e^x,$$

$$\tilde{y}'' = ae^x + (ax + a + b)e^x = (ax + 2a + b)e^x.$$

$$(ax + 2a + b)e^x + 2(ax + a + b)e^x = (3x + 7)e^x.$$

Сокращаем на e^x и приводим подобные члены:

$$3ax + 4a + 3b = 3x + 7.$$

Получаем систему

$$\begin{cases} 3a & = 3, \\ 4a + 3b & = 7. \end{cases}$$

Отсюда $a = 1$, $b = 1$; $\tilde{y} = (x + 1)e^x$.

Общее решение сопровождающего однородного уравнения:

$$y_0 = C_1 + C_2 e^{-2x}.$$

Окончательно получаем

$$y = C_1 + C_2 e^{-2x} + (x + 1)e^x.$$

(Обратите внимание: здесь $q = 0$, но мы этого не учитывали, в отличие от случая А, так как в этом случае учитывается лишь то, является ли α корнем характеристического уравнения.)

Пример **30.13.** Решить уравнение $y'' + y' - 6y = (10x + 2)e^{2x}$.

Решение. Решаем характеристическое уравнение:

$$k^2 + k - 6 = 0,$$

$$k_1 = 2, \quad k_2 = -3.$$

Итак, $\alpha = 2$ – однократный корень.

Общее решение сопровождающего однородного уравнения

$y_0 = C_1 e^{2x} + C_2 e^{-3x}$.

Находим $\tilde{y}$. Очевидно, $\tilde{y}$ надо искать в виде

$\tilde{y} = xQ_1(x)e^{2x} = x(ax + b)e^{2x} = (ax^2 + bx)e^{2x}$;

$\tilde{y}' = (2ax + b)e^{2x} + 2(ax^2 + bx)e^{2x} = (2ax^2 + 2ax + 2bx + b)e^{2x}$,

$\tilde{y}'' = (4ax + 2a + b)e^{2x} + 2(2ax^2 + 2ax + 2bx + b)e^{2x} =$

$= (4ax^2 + 8ax + 4bx + 2a + 3b)e^{2x}$.

Подставляем все эти выражения в исходное уравнение:

$(4ax^2 + 8ax + 4bx + 2a + 3b)e^{2x} + (2ax^2 + 2ax + 2bx + b)e^{2x} -$

$- 6(ax^2 + bx)e^{2x} = (10x + 2)e^{2x}$.

Сокращая на e^{2x} и приводя подобные члены, получаем

$$\begin{cases} 10a = 10, \\ 2a - 2b = 2, \end{cases}$$

т.е. $a = 1$, $b = 0$. Следовательно, $\tilde{y} = x^2 e^{2x}$. Получаем

$y = C_1 e^{2x} + C_2 e^{-3x} + x^2 e^{2x}$.

***Пример* 30.14.** Решить уравнение $y'' + 2y' + y = (6x - 2)e^{-x}$.

Решение. Здесь $\alpha = -1$ является двойным корнем характеристического уравнения $k^2 + 2k + 1 = 0$. Поэтому частное решение $\tilde{y}$ ищем в виде $\tilde{y} = x^2 Q_1(x)e^{-x} = x^2(ax + b)e^{-x} = (ax^3 + bx^2)e^{-x}$. Имеем:

$\tilde{y}' = (3ax^2 + 2bx)e^{-x} - (ax^3 + bx^2)e^{-x} =$

$= (-ax^3 + (3a - b)x^2 + 2bx)e^{-x}$;

$\tilde{y}'' = (-3ax^2 + 2(3a - b)x + 2b)e^{-x} - (-ax^3 + (3a - b)x^2 + 2bx)e^{-x} =$

$= (ax^3 + (-6a + b)x^2 + 6ax - 4bx + 2b)e^{-x}$.

Подставляем в исходное уравнение; после приведения подобных членов получаем

$(6ax + 2b)e^{-x} = (6x - 2)e^{-x}$.

Отсюда $a = 1$, $b = -1$; $\tilde{y} = (x^3 - x^2)e^{-x}$.

Находим также $y_0 = C_1 e^{-x} + C_2 x e^{-x}$.

Получаем общее решение:

$$y = C_1 e^{-x} + C_2 x e^{-x} + (x^3 - x^2)e^{-x}.$$

Естественно возникает вопрос о нахождении частного решения дифференциального уравнения вида

$$y'' + py' + qy = f_1(x) + f_2(x),$$

где $f_1(x)$ и $f_2(x)$ – функции различных видов (например, $f_1(x) = ax^2 + bx + c$, $f_2(x) = e^x$). Имеет место следующее утверждение.

Лемма **30.3.** Если y_1 – решение дифференциального уравнения

$$y'' + py' + qy = f_1(x),$$

а y_2 – решение дифференциального уравнения

$$y'' + py' + qy = f_2(x),$$

то сумма этих решений $y = y_1 + y_2$ является решением дифференциального уравнения

$$y'' + py' + qy = f_1(x) + f_2(x).$$

(Эту лемму называют **принципом наложения**.)

Доказательство. Лемма доказывается несложно – непосредственной подстановкой. Подставим сумму $y = y_1 + y_2$ в левую часть уравнения. Получим выражение

$$(y_1 + y_2)'' + p(y_1 + y_2)' + q(y_1 + y_2).$$

Перегруппировав члены, получаем

$$(y_1'' + py_1' + qy_1) + (y_2'' + py_2' + qy_2).$$

Но так как y_1 есть решение уравнения $y'' + py' + qy = f_1(x)$, то выражение в первых скобках тождественно равно $f_1(x)$. По аналогичной причине выражение во второй скобке равно $f_2(x)$. Итак, левая часть тождественно равна $f_1(x) + f_2(x)$. Уравнение обращается в тождество: $f_1(x) + f_2(x) \equiv f_1(x) + f_2(x)$. Доказательство закончено.

Заметим, что доказанное утверждение верно также и для случая, когда коэффициенты p и q зависят от x: $p = p(x)$, $q = q(x)$.

Пример **30.15.** Решить уравнение $y'' + 5y' = 10x + 12 + 6e^x$.

Решение. Решая уравнение $y'' + 5y' = 10x + 12$, получаем: $y_0 = C_1 + C_2 e^{-5x}$, $\tilde{y}_1 = x^2 + 2x$ (см. пример 30.8). Очевидно, $\tilde{y}_2 = e^x$. Поэтому общее решение $y = C_1 + C_2 e^{-5x} + x^2 + 2x + e^x$.

С. Пусть $f(x) = P_m(x)e^{\alpha x}\cos\beta x + Q_n(x)e^{\alpha x}\sin\beta x$. В этом случае можно воспользоваться приемом, примененным в предыдущем случае, если перейти от тригонометрических функций к показательным. В более подробном курсе математики рассматривается *формула Эйлера*, выражающая показательную функцию с мнимым показателем через тригонометрические функции (здесь i – мнимая единица, $i^2 = -1$):

$$e^{ix} = \cos x + i\sin x. \tag{30.22}$$

Заменяя в этой формуле x на $-x$, получим

$$e^{-ix} = \cos x - i\sin x. \tag{30.23}$$

Из равенств (30.22) и (30.23)легко найти $\cos x$ и $\sin x$:

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}, \qquad \sin x = \frac{e^{ix} - e^{-ix}}{2i}.$$

Эти формулы также называют *формулами Эйлера*. Применяя их, получаем

$$f(x) = P_m(x)e^{\alpha x}\,\frac{e^{i\beta x} + e^{-i\beta x}}{2} + Q_n(x)e^{\alpha x}\,\frac{e^{i\beta x} - e^{-i\beta x}}{2i},$$

или

$$f(x) = \left[\frac{1}{2}P_m(x) + \frac{1}{2i}Q_n(x)\right]e^{(\alpha+i\beta)x} + \left[\frac{1}{2}P_m(x) - \frac{1}{2i}Q_n(x)\right]e^{(\alpha-i\beta)x}.$$

Здесь в квадратных скобках стоят многочлены, степени которых равны наибольшей из степеней $P_m(x)$ и $Q_n(x)$, т.е. наибольшему из чисел m и n. Таким образом, правая часть уравнения имеет вид, рассмотренный в случае В. При этом можно доказать (мы не приводим этого доказательства), что существует частное решение $\tilde{y}$, не содержащее комплексных чисел.

Итак, если $\boldsymbol{f(x) = P_m(x)e^{\alpha x}\cos\beta x + Q_n(x)e^{\alpha x}\sin\beta x}$, то $\tilde{y}$ следует искать в виде:

- $\tilde{y} = u(x)e^{\alpha x}\cos\beta x + v(x)e^{\alpha x}\sin\beta x$, если $\alpha + i\beta$ – не корень характеристического уравнения;
- $\tilde{y} = x\,(u(x)e^{\alpha x}\cos\beta x + v(x)e^{\alpha x}\sin\beta x)$, если $\alpha + i\beta$ – корень характеристического уравнения.

Здесь $u(x)$ и $v(x)$ – многочлены, степень которых равна наибольшей степени многочленов $P_m(x)$ и $Q_n(x)$.

Замечание. Отметим, что указанные формы частных решений сохраняются и в том случае, когда в правой части дифференциального уравнения один из многочленов, $P_m(x)$ или $Q_n(x)$, тождественно равен нулю, т.е. либо

$$f(x) = P_m(x)e^{\alpha x}\cos\beta x,$$

либо

$$f(x) = Q_n(x)e^{\alpha x}\sin\beta x.$$

Рассмотрим подробнее более простой случай – частный вариант случая C.

C_0. Пусть $f(x) = M\cos\beta x + N\sin\beta x$. Применяя формулы Эйлера, перепишем дифференциальное уравнение

$$y'' + py' + qy = M\cos\beta x + N\sin\beta x$$

в виде

$$y'' + py' + qy = M\,\frac{e^{i\beta x} + e^{-i\beta x}}{2} + N\,\frac{e^{i\beta x} - e^{-i\beta x}}{2i}.$$

Положив для краткости

$$\frac{M}{2} + \frac{N}{2i} = M_1, \qquad \frac{M}{2} - \frac{N}{2i} = N_1,$$

получаем

$$y'' + py' + qy = M_1e^{i\beta x} + N_1e^{-i\beta x}. \tag{*}$$

Согласно лемме 30.3 решение $\tilde{y}$ нашего дифференциального уравнения является суммой решений $\tilde{y}_1$ и $\tilde{y}_2$ уравнений

$$y'' + py' + qy = M_1e^{i\beta x}, \qquad y'' + py' + qy = N_1e^{-i\beta x}.$$

Заметим, что мнимое число $i\beta$ не может быть двукратным корнем квадратного уравнения $k^2 + pk + q = 0$, коэффициенты которого p и q являются действительными числами.

Частное решение $\tilde{y}_1$ уравнения $y'' + py' + qy = M_1e^{i\beta x}$ имеет вид:

- $\tilde{y}_1 = Ae^{i\beta x}$, если $i\beta$ – не корень характеристического уравнения;
- $\tilde{y}_1 = xAe^{i\beta x}$, если $i\beta$ – корень характеристического уравнения.

Частное решение $\tilde{y}_2$ уравнения $y'' + py' + qy = N_1e^{-i\beta x}$ имеет вид:

- $\tilde{y}_2 = Be^{-i\beta x}$, если $-i\beta$ – не корень характеристического уравнения;
- $\tilde{y}_2 = xBe^{-i\beta x}$, если $-i\beta$ – корень характеристического уравнения.

Очевидно, числа $i\beta$ и $-i\beta$ либо оба являются, либо оба не являются корнями характеристического уравнения (так как если $\alpha + i\beta$ есть корень квадратного уравнения, то $\alpha - i\beta$ также есть корень этого уравнения).

Поэтому частное решение уравнения (*) будет иметь вид:

- $\tilde{y} = Ae^{i\beta x} + Be^{-i\beta x}$, если $\pm i\beta$ – не корни характеристического уравнения;
- $\tilde{y} = x(Ae^{i\beta x} + Be^{-i\beta x})$, если $\pm i\beta$ – корни характеристического уравнения.

Применим формулу Эйлера [см. (30.22) и (30.23)], получаем

$$Ae^{i\beta x} + Be^{-i\beta x} = (A + B)\cos\beta x + i(A - B)\sin\beta x =$$
$$= a\cos\beta x + b\sin\beta x,$$

где для краткости $A + B = a$, $i(A - B) = b$.

Отсюда получаем правило для нахождения частного решения $\tilde{y}$ дифференциального уравнения

$$y'' + py' + qy = M\cos\beta x + N\sin\beta x.$$

Итак, если $\boldsymbol{f(x) = M\cos\beta x + N\sin\beta x}$, то частное решение $\tilde{y}$ надо искать в виде:

- $\tilde{y} = a\cos\beta x + b\sin\beta x$, если $i\beta$ – не корень характеристического уравнения;
- $\tilde{y} = x(a\cos\beta x + b\sin\beta x)$, если $i\beta$ – корень характеристического уравнения.

***Пример* 30.16.** Решить уравнение $y'' + y' - 2y = 8\sin 2x$.

Решение. Характеристическое уравнение $k^2 + k - 2 = 0$ имеет корни $k_1 = 1$, $k_2 = -2$. Здесь $\beta = 2$, следовательно, $i\beta$ не корень характеристического уравнения. Стало быть, частное решение $\tilde{y}$ надо искать в виде

$$\tilde{y} = C\cos 2x + D\sin 2x,$$

$$\tilde{y}' = -2C\sin 2x + 2D\cos 2x,$$

$$\tilde{y}'' = -4C\cos 2x - 4D\sin 2x.$$

Подставляем:

$$-4C\cos 2x - 4D\sin 2x - 2C\sin 2x + 2D\cos 2x -$$
$$- 2(C\cos 2x + D\sin 2x) = 8\sin 2x,$$

$$(-6C + 2D)\cos 2x + (-2C - 6D)\sin 2x = 8\sin 2x.$$

Приравниваем коэффициенты при $\cos 2x$ и $\sin 2x$:

$$\begin{cases} -6C + 2D = 0, \\ -2C - 6D = 8. \end{cases}$$

Отсюда $C = -\frac{2}{5}$, $D = -\frac{6}{5}$. Следовательно,

$$\tilde{y} = -\frac{2}{5}\cos 2x - \frac{6}{5}\sin 2x.$$

Находим y_0:

$$y_0 = C_1 e^x + C_2 e^{-2x}.$$

Получаем общее решение:

$$y = C_1 e^x + C_2 e^{-2x} - \frac{1}{5}(2\cos 2x + 6\sin 2x).$$

***Пример* 30.17.** Решить уравнение $y'' + 4y = \cos 2x$.

Решение. Характеристическое уравнение $k^2 + 4k = 0$ имеет корни $k_1 = 2i$, $k_2 = -2i$. Здесь $\beta = 2$, следовательно, $i\beta$ есть корень характеристического уравнения, стало быть, частное решение $\tilde{y}$ надо искать в виде

$$\tilde{y} = x(C\cos 2x + D\sin 2x).$$

Дифференцируем:

$$\tilde{y}' = C\cos 2x + D\sin 2x + 2x(-C\sin 2x + D\cos 2x),$$

$$\tilde{y}'' = 2(-C\sin 2x + D\cos 2x) + 2(-C\sin 2x + D\cos 2x) +$$
$$+ 4x(-C\cos 2x - D\sin 2x) = 4(-C\sin 2x + D\cos 2x) +$$
$$+ 4x(-C\cos 2x - D\sin 2x).$$

Подставляем в уравнение, получаем:

$$4(-C\sin 2x + D\cos 2x) + 4x(-C\cos 2x - D\sin 2x) +$$
$$+ 4x(C\cos 2x + D\sin 2x) = \cos 2x.$$

Приводя подобные члены, получаем

$$-4C\sin 2x + 4D\cos 2x = \cos 2x.$$

Приравнивая коэффициенты при $\cos 2x$ и $\sin 2x$, получаем: $-4C = 0$, $4D = 1$; отсюда $C = 0$, $D = \frac{1}{4}$. Таким образом, частное решение данного уравнения есть

$$\tilde{y} = \frac{x}{4}\sin 2x.$$

Общее решение сопровождающего однородного уравнения:

$$y_0 = C_1\cos 2x + C_2\sin 2x.$$

Окончательно получаем

$$y = C_1\cos 2x + C_2\sin 2x + \frac{x}{4}\sin 2x.$$

Пример **30.18.** Решить уравнение $y'' - y = e^{2x}(6\cos x - 2\sin x)$.

Решение. Здесь $\alpha = 2$, $\beta = 1$. Характеристическое уравнение $k^2 - 1 = 0$ имеет корни $k_1 = 1$, $k_2 = -1$. Так как $\alpha + i\beta = 2 + i$ не является корнем характеристического уравнения, то частное решение ищем в виде

$$\tilde{y} = e^{2x}(C\cos x + D\sin x).$$

Находим $\tilde{y}'$ и $\tilde{y}''$:

$$\tilde{y}' = 2e^{2x}(C\cos x + D\sin x) + e^{2x}(-C\sin x + D\cos x),$$

$$\tilde{y}'' = 4e^{2x}(C\cos x + D\sin x) + 2e^{2x}(-C\sin x + D\cos x) +$$
$$+ 2e^{2x}(-C\sin x + D\cos x) + e^{2x}(-C\cos x - D\sin x) =$$
$$= e^{2x}(3C\cos x + 4D\cos x + 3D\sin x - 4C\sin x).$$

Подставляя полученные выражения в исходное уравнение и приводя подобные члены, получаем (после сокращения на e^{2x}):

$$(2C + 4D)\cos x + (2D - 4C)\sin x = 6\cos x - 2\sin x.$$

Приравнивая коэффициенты при $\cos x$ и $\sin x$, получаем систему

$$\begin{cases} 2C + 4D = 6, \\ -4C + 2D = -2. \end{cases}$$

Решая эту систему, находим $C = 1$, $D = 1$.

Частное решение данного уравнения:

$$\tilde{y} = e^{2x}(\cos x + \sin x).$$

Общее решение сопровождающего уравнения:

$y_0 = C_1 e^x + C_2 e^{-x}$.

Получаем общее решение данного уравнения:

$y = C_1 e^x + C_2 e^{-x} + e^{2x}(\cos x + \sin x)$.

Вопросы

1. Каков общий вид дифференциального уравнения второго порядка?
2. Что такое общее решение дифференциального уравнения второго порядка?
3. Какие дифференциальные уравнения допускают понижение порядка?
4. Что называется линейным дифференциальным уравнением n-го порядка?
5. Какое линейное дифференциальное уравнение n-го порядка называется однородным?
6. Каковы свойства решений линейного однородного дифференциального уравнения?
7. Какая система функций называется линейно независимой?
8. Как выглядит определитель Вронского для двух функций?
9. Какова структура общего решения однородного линейного дифференциального уравнения?
10. Что такое характеристическое уравнение?
11. Как выглядит общее решение однородного линейного дифференциального уравнения второго порядка в случае, когда корни характеристического уравнения совпадают?
12. Какова структура общего решения неоднородного линейного дифференциального уравнения? Как можно получить общее решение дифференциального уравнения $y'' + py' + qy = f(x)$, если известно общее решение дифференциального уравнения $y'' + py' + qy = 0$?
13. В каком виде следует искать частное решение дифференциального уравнения $y'' + py' + qy = P_n(x)$, когда правая часть $P_n(x)$ есть многочлен n-й степени? Всегда ли это частное решение есть также многочлен n-й степени?

14. В каком виде следует искать частное решение дифференциального уравнения $y'' + py' + qy = e^{\alpha x} P_n(x)$?
15. В чем заключается принцип наложения?
16. В чем заключается правило для нахождения частного решения дифференциального уравнения $y'' + py' + qy = M\cos\beta x + N\sin\beta x$?

Глава 31
РАЗНОСТНЫЕ УРАВНЕНИЯ

31.1. ОСНОВНЫЕ ПОНЯТИЯ

В математических приложениях наряду с функциями непрерывного аргумента приходится иметь дело также с функциями д и с к р е т н о г о аргумента, т.е. с функциями, заданными на конечном (или счетном) дискретном множестве. Примерами таких функций являются функции, заданные таблицами, числовые последовательности (см. главу 14), ряды (см. раздел IX).

Функции дискретного аргумента обычно обозначают $f(x_k)$ или $y(x_k)$. Расстояние $h_k = x_{k+1} - x_k$, $k = 1, 2, \ldots$ между соседними значениями аргумента могут быть любыми положительными числами. Однако наибольший интерес представляет случай, когда величины h_k являются одинаковыми: $h_k = h$ при всех $k = 1, 2, \ldots$. Это число h называют обычно **шагом дискретизации**. В этом случае $x_k = kh$, а функция $f(x_k)$ становится функцией номера k, т.е. $f(x_k) = f(kh)$, $k = 1, 2, \ldots$.

О п р е д е л е н и е. **Сеткой на отрезке** $[a, b]$ называется любое конечное множество точек этого отрезка. Точки сетки называются ее **узлами**.

Заметим, что мы уже имели дело с сетками и их узлами – когда определяли понятие определенного интеграла и когда занимались приближенным вычислением определенных интегралов по формулам прямоугольников и трапеций и по формуле Симпсона (см. § 22.5).

Сетка называется **равномерной**, если ее узлы делят отрезок $[a, b]$ на р а в н ы е отрезки. Длина h такого частичного отрезка называется **шагом сетки**. Очевидно, $h = \dfrac{b-a}{n}$, где n – число частичных отрезков.

Множество точек на $[a, b]$ вида

$$\{x_k = a + kh, \quad k = 0, 1, 2, \ldots, n\}$$

образует **равномерную сетку с шагом** h.

В случае когда узлы сетки делят отрезок $[a, b]$ на н е р а в н ы е отрезки, сетка называется **неравномерной**.

О п р е д е л е н и е. Функция, определенная в точках сетки, называется **сеточной функцией**.

Соответствующие значения сеточной функции в узлах сетки обычно обозначают через y_k или f_k. Если сеточная функция определена на равномерной сетке, то ее значения обозначают через $y(k)$, где k – номер узла сетки ($k = 0, 1, 2, ..., n$). В этом случае сеточная функция рассматривается как функция целочисленного аргумента.

Для того чтобы из функции непрерывного аргумента $y(x)$ получить соответствующую сеточную функцию $y(kh)$, надо аргумент x заменить на kh.

***П р и м е р* 31.1.** Для функции $y = 4x^2 + x$, определенной на отрезке $[0, 1]$, составить равномерную сетку при $n = 4$ и соответствующую сеточную функцию.

Р е ш е н и е. Очевидно, шаг сетки $h = 0{,}25$. Получаем сетку $\{0; 0{,}25; 0{,}5; 0{,}75; 1\}$. Сеточная функция также есть множество, состоящее из пяти чисел: $\{0; 0{,}5; 1{,}5; 3; 5\}$.

Аналогом *первой производной* функции непрерывного аргумента является *первая разность* сеточной функции.

Разность первого порядка, или **первая разность**, сеточной функции $y(k)$, обозначаемая через $\Delta y(k)$, определяется так:

$$\Delta y(k) = y(k+1) - y(k). \tag{31.1}$$

Вторая разность $\Delta^2 y(k)$ функции $y(k)$ определяется как первая разность от ее первой разности:

$$\Delta^2 y(k) = \Delta y(k+1) - \Delta y(k). \tag{31.2}$$

Подставляя сюда значения $\Delta y(k)$ и $\Delta y(k+1)$, определяемые по формуле (31.1), получаем

$$\Delta^2 y(k) = y(k+2) - 2y(k+1) + y(k).$$

Аналогично определяются $\Delta^3 y(k)$ и вообще разность любого порядка. При этом разность m-го порядка $\Delta^m y(k)$ можно представить как *линейную комбинацию* значений $y(k)$, $y(k+1)$, ..., $y(k+m)$. В частности,

$$\Delta^3 y(k) = \Delta^2 y(k+1) - \Delta^2 y(k) =$$

$$= y(k+3) - 3y(k+2) + 3y(k+1) - y(k).$$

Пример **31.2.** Найти все разности до m-го порядка включительно для функции $y(k) = e^{\alpha k}$.

Решение: $\Delta y(k) = e^{\alpha(k+1)} - e^{\alpha k} = (e^{\alpha} - 1)e^{\alpha k}$.

Мы видим, что первая разность пропорциональна самой функции $e^{\alpha k}$. Следовательно,

$$\Delta^2 y(k) = (e^{\alpha} - 1)^2 e^{\alpha k}, \quad \Delta^3 y(k) = (e^{\alpha} - 1)^3 e^{\alpha k}, \ \dots,$$

$$\Delta^m y(k) = (e^{\alpha} - 1)^m e^{\alpha k}.$$

Определение. Уравнение вида

$$F(k, \ y(k), \ \Delta y(k), \ \dots, \ \Delta^m y(k)) = 0, \tag{31.3}$$

где $y(k)$ – неизвестная функция целочисленного аргумента, а $\Delta y(k), \ \dots, \ \Delta^m y(k)$ – ее разности, называется **разностным уравнением**, или **уравнением в конечных разностях**, m-го порядка.

Решением разностного уравнения называется всякая сеточная функция, обращающая его в тождество.

Ранее мы убедились, что конечные разности различных порядков могут быть выражены через значения исходной сеточной функции. Поэтому уравнение (31.3) можно представить в виде

$$F_1(k, \ y(k+m), \ \dots, \ y(k+1), \ y(k)) = 0. \tag{31.4}$$

Разностные уравнения имеют многочисленные приложения в моделях экономической динамики с дискретным временем.

31.2. ЛИНЕЙНЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ

Определение. Разностное уравнение вида

$$a_0(k)y(k+m) + a_1(k)y(k+m-1) + \dots + a_m(k)y(k) = f(k), \tag{31.5}$$

где $a_j(k)$ и $f(k)$ – известные функции, а $y(k+j)$ – неизвестная функция от k $(j = 0, 1, 2, \dots, m)$, причем $a_m(k)$ и $a_0(k)$ не равны нулю ни при каком k, называется **линейным разностным уравнением m-го порядка.**

В случае когда коэффициенты $a_0, \ a_1, \dots, \ a_m$ являются константами, методы решения таких уравнений аналогичны методам решения линейных дифференциальных уравнений с постоянными коэффициентами.

Вместе с неоднородным уравнением

$$a_0 y(k+m) + a_1 y(k+m-1) + \dots + a_m y(k) = f(k) \tag{31.6}$$

рассматривается соответствующее однородное уравнение

$$a_0 y(k+m) + a_1 y(k+m-1) + \dots + a_m y(k) = 0. \tag{31.7}$$

Для разностных уравнений (в частности, для линейных разностных уравнений), так же как и для их дифференциальных аналогов, определяются понятия общего и частного решений.

Общее решение уравнения (31.6) имеет вид

$$y(k) = \varphi(k, c_1, ..., c_m),$$

где $c_1, ..., c_m$ – произвольные постоянные; их число равно порядку уравнения.

Частное решение уравнения (31.6) выделяется заданием значений функции $y(k)$ в m произвольных, но расположенных подряд точках.

Так же как и для линейных дифференциальных уравнений, определяется понятие линейно независимой системы решений, доказывается, что общее решение уравнения (31.6) имеет вид

$$y(k) = y_0(k) + \tilde{y}(k), \tag{31.8}$$

где $y_0(k)$ – общее решение соответствующего однородного уравнения (31.7), а $\tilde{y}(k)$ – некоторое частное решение исходного уравнения (31.6).

В дальнейшем ограничимся рассмотрением разностных уравнений второго порядка. Результаты, которые будут получены, могут быть распространены и на разностные уравнения более высоких порядков.

Итак, рассмотрим *однородное линейное разностное уравнение второго порядка*:

$$y(k+2) + py(k+1) + qy(k) = 0. \tag{31.9}$$

Решение этого уравнения будем искать в виде

$$y(k) = \lambda^k.$$

После очевидных упрощений получим *характеристическое уравнение*:

$$\lambda^2 + p\lambda + q = 0. \tag{31.10}$$

Возможны три варианта.

1. Оба корня λ_1 и λ_2 уравнения (31.10) действительны и различны. В этом случае *общее решение* имеет вид

$$y_0(k) = c_1\lambda_1^k + c_2\lambda_2^k. \tag{31.11}$$

Пример **31.3.** Дано разностное уравнение

$$y(k+2) - 5y(k+1) + 6y(k) = 0.$$

Найти его общее решение.

Р е ш е н и е. Характеристическое уравнение $\lambda^2 - 5\lambda + 6 = 0$ имеет два различных действительных корня: $\lambda_1 = 3$, $\lambda_2 = 2$. Следовательно, в соответствии с формулой (31.11) общим решением заданного уравнения будет сетчатая функция

$$y_0(k) = c_1 \cdot 3^k + c_2 \cdot 2^k.$$

2. Оба корня д е й с т в и т е л ь н ы и р а в н ы между собой: $\lambda_1 = \lambda_2 = \lambda$. Тогда *общее решение* имеет вид

$$y_0(k) = c_1\lambda^k + c_2 k\lambda^k.$$

3. Характеристическое уравнение имеет к о м п л е к с н ы е с о п р я ж е н н ы е корни $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$.

Представим корни уравнения в тригонометрической форме: $\lambda_1 = r(\cos\varphi + i\sin\varphi)$, $\lambda_2 = r(\cos\varphi - i\sin\varphi)$, где модуль $r = \sqrt{\alpha^2 + \beta^2}$, а аргумент φ определяется соотношением $\operatorname{tg}\varphi = \frac{\beta}{\alpha}$.

Общее решение имеет вид

$$y_0(k) = r^k(c_1\cos k\varphi + c_2\sin k\varphi).$$

Обратимся теперь к *н е о д н о р о д н о м у линейному разностному уравнению второго порядка*:

$$y(k+2) + py(k+1) + qy(k) = f(k). \qquad (31.12)$$

Его общее решение имеет вид (31.8). Для нахождения частного решения $\tilde{y}(k)$ уравнения (31.12) часто применяют *метод неопределенных коэффициентов*.

***П р и м е р* 31.4.** Решить уравнение

$$y(k+2) - 7y(k+1) + 10y(k) = 4 \cdot 6^k.$$

Р е ш е н и е. Для нахождения общего решения соответствующего однородного уравнения составим характеристическое уравнение:

$$\lambda^2 - 7\lambda + 10 = 0.$$

Его корни $\lambda_1 = 5$, $\lambda_2 = 2$. Следовательно,

$$y_0(k) = c_1 \cdot 5^k + c_2 \cdot 2^k.$$

Для нахождения частного решения $\tilde{y}(k)$ исходного уравнения воспользуемся методом неопределенных коэффициентов. Будем искать $\tilde{y}(k)$ в виде $\tilde{y}(k) = c \cdot 6^k$. Подставляя это выражение в данное уравнение, получим:

$$c \cdot 6^{k+2} - 7c \cdot 6^{k+1} + 10\, c \cdot 6^k = 4 \cdot 6^k,$$

$$c(36 - 42 + 10) \cdot 6^k = 4 \cdot 6^k.$$

Следовательно, $c = 1$, а значит,

$$\widetilde{y}(k) = 6^k.$$

Складывая $y_0(k)$ и $\widetilde{y}(k)$, получаем общее решение уравнения:

$$y(k) = c_1 \cdot 5^k + c_2 \cdot 2^k + 6^k.$$

31.3. МОДЕЛЬ ДЕЛОВОГО ЦИКЛА САМУЭЛЬСОНА — ХИКСА

Рассмотрим в качестве примера применения разностных уравнений известную в макроэкономической теории **модель делового цикла Самуэльсона – Хикса**. В этой модели используется предположение, что объем инвестиций прямо пропорционален приросту национального дохода. Это предположение – уже известный нам *принцип акселерации* – описывается следующим уравнением:

$$I(k) = \chi\,[y(k-1) - y(k-2)], \qquad (31.13)$$

где χ – коэффициент пропорциональности, называемый фактором акселерации ($\chi > 0$), $I(k)$ – величина инвестиций в период k (в k-й календарный год), а $y(k-1)$, $y(k-2)$ – величины национального дохода в предшествующих периодах – в $(k-1)$-м и $(k-2)$-м соответственно.

Кроме того, предполагается, что потребление $C(k)$ в рассматриваемом k-м периоде также линейно зависит от величины национального дохода $y(k-1)$ за предшествующий период:

$$C(k) = ay(k-1) + b, \qquad (31.14)$$

а доход $y(k)$ делится между производителями и потребителями. Поэтому

$$y(k) = C(k) + I(k). \qquad (31.15)$$

Подставим в (31.15) выражение для $I(k)$ из (31.13), а также выражение для $C(k)$ из (31.14):

$$y(k) = ay(k-1) + b + \chi\,[y(k-1) - y(k-2)].$$

Получаем так называемое **уравнение Хикса**:

$$y(k) - (a + \chi)y(k-1) + \chi y(k-2) = b. \qquad (31.16)$$

Если допустить, что на протяжении рассматриваемых периодов времени величины a и χ постоянны, то уравнение (31.16) есть *линейное неоднородное разностное уравнение второго порядка с постоянными коэффициентами*.

Если предположить, что величина национального дохода оставалась постоянной на протяжении рассматриваемых периодов, т.е.

$$y(k) = y(k-1) = y(k-2) = \tilde{y},$$

то можно найти простое частное решение уравнения (31.16):

$$\tilde{y} = (a+\chi)\tilde{y} - \chi\tilde{y} + b.$$

Отсюда

$$\tilde{y} = b(1-a)^{-1}. \qquad (31.17)$$

Выражение $(1-a)^{-1}$ в формуле (31.17) называется *мультипликатором Кейнса*.

***Пример* 31.5.** Рассмотреть модель Самуэльсона – Хикса при условии, что $a = 0{,}48$, $\chi = 0{,}72$, $b = 1{,}3$. Найти общее решение уравнения Хикса.

Решение. В данном случае уравнение (31.16) имеет вид

$$y(k) - 1{,}2y(k-1) + 0{,}72y(k-2) = 1{,}3.$$

Частное решение этого уравнения, согласно (31.17), есть

$$\tilde{y}(k) = \frac{1{,}3}{1-0{,}48} = 2{,}5.$$

Напишем характеристическое уравнение:

$$\lambda^2 - 1{,}2\lambda + 0{,}72 = 0.$$

Его корни $\lambda_{1,2} = 0{,}6 \pm 0{,}6i = 0{,}6\sqrt{2}\left(\cos\frac{\pi}{4} \pm i\sin\frac{\pi}{4}\right)$.

Общее решение соответствующего однородного уравнения есть

$$y_0(k) = (0{,}6\sqrt{2})^k\left(c_1\cos\frac{\pi k}{4} + c_2\sin\frac{\pi k}{4}\right).$$

Получаем общее решение данного уравнения:

$$y(k) = 2{,}5 + (0{,}6\sqrt{2})^k\left(c_1\cos\frac{\pi k}{4} + c_2\sin\frac{\pi k}{4}\right).$$

В рассмотренном примере динамика носит колебательный характер с затухающей амплитудой. Очевидно, при комплексных сопряженных корнях характеристического уравнения, по модулю превышающих единицу, динамика была бы растущей. Вообще, в зависимости от значений a и χ динамика может быть растущей или затухающей и при этом иметь или не иметь колебательный характер.

Вопросы

1. Какие функции называются сеточными?
2. Как определяются первая, вторая и последующие разности сеточной функции?
3. Какие уравнения называются конечно-разностными?
4. Что такое характеристическое уравнение для однородного линейного разностного уравнения?
5. Как найти общее решение неоднородного разностного уравнения?

Раздел IX
РЯДЫ

Глава 32
ЧИСЛОВЫЕ РЯДЫ

32.1. ПОНЯТИЕ ЧИСЛОВОГО РЯДА

Определение. Пусть дана числовая последовательность $a_1, a_2, ..., a_n, ...$. Выражение вида

$$a_1 + a_2 + ... + a_n + ... \qquad (32.1)$$

называется **числовым рядом**, или просто **рядом**.

Числа $a_1, a_2, ..., a_n, ...$ называются **членами** ряда, n-й член ряда a_n – **общим членом** ряда.

Ряд считается заданным, если задан его общий член a_n.

Ряд (32.1) записывается также в виде $\sum_{n=1}^{\infty} a_n$.

Следует заметить, что из теории действительных чисел известно лишь, что означает сумма конечного числа чисел. Сумма же бесконечного множества слагаемых пока еще не определена.

Рассмотрим суммы конечного числа первых членов ряда

$$S_1 = a_1,$$
$$S_2 = a_1 + a_2,$$
$$S_3 = a_1 + a_2 + a_3,$$
$$\dots\dots\dots\dots\dots\dots$$
$$S_n = a_1 + a_2 + a_3 + ... + a_n,$$

называемые **частичными суммами** ряда (32.1).

Очевидно, частичные суммы образуют числовую последовательность:

$$S_1,\ S_2,\ S_3, ...,\ S_n, ...$$

Определение. Ряд называется **сходящимся**, если существует конечный предел последовательности его частичных сумм, т.е.

$$\lim_{n\to\infty} S_n = S. \tag{32.2}$$

В противном случае ряд (32.1) называется **расходящимся**.

Число S, определяемое равенством (32.2), называется **суммой** ряда.

Пусть ряд (32.1) сходится, S – его сумма. Рассмотрим разность между величиной S и частичной суммой S_n этого ряда: $S - S_n = r_n$. Величина r_n называется **остатком** ряда. Очевидно, что $\lim\limits_{n\to\infty} r_n = \lim\limits_{n\to\infty} (S - S_n) = 0$.

В случае сходимости ряда (32.1) его сумму записывают в виде символического равенства

$$S = a_1 + a_2 + \ldots + a_n + \ldots, \quad \text{или} \quad S = \sum_{n=1}^{\infty} a_n.$$

***Пример* 32.1.** Рассмотрим ряд, составленный из членов бесконечной геометрической прогрессии:

$$b + bq + bq^2 + \ldots + bq^{n-1} + \ldots . \tag{32.3}$$

Частичная сумма S_n этого ряда есть сумма n членов геометрической прогрессии:

$$S_n = b + bq + bq^2 + \ldots + bq^{n-1}.$$

Эта сумма, как известно из школьного курса, при $q \neq 1$ имеет вид

$$S_n = \frac{b(1-q^n)}{1-q} = \frac{b}{1-q} - \frac{bq^n}{1-q}.$$

Если $|q| < 1$, то $\lim\limits_{n\to\infty} q^n = 0$,

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty}\left(\frac{b}{1-q} - \frac{bq^n}{1-q}\right) = \frac{b}{1-q},$$

т.е. ряд (32.3) с х о д и т с я и его сумма $S = \dfrac{b}{1-q}$.

Легко проверить, что при $|q| \geq 1$ ряд (32.3) р а с х о д и т с я.

***Пример* 32.2.** Рассмотрим ряд

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{n(n+1)} + \ldots .$$

Очевидно, $a_n = \dfrac{1}{n(n+1)} = \dfrac{1}{n} - \dfrac{1}{n+1}$. Поэтому

$$S_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n} - \frac{1}{n+1}\right).$$

Если раскрыть скобки, то все слагаемые, кроме первого и последнего, взаимно уничтожатся. Получим

$$S_n = 1 - \frac{1}{n+1}.$$

Отсюда $\lim\limits_{n\to\infty} S_n = 1$. Итак, ряд сходится и его сумма равна единице.

32.2. ОСНОВНЫЕ СВОЙСТВА РЯДОВ

Свойство 1. Если ряд $a_1 + a_2 + \ldots + a_n + \ldots$ сходится и его сумма равна S, то для любого числа λ ряд $\lambda a_1 + \lambda a_2 + \ldots + \lambda a_n + \ldots$ также сходится и его сумма равна λS.

Доказательство. Пусть дан сходящийся ряд

$$a_1 + a_2 + \ldots + a_n + \ldots \tag{32.1}$$

и пусть S – его сумма. Образуем ряд

$$\lambda a_1 + \lambda a_2 + \ldots + \lambda a_n + \ldots . \tag{32.4}$$

Пусть

$$S_n = a_1 + a_2 + \ldots + a_n,$$

$$S'_n = \lambda a_1 + \lambda a_2 + \ldots + \lambda a_n.$$

Тогда, очевидно, $S'_n = \lambda S_n$. Но по условию $\lim\limits_{n\to\infty} S_n = S$. Поэтому $\lim\limits_{n\to\infty} S'_n = \lim\limits_{n\to\infty} \lambda S_n = \lambda \lim\limits_{n\to\infty} S_n = \lambda S$, что и требовалось доказать.

Свойство 2. Если ряды $a_1 + a_2 + \ldots + a_n + \ldots$ и $b_1 + b_2 + \ldots + b_n + \ldots$ сходятся и их суммы соответственно равны S и S', то ряд $(a_1 \pm b_1) + (a_2 \pm b_2) + \ldots + (a_n \pm b_n) + \ldots$ также сходится и его сумма равна $S \pm S'$.

Доказательство. Рассмотрим ряды

$$a_1 + a_2 + \ldots + a_n + \ldots , \tag{32.1}$$

$$b_1 + b_2 + \ldots + b_n + \ldots , \tag{32.5}$$

$$(a_1 + b_1) + (a_2 + b_2) + \ldots + (a_n + b_n) + \ldots . \tag{32.6}$$

Пусть S_n, S'_n, S''_n – частичные суммы рядов (32.1), (32.5), (32.6) соответственно:

$$S_n = a_1 + a_2 + \ldots + a_n,$$

$$S'_n = b_1 + b_2 + \ldots + b_n,$$

$$S''_n = (a_1 + b_1) + (a_2 + b_2) + \ldots + (a_n + b_n).$$

Пусть S – сумма ряда (32.1): $\lim\limits_{n\to\infty} S_n = S$ и пусть S' – сумма ряда (31.5): $\lim\limits_{n\to\infty} S'_n = S'$. Тогда существует предел

$$\lim_{n\to\infty} S''_n = \lim_{n\to\infty} (S_n + S'_n) = \lim_{n\to\infty} S_n + \lim_{n\to\infty} S'_n = S + S',$$

т.е. ряд (32.6) сходится и его сумма равна $S + S'$.

Свойство 3. Отбрасывание конечного числа членов ряда не влияет на его сходимость (расходимость).

Доказательство. В этом случае все частичные суммы ряда, начиная с некоторой, изменятся на одно и то же постоянное число, равное сумме отбрасываемых членов. Поэтому последовательность частичных сумм исходного ряда имеет конечный предел тогда и только тогда, когда существует конечный предел последовательности частичных сумм ряда, полученного из исходного отбрасыванием конечного числа членов.

Отметим отдельно еще одно свойство.

Необходимый признак сходимости ряда

Свойство 4. Общий член a_n сходящегося ряда стремится к нулю при $n \to \infty$.

Доказательство. Пусть ряд сходится и его сумма равна S, т.е. $\lim\limits_{n\to\infty} S_n = S$. Очевидно, что $a_n = S_n - S_{n-1}$. Также очевидно, что $\lim\limits_{n\to\infty} S_{n-1} = S$. Поэтому $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} (S_n - S_{n-1}) = \lim\limits_{n\to\infty} S_n - \lim\limits_{n\to\infty} S_{n-1} =$ $= S - S = 0$.

Подчеркнем, что мы установили только необходимое условие сходимости, не являющееся достаточным. Из того, что $\lim\limits_{n\to\infty} a_n = 0$, еще не следует, что ряд сходится.

Пример **32.3.** Рассмотрим ряд $1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots$, который называется *гармоническим*.

Очевидно, для гармонического ряда выполнено необходимое условие сходимости: $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \frac{1}{n} = 0.$ Докажем, что, несмотря на это, гармонический ряд расходится. Предположим противное, т.е. что ряд сходится и $\lim\limits_{n\to\infty} S_n = S.$ В этом случае, очевидно, и $\lim\limits_{n\to\infty} S_{2n} = S,$ следовательно,

$$\lim_{n\to\infty}(S_{2n} - S_n) = \lim_{n\to\infty} S_{2n} - \lim_{n\to\infty} S_n = S - S = 0. \qquad (*)$$

Однако

$$S_{2n} - S_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} > \frac{1}{2n} + \frac{1}{2n} + \ldots + \frac{1}{2n} = n\frac{1}{2n} = \frac{1}{2},$$

т.е. $S_{2n} - S_n > \frac{1}{2},$ а это противоречит равенству (*). Полученное противоречие означает, что наше предположение о сходимости гармонического ряда неверно.

Отметим, что с помощью необходимого признака сходимости можно доказать расходимость ряда.

Пример **32.4.** Исследовать сходимость ряда

$$\sum_{n=1}^{\infty} \frac{2n-1}{7n+5} = \frac{1}{12} + \frac{3}{19} + \frac{5}{26} + \ldots + \frac{2n-1}{7n+5} + \ldots .$$

Решение: $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \frac{2n-1}{7n+5} = \frac{2}{7} \neq 0,$ поэтому данный ряд расходится.

32.3. РЯДЫ С НЕОТРИЦАТЕЛЬНЫМИ ЧЛЕНАМИ

Ряды с неотрицательными членами (т.е. ряды, все члены которых неотрицательны) – это наиболее простой тип числовых рядов. **Основное свойство рядов с неотрицательными членами**: последовательность частичных сумм такого ряда является неубывающей.

Критерий сходимости

Теорема **32.1.** Для того чтобы ряд с неотрицательными членами сходился, необходимо и достаточно, чтобы последовательность его частичных сумм была ограничена.

Доказательство. 1. *Необходимость.* Пусть ряд сходится. Это означает, что последовательность его частичных сумм сходится. А сходящаяся последовательность, как известно, ограничена.

2. *Достаточность.* Так как последовательность частичных сумм ряда является ограниченной и монотонной, то она сходится в силу теоремы 14.4.

Признаки сравнения

Теорема **32.2 (первый признак сравнения).** Пусть даны два ряда с положительными членами:

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \ldots + a_n + \ldots\ , \tag{32.7}$$

$$\sum_{n=1}^{\infty} b_n = b_1 + b_2 + \ldots + b_n + \ldots\ , \tag{32.8}$$

причем $a_n \le b_n$ для всех n. Тогда из сходимости ряда (32.8) следует сходимость ряда (32.7), а из расходимости ряда (32.7) следует расходимость ряда (32.8).

Доказательство. Пусть S_n – частичная сумма ряда (32.7), S'_n – частичная сумма ряда (32.8). Из условия теоремы следует, что $S_n \le S'_n$. Если ряд (32.8) сходится, то последовательность $\{S'_n\}$ ограничена. Следовательно, последовательность $\{S_n\}$ также ограничена и в силу теоремы 14.4 сходится, т.е. сходится ряд (32.7).

Если же ряд (32.7) расходится, то ряд (32.8) также расходится. Действительно, если бы ряд (32.8) сходился, то, как только что доказано, ряд (32.7) также должен был бы сходиться. Теорема доказана.

Заметим, что в условиях теоремы 32.2 ряд (32.8) называется **мажорантой** ряда (32.7), а ряд (32.7) – **минорантой** ряда (32.8).

Теорему 32.2 можно изложить и в такой, удобной для запоминания форме:

- если мажоранта сходится, то и миноранта сходится;
- если миноранта расходится, то и мажоранта расходится.

Рассмотрим примеры применения теоремы 32.2.

Пример **32.5.** Исследовать сходимость ряда $\sum\limits_{n=1}^{\infty} \frac{1}{2^n + n}$.

Решение. Очевидно, $\frac{1}{2^n + n} < \frac{1}{2^n}$, а ряд $\sum\limits_{n=1}^{\infty} \frac{1}{2^n}$ сходится (сумма бесконечной убывающей геометрической прогрессии). Следовательно, сходится и данный ряд.

Пример **32.6.** Исследовать сходимость ряда $\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}$.

Решение. Так как $\frac{1}{\sqrt{n}} \geq \frac{1}{n}$, а ряд $\sum_{n=1}^{\infty}\frac{1}{n}$ расходится (это гармонический ряд), то данный ряд расходится.

Теорема **32.3 (второй признак сравнения).** Пусть (32.7) – ряд с неотрицательными членами, а (32.8) – с положительными членами и пусть существует отличный от нуля конечный предел

$$\lim_{n\to\infty}\frac{a_n}{b_n} = l.$$

Тогда оба ряда (32.7) и (32.8) сходятся или расходятся одновременно.

Доказательство. В соответствии с определением предела для произвольного $\varepsilon > 0$ существует такое N, что для всех $n > N$ выполняется неравенство

$$l - \varepsilon < \frac{a_n}{b_n} < l + \varepsilon, \quad \text{или} \quad (l - \varepsilon)b_n < a_n < (l + \varepsilon)b_n.$$

Пусть ряд (32.8) сходится. Тогда по свойству 1 (см. § 32.2) сходится ряд $\sum_{n=1}^{\infty}(l+\varepsilon)b_n$, а по теореме 32.2 сходится и ряд (32.7). Аналогично, если сходится ряд (32.7), то по теореме 32.2 сходится ряд $\sum_{n=1}^{\infty}(l-\varepsilon)b_n$ и по тому же свойству 1 сходится ряд (32.8).

Утверждение теоремы о расходимости рядов доказывается аналогично.

Замечание. Можно предполагать, что общие члены рядов (32.7) и (32.8), т.е. a_n и b_n, – бесконечно малые при $n \to \infty$ (в противном случае все было бы ясно само собой: ряд, у которого общий член не стремится к нулю, расходится). Поэтому ***теорему* 32.3** можно переформулировать так: если члены a_n и b_n двух положительных рядов являются бесконечно малыми одного порядка, то эти ряды сходятся или расходятся одновременно.

Пример **32.7.** Исследовать сходимость ряда $\sum_{n=1}^{\infty}\frac{3n-7}{n^2}$.

Решение. Сравним данный ряд с расходящимся гармоническим рядом $\sum_{n=1}^{\infty}\frac{1}{n}$ (см. пример 32.3). Так как

$$\lim_{n\to\infty}\left(\frac{3n-7}{n^2}:\frac{1}{n}\right)=\lim_{n\to\infty}\frac{3n^2-7n}{n^2}=3\neq 0,$$

то данный ряд расходится.

Другие признаки сходимости

Заметим, что оба признака сравнения, рассмотренные выше (см. теоремы 32.2 и 32.3), имеют один и тот же недостаток: чтобы исследовать сходимость какого-нибудь положительного ряда с помощью этих признаков, надо подобрать для сравнения с этим рядом некоторый другой ряд, сходимость (или расходимость) которого известна. Никаких общих методов для нахождения таких рядов нет. Здесь все зависит от интуиции, от того, насколько обширен у исследователя запас «эталонных» рядов, сходимость или расходимость которых известна. Поэтому весьма полезно иметь в распоряжении такие признаки сходимости, для которых не надо привлекать никаких новых рядов, кроме исследуемого.

Теорема **32.4 (признак Даламбера).** Пусть для ряда $\sum_{n=1}^{\infty}a_n$ существует предел

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=l, \tag{32.9}$$

тогда при $l<1$ ряд сходится, а при $l>1$ – расходится.

Доказательство. По определению предела для любого $\varepsilon>0$ существует такой номер N, что для всех $n>N$ выполняются неравенства

$$l-\varepsilon<\frac{a_{n+1}}{a_n}<l+\varepsilon. \tag{32.10}$$

1. Пусть $l<1$. Тогда возьмем такое ε, что $l+\varepsilon<1$. Обозначим $l+\varepsilon=q$. Из неравенств (32.10) имеем

$$\frac{a_{n+1}}{a_n}<q, \quad \text{или} \quad a_{n+1}<a_nq$$

для всех $n>N$. Получаем систему неравенств

$$a_{N+2}<a_{N+1}q, \quad a_{N+3}<a_{N+2}q<a_{N+1}q^2, \ \ldots\ .$$

Итак, члены ряда, начиная с a_{N+2}, меньше соответствующих членов убывающей геометрической прогрессии. Следовательно, ряд сходится.

2. Пусть $l > 1$. Возьмем такое ε, что $l - \varepsilon > 1$. Тогда из левого неравенства (32.10) следует, что $a_{n+1} > a_n$ для всех $n > N$, т.е. члены ряда, начиная с $(N+1)$-го, возрастают, поэтому предел общего члена не равен нулю; следовательно, ряд расходится.

***Пример* 32.8.** Исследовать сходимость ряда $\sum_{n=1}^{\infty} \frac{n^2}{2^n}$.

Решение: $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left(\frac{(n+1)^2}{2^{n+1}} : \frac{n^2}{2^n} \right) = \lim_{n\to\infty} \frac{(n+1)^2}{2n^2} = \frac{1}{2} < 1.$

Следовательно, ряд сходится по признаку Даламбера.

Замечание. При $l = 1$ ряд может как сходиться, так и расходиться. В частности, для рядов $\sum_{n=1}^{\infty} \frac{1}{n}$ и $\sum_{n=1}^{\infty} \frac{1}{n^2}$, как нетрудно убедиться, $l = 1$, но первый из них (гармонический ряд), как известно, расходится, а второй, как мы узнаем позднее, сходится.

***Теорема* 32.5 (признак Коши).** Если для членов ряда $\sum_{n=1}^{\infty} a_n$ существует предел

$$\lim_{n\to\infty} \sqrt[n]{a_n} = l,$$

то ряд сходится при $l < 1$ и расходится при $l > 1$.

Доказательство этой теоремы также основано на том, что при $l < 1$ члены ряда, начиная с некоторого номера, меньше членов некоторой бесконечной убывающей геометрической прогрессии, а при $l > 1$ общий член ряда не стремится к нулю.

1. Пусть $l < 1$. Возьмем какое-нибудь число q, удовлетворяющее соотношению $l < q < 1$.

Так как $\lim_{n\to\infty} \sqrt[n]{a_n} = l$, то, начиная с некоторого номера $n = N$, будет выполняться неравенство

$$\left| \sqrt[n]{a_n} - l \right| < q - l.$$

Отсюда следует, что

$$\sqrt[n]{a_n} < q,$$

или, что то же самое,

$$a_n < q^n \qquad (*)$$

для всех $n \geq N$.

Сравним следующие два ряда:

$$a_1 + a_2 + \ldots + a_N + a_{N+1} + a_{N+2} + \ldots , \tag{32.1′}$$

$$q^N + q^{N+1} + q^{N+2} + \ldots . \tag{**}$$

Члены ряда (**) образуют убывающую геометрическую прогрессию (ее знаменатель равен q, по условию $q < 1$). Следовательно, ряд (**) сходится. Мы знаем, что отбрасывание конечного числа членов ряда не влияет на его сходимость (см. свойство 3 в § 32.2). Из условия (*) следует, что члены ряда (32.1′), начиная с a_N, меньше соответствующих членов сходящегося ряда (**). Следовательно, ряд (32.1′) сходится.

2. Пусть $l > 1$. В этом случае, как легко убедиться, предел общего члена ряда не равен нулю, следовательно, ряд расходится. Теорема доказана.

***Пример* 32.9.** Исследовать сходимость ряда $\sum\limits_{n=1}^{\infty}\left(1-\frac{1}{n}\right)^{n^2}$.

Решение. Применим признак Коши:

$$\lim_{n\to\infty}\sqrt[n]{a_n} = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n = \mathrm{e}^{-1} = \frac{1}{\mathrm{e}} < 1.$$

Следовательно, ряд сходится.

Заметим, что при $l = 1$ признак Коши, так же как и признак Даламбера, не дает ответа на вопрос о сходимости ряда.

Признаки Коши и Даламбера эффективны в основном для выяснения сходимости «быстро» сходящихся рядов – рядов, члены которых являются бесконечно малыми того же (или более высокого) порядка, что и члены убывающих геометрических прогрессий. Мы отмечали уже, что признак Даламбера не дает ответа на вопрос о сходимости или расходимости гармонического ряда

$$1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots ,$$

который, как известно, расходится, а также о сходимости или расходимости обобщенного гармонического ряда

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{n^2} + \ldots ,$$

который, как мы вскоре узнаем, сходится.

***Теорема* 32.6 (интегральный признак сходимости).** Пусть члены ряда $\sum\limits_{n=1}^{\infty} a_n$ не возрастают, т.е. $a_1 \ge a_2 \ge \ldots \ge a_n \ge \ldots$, и пусть

$f(x)$ – такая непрерывная невозрастающая функция, определенная при $x \geq 1$, что

$$f(1) = a_1,\ f(2) = a_2,\ \ldots,\ f(n) = a_n,\ \ldots .$$

Тогда для сходимости ряда $\sum\limits_{n=1}^{\infty} a_n$ необходимо и достаточно, чтобы сходился интеграл

$$\int\limits_1^{\infty} f(x)dx.$$

Доказательство. Так как $f(x)$ монотонна, то для $x \in [n, n+1]$ выполняется неравенство $f(n) \geq f(x) \geq f(n+1)$, или

$$a_n \geq f(x) \geq a_{n+1} \tag{32.11}$$

при любом n.

Интегрируем (32.11) на отрезке $[n, n+1]$:

$$\int\limits_n^{n+1} a_n dx \geq \int\limits_n^{n+1} f(x)dx \geq \int\limits_n^{n+1} a_{n+1}dx.$$

Получаем

$$a_n \geq \int\limits_n^{n+1} f(x)dx \geq a_{n+1}. \tag{32.12}$$

Рассмотрим ряд

$$\int\limits_1^2 f(x)dx + \int\limits_2^3 f(x)dx + \ldots + \int\limits_n^{n+1} f(x)dx + \ldots . \tag{32.13}$$

Его n-я частичная сумма S_n имеет вид

$$S_n = \int\limits_1^2 f(x)dx + \int\limits_2^3 f(x)dx + \ldots + \int\limits_n^{n+1} f(x)dx = \int\limits_1^{n+1} f(x)dx. \tag{32.14}$$

Сходимость ряда (32.13) означает существование конечного предела последовательности его частичных сумм (32.14), т.е. сходимость несобственного интеграла $\int\limits_1^{\infty} f(x)dx$, так как

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \int\limits_1^{n+1} f(x)dx = \int\limits_1^{\infty} f(x)dx.$$

Если ряд сходится, то по теореме 32.2 в силу левого неравенства (32.12) должен сходиться и ряд (32.13), а следовательно, и несобственный интеграл $\int_1^{\infty} f(x)dx.$ И наоборот, если сходится интеграл $\int_1^{\infty} f(x)dx,$ т.е. ряд (32.13), то по той же теореме 32.2 должен сходиться ряд

$$\sum_{n=1}^{\infty} a_{n+1} = a_2 + a_3 + \ldots + a_n + a_{n+1} + \ldots ,$$

а следовательно, и данный ряд $\sum_{n=1}^{\infty} a_n.$

***Пример* 32.10.** Выяснить, при каких $\alpha > 0$ сходится ряд $\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}$ (его называют *обобщенным гармоническим*).

Решение. Рассмотрим функцию $f(x) = \frac{1}{x^{\alpha}},$ $x \geq 1.$ Эта функция монотонно убывает. Поэтому сходимость данного ряда равносильна сходимости несобственного интеграла $\int_1^{\infty} \frac{dx}{x^{\alpha}}.$ Ранее было установлено (см. пример 24.2), что этот интеграл сходится при $\alpha > 1$ и расходится при $0 < \alpha \leq 1.$ Следовательно, данный ряд сходится при $\alpha > 1$ и расходится при $\alpha \leq 1.$

32.4. РЯДЫ С ЧЛЕНАМИ ПРОИЗВОЛЬНОГО ЗНАКА

Перейдем к изучению рядов, содержащих как положительные, так и отрицательные члены.

Определение. Числовой ряд называется **знакопеременным**, если он имеет бесконечное число как положительных, так и отрицательных членов.

Рассмотрим ряд

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \ldots + a_n + \ldots \tag{*}$$

и, кроме того, ряд

$$\sum_{n=1}^{\infty} |a_n| = |a_1| + |a_2| + \ldots + |a_n| + \ldots \tag{**}$$

Можно доказать, что из сходимости ряда (**) следует сходимость ряда (*).

Ряд (*) называется **абсолютно сходящимся**, если сходится ряд (**).

Ряд (*) называется **условно сходящимся**, если он сходится, а ряд (**), составленный из модулей его членов, расходится.

Определение. Числовой ряд $\sum\limits_{n=1}^{\infty} a_n$ называется **знакочередующимся**, если для любого n члены ряда a_n и a_{n+1} имеют разные знаки. Полагая $a_1 > 0$, можно записать знакочередующийся ряд в виде

$$\sum_{n=1}^{\infty}(-1)^{n+1}c_n = c_1 - c_2 + c_3 - c_4 + ... + (-1)^{n+1}c_n + ... , \tag{32.15}$$

где $c_n > 0$.

Сформулируем и докажем **достаточный признак сходимости для знакочередующихся рядов**.

Теорема **32.7 (признак Лейбница).** Если члены знакочередующегося ряда (32.15) убывают по абсолютной величине:

$$c_1 > c_2 > ... > c_n > ...$$

и предел общего члена этого ряда при $n \to \infty$ равен нулю, т.е. $\lim\limits_{n\to\infty} c_n = 0$, то ряд сходится, а его сумма не превосходит первого члена: $S \le c_1$.

Доказательство. Рассмотрим частичную сумму ряда (32.15) с четным числом членов $S_{2m} = c_1 - c_2 + c_3 - c_4 + ... + c_{2m-1} - c_{2m}$. Ее можно представить в виде

$$S_{2m} = (c_1 - c_2) + (c_3 - c_4) + ... + (c_{2m-1} - c_{2m}).$$

В силу условия теоремы все разности в скобках положительны, поэтому последовательность $\{S_{2m}\}$ является возрастающей.

С другой стороны, S_{2m} может быть представлена в виде

$$S_{2m} = c_1 - (c_2 - c_3) - (c_4 - c_5) - ... - (c_{2m-2} - c_{2m-1}) - c_{2m},$$

откуда следует, что $S_{2m} < c_1$.

Итак, последовательность $\{S_{2m}\}$ возрастает и ограничена, следовательно, она имеет предел $\lim\limits_{m\to\infty} S_{2m} = S$. При этом из неравенства $S_{2m} < c_1$ следует, что $S \le c_1$.

Так как $S_{2m+1} = S_{2m} + c_{2m+1}$ и, кроме того, по условию $\lim\limits_{m\to\infty} c_{2m+1} = 0$, то $\lim\limits_{m\to\infty} S_{2m+1} = \lim\limits_{m\to\infty} S_{2m} = S$.

Итак, при любом n (как при $n = 2m$, так и при $n = 2m + 1$) $\lim\limits_{n\to\infty} S_n = S$, т.е. ряд сходится.

***Пример* 32.11.** Исследовать сходимость ряда $\sum\limits_{n=1}^{\infty}\frac{(-1)^n}{n}$.

Решение. В данном случае $c_n = \frac{1}{n}$ и условия признака Лейбница выполнены. Поэтому данный ряд сходится. Однако ряд, составленный из модулей членов данного ряда, есть гармонический ряд $\sum\limits_{n=1}^{\infty}\frac{1}{n}$, который, как известно, расходится. Следовательно, данный ряд сходится условно.

***Замечание* 1.** В теореме Лейбница существенно не только условие $\lim\limits_{n\to\infty} c_n = 0$, но и условие $c_1 > c_2 > \ldots > c_n > \ldots$.

Рассмотрим, например, ряд

$$\frac{1}{\sqrt{2}-1}-\frac{1}{\sqrt{2}+1}+\ldots+\frac{1}{\sqrt{n+1}-1}-\frac{1}{\sqrt{n+1}+1}+\ldots .$$

Члены этого ряда стремятся к нулю, но нарушено условие монотонности. Очевидно,

$$\frac{1}{\sqrt{2}-1}-\frac{1}{\sqrt{2}+1}=2, \quad \frac{1}{\sqrt{3}-1}-\frac{1}{\sqrt{3}+1}=1, \quad \ldots,$$

$$\frac{1}{\sqrt{n+1}-1}-\frac{1}{\sqrt{n+1}+1}=\frac{2}{n}, \quad \ldots .$$

Поэтому данный ряд можно представить в виде

$$2+1+\frac{2}{3}+\ldots+\frac{2}{n}+\ldots, \quad \text{т.е.} \quad \sum_{n=1}^{\infty}\frac{2}{n}.$$

Этот ряд расходится, так как получен из гармонического ряда удвоением всех его членов.

***Замечание* 2.** В процессе доказательства теоремы Лейбница мы видели, что S_{2m} приближается к S возрастая. Сумма S_{2m+1}, напротив, убывает.

Рассмотрим S_{2m+1} подробнее:

$S_1 = c_1$,

$S_3 = c_1 - (c_2 - c_3)$,

$S_5 = c_1 - (c_2 - c_3) - (c_4 - c_5)$,

....

Так как каждая из разностей, написанных в скобках, в соответствии с условием, положительна, то, очевидно,

$$S_1 > S_3 > S_5 > \dots .$$

Таким образом, если ряд удовлетворяет условиям теоремы Лейбница, то суммы S_{2m} являются приближенными значениями суммы S с недостатком, а суммы S_{2m+1} – с избытком: $S_{2m} < S < S_{2m+1}$.

***Замечание* 3.** Сумма S ряда, удовлетворяющая условиям теоремы Лейбница, не превышает по модулю его первого члена и имеет тот же знак, что и первый член. Действительно, $S_1 = c_1$, $S_2 = c_1 - c_2 > 0$. Поэтому из неравенства

$$S_2 < S < S_1$$

получаем

$$0 < S < c_1.$$

Далее заметим, что остаток ряда, удовлетворяющего условиям теоремы Лейбница, сам является рядом, удовлетворяющим этим условиям, и к нему также применимо замечание 3. Если сумма n-го остатка равна r_n, то равенство

$$S = S_n + r_n$$

позволяет сделать следующее ***заключение***: ошибка, допущенная при замене суммы ряда, удовлетворяющего условиям теоремы Лейбница, его частичной суммой, имеет тот же знак, что и первый отброшенный член, а по модулю меньше его.

Вопросы

1. Что такое общий член ряда?
2. Как определяется седьмая частичная сумма ряда S_7?
3. Какой числовой ряд называется сходящимся?
4. Чему равен предел общего члена сходящегося ряда?
5. Что такое остаток ряда?
6. Какой числовой ряд называется гармоническим? Сходится или расходится гармонический ряд?
7. Будет ли сходящимся ряд с положительными членами, для которого предел отношения последующего члена к предыдущему равен 2?
8. Какими свойствами должна обладать функция, применяемая в интегральном признаке сходимости?

9. Предел какого выражения используется в признаке Коши?
10. При каких значениях α обобщенный гармонический ряд $\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}$ является сходящимся?
11. Можно ли с помощью признака Даламбера установить сходимость или расходимость гармонического (обобщенного гармонического) ряда?
12. Будет ли сходящимся знакопеременный ряд, для которого ряд из модулей его членов сходится?
13. Какой ряд называется условно сходящимся?
14. Каких условий достаточно для сходимости знакочередующегося ряда?
15. Пусть ряд удовлетворяет условиям теоремы Лейбница. Как оценить ошибку, допущенную при замене суммы этого ряда его частичной суммой?

Глава 33
ФУНКЦИОНАЛЬНЫЕ РЯДЫ

33.1. ОСНОВНЫЕ ПОНЯТИЯ

Рассмотрим ряд, членами которого являются функции, определенные в некоторой области D:

$$\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots . \tag{33.1}$$

Такой ряд называется **функциональным**.

Придавая x определенные числовые значения, получаем различные числовые ряды, которые могут либо сходиться, либо расходиться.

Совокупность всех значений x, при которых функциональный ряд сходится, называется **областью сходимости** ряда. Очевидно, если D' – область сходимости ряда (33.1), то $D' \subseteq D$. В области сходимости сумма ряда является некоторой функцией от x. Обычно ее обозначают $S(x)$.

Для ряда (33.1) составим частичные суммы $S_n(x)$ – точно так же, как для числовых рядов. Если ряд (33.1) сходится и его сумма равна $S(x)$, то

$$S(x) = S_n(x) + r_n(x), \tag{33.2}$$

где $r_n(x)$ есть сумма ряда $u_{n+1}(x) + u_{n+2}(x) + \ldots$, т.е.

$$r_n(x) = u_{n+1}(x) + u_{n+2}(x) + \ldots . \tag{33.3}$$

В этом случае величина $r_n(x)$ называется **остатком** ряда (33.1).

Так как для каждого x, принадлежащего области сходимости ряда, имеет место равенство $\lim\limits_{n\to\infty} S_n(x) = S(x)$, то с учетом (33.2) получаем

$$\lim_{n\to\infty} r_n(x) = \lim_{n\to\infty} [S(x) - S_n(x)] = 0.$$

Итак, остаток $r_n(x)$ сходящегося ряда стремится к нулю при $n \to \infty$.

Сходимость ряда (33.1) в области D' означает, что для каждого $x \in D'$ последовательность частичных сумм $\{S_n(x)\}$ сходится: $\lim\limits_{n\to\infty} S_n(x) = S_n$. В соответствии с определением предела числовой последовательности это означает, что для всякого $\varepsilon > 0$ существует такой номер N, что для всех $n > N$ выполняется неравенство

$$|S(x) - S_n(x)| < \varepsilon. \tag{33.4}$$

Этот номер N зависит от ε. Действительно, для того же $\varepsilon > 0$, но другого значения x надо подбирать, вообще говоря, другое значение N, чтобы выполнялось неравенство (33.4).

Рассмотрим это на примере. Пусть дан ряд

$$1 + x + x^2 + \ldots + x^n + \ldots .$$

Он, очевидно, сходится и при $x = \frac{1}{5}$, и при $x = \frac{1}{10}$.

Пусть задано $\varepsilon = 0{,}0004$. При $x = \frac{1}{5}$ получаем ряд

$$1 + \frac{1}{5} + \frac{1}{5^2} + \ldots + \frac{1}{5^n} + \ldots .$$

Его сумма (по формуле суммы бесконечно убывающей прогрессии) равна $S = \dfrac{1}{1 - \frac{1}{5}} = \dfrac{5}{4} = 1{,}25.$

При заданном $\varepsilon = 0{,}0004$, как нетрудно подсчитать, надо взять $N = N_1 = 5$ для того, чтобы выполнялось (33.4). Действительно, $S_5 = 1{,}2496$, $S - S_5 = 0{,}0004$, и если $n > 5$, то $|S - S_n| < 0{,}0004$.

При $x = \frac{1}{10}$ ряд выглядит так:

$$1 + \frac{1}{10} + \frac{1}{10^2} + \ldots + \frac{1}{10^n} + \ldots .$$

Его сумма $S = \dfrac{10}{9} = 1{,}11111\ldots$.

Если взять $N = N_2 = 3$, то для $n > N_2$, в частности для $n = 4$, $|S - S_4| = 1{,}11111\ldots - 1{,}111 = 0{,}00011\ldots < 0{,}0004$.

Итак, при одном и том же заданном $\varepsilon = 0{,}0004$ для $x = \frac{1}{5}$ надо, чтобы было $N = 5$, а для $x = \frac{1}{10}$, чтобы $N = 3$. При этом ясно, что если взять $N=5$, то неравенство (33.4) будет выполняться для рассматриваемого ряда и при $x=\frac{1}{5}$, и при $x=\frac{1}{10}$.

Естественно возникает вопрос: всегда ли можно по заданному $\varepsilon > 0$ найти такое N, чтобы для любого $n > N$ и *для всех* $x \in D'$ выполнялось неравенство (33.4)? Ответ на этот вопрос – отрицательный. Существуют функциональные ряды, для которых это невозможно.

Определение. Функциональный ряд (33.1) называется **равномерно сходящимся** в области D', если для всякого $\varepsilon > 0$ существует такой номер N, что для любого $n > N$ и для всех $x \in D'$ выполняется неравенство

$$|S(x) - S_n(x)| < \varepsilon.$$

Подчеркнем, что номер N, упоминаемый в сформулированном определении, зависит только от ε и не зависит от x: $N = N(\varepsilon)$.

Понятие равномерной сходимости весьма глубокое и сложное. Мы здесь не имеем возможности изучать подробно равномерно сходящиеся ряды в общем виде. Рассмотрим важный частный случай равномерно сходящихся рядов – мажорируемые ряды.

Определение. Функциональный ряд

$$u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots$$

называется **мажорируемым** в некоторой области, если существует такой сходящийся числовой ряд

$$c_1 + c_2 + \ldots + c_n + \ldots \tag{33.5}$$

с положительными членами, что для всех значений x из данной области выполняются неравенства

$$|u_1(x)| \le c_1, \quad |u_2(x)| \le c_2, \ \ldots, \ |u_n(x)| \le c_n, \ \ldots . \tag{33.6}$$

(Обратим внимание на то, что ряд называется мажорируемым, если для него существует именно с х о д я щ а я с я числовая мажоранта.)

Например, функциональный ряд

$$\frac{\sin x}{2} + \frac{\sin 2x}{2^2} + \frac{\sin 3x}{2^3} + \ldots + \frac{\sin nx}{2^n} + \ldots$$

есть ряд, мажорируемый на всей числовой прямой, так как для всех значений x выполняются неравенства

$$\left|\frac{\sin nx}{2^n}\right| \le \frac{1}{2^n} \qquad (n = 1, 2, \ldots),$$

а ряд

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots + \frac{1}{2^n} + \ldots ,$$

являясь геометрической прогрессией, сходится.

Из определения следует, что если ряд мажорируем в некоторой области, то он абсолютно сходится в этой области. Кроме того, имеет место следующее важное утверждение.

Теорема **33.1 (признак Вейерштрасса).** Пусть функциональный ряд

$$u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots \tag{33.7}$$

мажорируем на отрезке $[a, b]$. Тогда он сходится равномерно на $[a, b]$.

Доказательство. Пусть ряд (33.5) – сходящаяся числовая мажоранта ряда (33.7). Обозначим через S' сумму ряда (33.5):

$$S' = c_1 + c_2 + \ldots + c_n + c_{n+1} + \ldots , \tag{33.8}$$

через S'_n обозначим n-ю частичную сумму, а через r'_n – остаток ряда (33.5) после n-го члена. Тогда

$$S' = S'_n + r'_n .$$

Так как ряд (33.5) сходится, то

$$\lim_{n\to\infty} S'_n = S'.$$

Следовательно,

$$\lim_{n\to\infty} r'_n = 0.$$

Как уже отмечалось, сумму функционального ряда (33.1) можно представить в виде

$$S(x) = S_n(x) + r_n(x), \tag{33.2}$$

где $S_n(x)$ – n-я частичная сумма, а $r_n(x)$ – остаток ряда (33.1):

$$r_n(x) = u_{n+1}(x) + u_{n+2}(x) + \ldots .$$

Согласно условию (33.6)

$$|u_{n+1}(x)| \le c_{n+1}, \quad |u_{n+2}(x)| \le c_{n+2}, \quad \ldots,$$

следовательно, $|r_n(x)| \le r'_n$ для всех $x \in [a, b]$. Так как $\lim\limits_{n\to\infty} r'_n = 0$ и $\{r'_n(x)\}$ – числовая последовательность, то для всякого $\varepsilon > 0$ существует такое N, не зависящее от x, что для всех $n > N$ выполняется неравенство $|r'_n| < \varepsilon$. Следовательно,

$$|S(x) - S_n(x)| < \varepsilon$$

для всех $n > N$ и всех $x \in [a, b]$. Итак, ряд (33.7) сходится на $[a, b]$ равномерно. Теорема доказана.

33.2. СВОЙСТВА РАВНОМЕРНО СХОДЯЩИХСЯ РЯДОВ

Непрерывность суммы ряда

Рассмотрим ряд

$$u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots, \tag{33.1}$$

все члены которого – непрерывные на $[a, b]$ функции. Хорошо известно, что сумма непрерывных функций есть непрерывная функция, но это верно для конечного числа слагаемых. Всякая частичная сумма ряда (33.1)

$$S_n(x) = u_1(x) + u_2(x) + \ldots + u_n(x)$$

есть функция, непрерывная на $[a, b]$. Естественно возникает вопрос: непрерывна ли сумма ряда (33.1), который состоит из бесконечного числа функций, непрерывных на $[a, b]$? Оказывается, существуют ряды из непрерывных функций, имеющие *разрывную сумму*.

Пример **33.1.** Рассмотрим ряд

$$1 + (x-1) + (x^2 - x) + (x^3 - x^2) + \ldots + (x^n - x^{n-1}) + \ldots . \tag{*}$$

Члены этого ряда – непрерывные функции при всех значениях x.

Убедимся в том, что данный ряд сходится на отрезке $[0, 1]$, но сумма его есть разрывная функция. Действительно, частичная сумма этого ряда $S_n(x)$ имеет вид $S_n(x) = x^{n-1}$. Очевидно,

$$S(x) = \lim_{n\to\infty} S_n(x) = 0 \quad \text{при} \quad 0 \le x < 1,$$

$$S(1) = 1.$$

Итак, сумма $S(x)$ данного ряда имеет разрыв в точке $x = 1$. Данный ряд сходится на $[0, 1]$ неравномерно. Действительно, при $x < 1$ имеем $|S(x) - S_n(x)| = x^{n-1}$. При любом фиксированном n, очевидно, $\lim\limits_{x\to 1} x^{n-1} = 1$. Поэтому если задано $\varepsilon < 1$, то невозможно обеспечить неравенство $|S(x) - S_n(x)| < \varepsilon$ для всех $x \in [0, 1]$ одновременно.

Итак, ряд (*) состоит из непрерывных функций, но имеет сумму, не являющуюся непрерывной функцией, и сходится неравномерно на $[0, 1]$, а потому этот ряд не мажорируем.

***Теорема* 33.2.** Если функции $u_n(x)$ определены и непрерывны на $[a, b]$ и ряд

$$u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots$$

сходится равномерно на $[a, b]$ к сумме $S(x)$, то и эта сумма $S(x)$ также непрерывна на $[a, b]$.

Доказательство. Возьмем на отрезке $[a, b]$ какую-нибудь точку x_0 и установим непрерывность суммы $S(x)$ в этой точке. Так как при любом n и при любом $x \in [a, b]$ выполняется равенство

$$S(x) = S_n(x) + r_n(x),$$

то, в частности,

$$S(x_0) = S_n(x_0) + r_n(x_0).$$

Отсюда

$$|S(x) - S(x_0)| \le |S_n(x) - S_n(x_0)| + |r_n(x) + r_n(x_0)|. \qquad (**)$$

Чтобы доказать непрерывность $S(x)$, надо доказать, что для любого заданного $\varepsilon > 0$ существует такое $\delta > 0$, что для всех x, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство $|S(x) - S(x_0)| < \varepsilon$. Так как данный ряд сходится равномерно, то при заданном $\varepsilon > 0$ существует такой номер N, что для всех $n > N$ выполняется неравенство

$$|r_n(x)| < \frac{\varepsilon}{3} \qquad (33.9)$$

для всех $x \in [a, b]$, в том числе и для $x = x_0$. Это неравенство выполняется, в частности, при $n = N + 1$. Кроме того, функция $S_n(x)$, являясь при $n = N + 1$ суммой конечного числа непрерывных функций, есть функция, непрерывная в точке x_0. Поэтому для заданного $\varepsilon > 0$ существует такое $\delta > 0$, что при $|x - x_0| < \delta$ будет

$$|S_n(x) - S_n(x_0)| < \frac{\varepsilon}{3}. \qquad (33.10)$$

Из (**), (33.9) и (33.10) следует, что для всех x, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство

$$|S(x) - S(x_0)| < \varepsilon.$$

Итак, мы доказали, что в произвольно взятой точке $x_0 \in [a, b]$ сумма $S(x)$ непрерывна. Следовательно, она непрерывна на $[a, b]$. Теорема доказана.

Доказанная теорема 33.2 справедлива (в силу теоремы 33.1) для мажорируемых рядов. Иначе говоря, имеет место следующее утверждение.

***Теорема* 33.3.** Сумма ряда непрерывных функций, мажорируемых на отрезке $[a, b]$, есть функция, непрерывная на этом отрезке.

Почленное интегрирование и дифференцирование рядов

***Теорема* 33.4.** Если функции $u_n(x)$ $(n = 1, 2, ...)$ непрерывны на $[a, b]$ и ряд, составленный из этих функций, сходится равномерно на $[a, b]$ к сумме $S(x)$, то его можно почленно интегрировать в пределах от a до b; при этом интеграл от суммы ряда равен сумме интегралов от членов ряда:

$$\int_a^b S(x)dx = \int_a^b u_1(x)dx + \int_a^b u_2(x)dx + ... + \int_a^b u_n(x)dx + ... \qquad (33.11)$$

Доказательство. Обозначим, как обычно, через $S_n(x)$ n-ю частичную сумму данного ряда (33.11). Его равномерная сходимость означает, что для любого $\varepsilon > 0$ существует такое N, что для всякого $n > N$ и всех $x \in [a, b]$ выполняется неравенство

$$|S(x) - S_n(x)| < \frac{\varepsilon}{b-a}.$$

Сумма $S(x)$ есть непрерывная на $[a, b]$ функция по теореме 33.3. Частичная сумма $S_n(x)$ также непрерывна на $[a, b]$, так как представляет собой сумму конечного числа непрерывных функций. Поэтому функции $S(x)$ и $S_n(x)$ интегрируемы на $[a, b]$ и

$$\left|\int_a^b S(x)dx - \int_a^b S_n(x)dx\right| < \int_a^b |S(x) - S_n(x)|dx < (b-a)\frac{\varepsilon}{b-a} = \varepsilon.$$

Итак,

$$\lim_{n\to\infty}\left[\int_a^b S(x)dx - \left(\int_a^b u_1(x)dx + \int_a^b u_2(x)dx + ... + \int_a^b u_n(x)dx\right)\right] = 0,$$

а это означает, что ряд (33.11) сходится к сумме $\int_a^b S(x)dx$. Теорема доказана.

Замечание. Очевидно, ряд, сходящийся равномерно на $[a, b]$, сходится равномерно на любом отрезке $[a, x]$, где $a < x < b$. Поэтому в условиях теоремы 33.4 выполняется равенство

$$\int_a^x S(t)dt = \int_a^x u_1(t)dt + \int_a^x u_2(t)dt + \ldots + \int_a^x u_n(t)dt + \ldots \,. \qquad (33.12)$$

Теорема **33.5.** Если ряд

$$u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots, \qquad (33.1)$$

составленный из функций, имеющих непрерывные на $[a, b]$ производные, сходится на $[a, b]$ к сумме $S(x)$ и ряд

$$u_1'(x) + u_2'(x) + \ldots + u_n'(x) + \ldots, \qquad (33.13)$$

составленный из производных его членов, сходится равномерно на $[a, b]$, то и сумма $S(x)$ ряда (33.1) имеет на $[a, b]$ производную, причем

$$S'(x) = u_1'(x) + u_2'(x) + \ldots + u_n'(x) + \ldots \,.$$

Доказательство. Обозначим через $\widetilde{S}(x)$ сумму ряда (33.13). Так как этот ряд сходится равномерно на $[a, b]$, то, в силу замечания к теореме 33.4, его можно почленно интегрировать на любом отрезке $[a, x]$, где $a < x < b$:

$$\int_a^x \widetilde{S}(t)dt = \int_a^x u_1'(t)dt + \int_a^x u_2'(t)dt + \ldots + \int_a^x u_n'(t)dt + \ldots \,.$$

Очевидно, для всех n

$$\int_a^x u_n'(t)dt = u_n(x) - u_n(a).$$

Следовательно,

$$\int_a^x \widetilde{S}(t)dt = [u_1(x) - u_1(a)] + [u_2(x) - u_2(a)] + \ldots + [u_n(x) - u_n(a)] + \ldots \,.$$

Но по условию

$$S(x) = u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots,$$

$$S(a) = u_1(a) + u_2(a) + \ldots + u_n(a) + \ldots \,.$$

Отсюда получаем

$$\widetilde{S}(x) = \left(\int_a^x \widetilde{S}(t)dt \right)' = [S(x) - S(a)]' = S'(x).$$

Следовательно,

$$S'(x) = u_1'(x) + u_2'(x) + \ldots + u_n'(x) + \ldots ,$$

что и требовалось доказать.

33.3. СТЕПЕННЫЕ РЯДЫ

О п р е д е л е н и е. Функциональный ряд вида

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots \qquad (33.14)$$

называется **степенным рядом**. Постоянные числа $a_0, a_1, \ldots, a_n, \ldots$ называются **коэффициентами** степенного ряда (33.14).

Область сходимости степенного ряда

Заметим, что область сходимости ряда (33.14) всегда является непустым множеством, поскольку этот ряд сходится при $x = 0$.

Т е о р е м а **33.6 (теорема Абеля).** 1. Если степенной ряд (33.14) с х о д и т с я при некотором $x = x_0$ $(x_0 \neq 0)$, то он сходится, и притом абсолютно, при всех x, удовлетворяющих условию $|x| < |x_0|$. 2. Если ряд (33.14) р а с х о д и т с я при некотором $x = x_1$, то он расходится при всех x, удовлетворяющих условию $|x| > |x_1|$.

Д о к а з а т е л ь с т в о. 1. По условию числовой ряд

$$\sum_{n=0}^{\infty} a_n x_0^n = a_0 + a_1 x_0 + a_2 x_0^2 + \ldots + a_n x_0^n + \ldots$$

с х о д и т с я, поэтому его общий член $a_n x_0^n$ стремится к нулю при $n \to \infty$. Следовательно, последовательность $\{a_n x_0^n\}$ ограничена, т.е. существует такое число $M > 0$, что для всех n выполняется неравенство

$$|a_n x_0^n| < M. \qquad (33.15)$$

Рассмотрим ряд, составленный из абсолютных величин членов ряда (33.14):

$$\sum_{n=0}^{\infty} |a_n x^n| = |a_0| + |a_1 x| + |a_2 x^2| + \ldots + |a_n x^n| + \ldots . \qquad (33.16)$$

Представим его в виде

$$\sum_{n=0}^{\infty} |a_n x^n| = |a_0| + |a_1 x_0| \left|\frac{x}{x_0}\right| + |a_2 x_0^2| \left|\frac{x}{x_0}\right|^2 + \ldots + |a_n x_0^n| \left|\frac{x}{x_0}\right|^n + \ldots . \qquad (33.17)$$

Пусть $x < x_0$. Тогда $q = \left|\frac{x}{x_0}\right| < 1$. Из (33.15) и (33.17) следует, что члены ряда (33.16) меньше соответствующих членов ряда

$$\sum_{n=0}^{\infty} M\left|\frac{x}{x_0}\right|^n = M + M\left|\frac{x}{x_0}\right| + M\left|\frac{x}{x_0}\right|^2 + \ldots + M\left|\frac{x}{x_0}\right|^n + \ldots,$$

который является суммой бесконечной убывающей геометрической прогрессии со знаменателем $q < 1$ и потому сходится. Следовательно, на основании первого признака сравнения (см. теорему 32.2) ряд (33.16) сходится, т.е. ряд (33.14) сходится абсолютно.

2. По условию ряд (33.14) расходится при $x = x_1$. Докажем, что он расходится для всех x, удовлетворяющих условию $|x| > |x_1|$. Предположим противное, т.е. что при некотором x, таком, что $|x| > |x_1|$, ряд (33.14) сходится. Тогда по доказанному первому утверждению теоремы он должен сходиться при $x = x_1$, что противоречит условию. Теорема доказана.

Теорема Абеля позволяет судить о расположении точек сходимости и расходимости степенного ряда. Из нее следует, что если степенной ряд сходится при $x = x_0$, то он сходится абсолютно на интервале $(-|x_0|, |x_0|)$; если же ряд расходится при $x = x_1$, то он расходится всюду вне отрезка $[-|x_1|, |x_1|]$. Отсюда следует, что существует такое число R, что степенной ряд сходится (и притом абсолютно) *в интервале* $(-R, R)$ и расходится *вне отрезка* $[-R, R]$.

Число R называется **радиусом сходимости**, а интервал $(-R, R)$ – **интервалом сходимости** степенного ряда. На концах интервала сходимости (т.е. при $x = -R$ и $x = R$) ряд может как сходиться, так и расходиться.

Укажем **способ нахождения радиуса сходимости степенного ряда**. Рассмотрим ряд (33.14). Для каждого фиксированного x применим к этому ряду признак Даламбера (см. теорему 32.4). Ряд сходится, если

$$\lim_{n\to\infty}\left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| = \lim_{n\to\infty}|x|\left|\frac{a_{n+1}}{a_n}\right| < 1.$$

Пусть $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = L$. Тогда по признаку Даламбера ряд (33.16) сходится, если $|x|L < 1$, и расходится, если $|x|L > 1$. Следовательно, ряд (33.14) сходится абсолютно при $x<\dfrac{1}{L}$ и расходится при $x>\dfrac{1}{L}$. Таким образом, число $\dfrac{1}{L}$ является *радиусом сходимости* ряда (33.14), т.е.

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|. \qquad (33.18)$$

При этом, в частности, может быть $R=0$ или $R=\infty$, т.е. область сходимости может состоять из одной точки или совпадать со всей числовой прямой.

***Пример* 33.2.** Найти область сходимости степенного ряда $\sum\limits_{n=1}^{\infty}\dfrac{x^n}{n}$.

Решение. Здесь $a_n=\dfrac{1}{n}$, $a_{n+1}=\dfrac{1}{n+1}$. Найдем радиус сходимости по формуле (33.18):

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|=\lim_{n\to\infty}\frac{n+1}{n}=1.$$

Интервал сходимости есть $(-1, 1)$.

Выясним теперь поведение ряда на концах интервала сходимости: а) при $x=-1$ получаем знакочередующийся ряд $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^n}{n}$, который сходится по признаку Лейбница (см. теорему 32.7); б) при $x=1$ получаем гармонический ряд $\sum\limits_{n=1}^{\infty}\dfrac{1}{n}$, который, как известно, расходится.

Итак, на левом конце интервала сходимости ряд сходится (условно), на правом конце – расходится.

***Пример* 33.3.** Найти область сходимости степенного ряда $\sum\limits_{n=0}^{\infty}\dfrac{x^n}{n!}$. (Напомним, что $0!=1$.)

Решение. Найдем радиус сходимости по формуле (33.18):

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|=\lim_{n\to\infty}\frac{(n+1)!}{n!}=\lim_{n\to\infty}(n+1)=\infty.$$

Данный ряд сходится (абсолютно) на всей числовой прямой.

Свойства степенных рядов

Пусть функция $f(x)$ является суммой степенного ряда:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \tag{33.19}$$

интервал сходимости которого $(-R, R)$. В этом случае говорят, что *функция* $f(x)$ *разлагается на интервале* $(-R, R)$ *в степенной ряд*.

Из теоремы Абеля следует, что степенной ряд мажорируем на любом отрезке, целиком лежащем внутри интервала сходимости. Поэтому степенные ряды обладают рядом свойств, аналогичных свойствам обыкновенных многочленов.

Доказательства **теорем о свойствах степенных рядов** основаны на равномерной сходимости степенного ряда на любом отрезке, содержащемся в интервале сходимости.

***Теорема* 33.7.** Степенной ряд

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots \tag{33.14}$$

мажорируем на любом отрезке $[-r, r]$, целиком лежащем внутри его интервала сходимости.

Доказательство. Рассмотрим числовой ряд

$$|a_0| + |a_1| r + |a_2| r^2 + \ldots + |a_n| r^n + \ldots\,. \tag{33.20}$$

Надо доказать, что при $r < R$, где R – радиус сходимости ряда, существует для ряда (33.20) сходящаяся числовая мажоранта (с положительными членами).

Так как $r < R$, то по теореме Абеля этот ряд сходится. Но при $x \in [-r, r]$, т.е. при $|x| \le r$, для членов ряда (33.14) выполняются неравенства $|a_n x^n| \le |a_n| r^n$. Следовательно, ряд (33.14) мажорируем на отрезке $[-r, r]$, что и требовалось доказать.

Следствие. Степенной ряд сходится равномерно на всяком отрезке, целиком лежащем внутри интервала сходимости.

Действительно, так как ряд мажорируем на $[-r, r]$ при $r < R$, то по признаку Вейерштрасса (см. теорему 33.1) он сходится равномерно на $[-r, r]$.

***Теорема* 33.8.** На всяком отрезке, целиком лежащем внутри интервала сходимости, сумма степенного ряда есть непрерывная функция.

Эта теорема есть непосредственное следствие теорем 33.1, 33.2 и 33.7.

Теорема **33.9.** Степенной ряд можно почленно интегрировать на любом отрезке $[0, x]$, где $-R < x < R$. При этом интеграл от суммы ряда равен сумме интегралов от членов ряда:

$$\int_0^x \left(\sum_{n=1}^{\infty} a_n x^n \right) dx = a_0 x + \frac{a_1}{2} x^2 + \frac{a_2}{3} x^3 + \ldots + \frac{a_n}{n+1} x^{n+1} + \ldots .$$

Эта теорема также есть непосредственное следствие предыдущих утверждений – теорем 33.4 и 33.6.

Выясним теперь возможность **почленного дифференцирования степенного ряда**.

Теорема **33.10.** Пусть функция $f(x)$ разлагается на интервале $(-R, R)$ в степенной ряд

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots . \tag{33.19}$$

Тогда степенной ряд

$$\sum_{n=1}^{\infty} n a_n x^{n-1} = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1} + \ldots , \tag{33.21}$$

полученный почленным дифференцированием ряда (33.19), имеет тот же интервал сходимости $(-R, R)$ и функция $f(x)$ на всем интервале $(-R, R)$ имеет производную $f'(x)$, которая разлагается в степенной ряд (33.21):

$$f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1} + \ldots . \tag{33.22}$$

Доказательство. Докажем, что ряд (33.21) мажорируем на любом отрезке $[-r, r]$ при $r < R$. Возьмем произвольно точку r_0, удовлетворяющую условиям $r < r_0 < R$. В этой точке ряд (33.19) сходится, следовательно, общий член этого ряда при $x = r_0$ стремится к нулю, а потому ограничен. Иначе говоря, $\lim\limits_{n\to\infty} a_n r_0^n = 0$, поэтому существует такое $M > 0$, что

$$|a_n| r_0^n < M \quad (n = 1, 2, \ldots).$$

Тогда при $x \in [-r, r]$ имеем

$$|n a_n x^{n-1}| \le |n a_n r^{n-1}| = n|a_n| r_0^{n-1} \left(\frac{r}{r_0} \right)^{n-1} \le n \frac{M}{r_0} \left(\frac{r}{r_0} \right)^{n-1} .$$

Обозначим $\dfrac{M}{r_0} = M_0$, $\dfrac{r}{r_0} = q$; по условию $q < 1$.

Итак, $|n a_n x^{n-1}| \le n M_0 q^{n-1}$. Следовательно, все члены ряда (33.21) для указанных значений x не превосходят соответствующих членов мажорантного числового ряда

$$M_0 + 2M_0 q + 3M_0 q^2 + \ldots + n M_0 q^{n-1} + \ldots .$$

А этот последний ряд сходится, в чем нетрудно убедиться с помощью признака Даламбера. Обозначим $a'_n = n M_0 q^{n-1}$, $a'_{n+1} = (n+1) M_0 q^n$. Тогда

$$\lim_{n\to\infty} \frac{a'_{n+1}}{a'_n} = \lim_{n\to\infty} \frac{(n+1)q^n}{nq^{n-1}} = q < 1.$$

Итак, ряд (33.21) мажорируем на отрезке $[-r, r]$, и на основании теоремы 33.5 можно утверждать, что ряд (33.19) можно почленно дифференцировать и верно равенство (33.22).

Так как для всякой точки $x \in (-R, R)$ существует такое $r < R$, что $x \in [-r, r]$, то из доказанного следует, что ряд (33.21) сходится в любой внутренней точке интервала $(-R, R)$.

Мы доказали, что при дифференцировании степенного ряда его радиус сходимости *не может уменьшиться*.

Для завершения доказательства теоремы надо доказать теперь, что радиус сходимости *не может увеличиться* в результате дифференцирования. Предположим противное, т.е. что ряд (33.21) сходится при некотором $x_1 > R$. Интегрируя этот ряд в пределах от 0 до x_2, где $R < x_2 < x_1$, мы получили бы, что исходный ряд (33.19) сходится в точке x_2, а это противоречит условию теоремы. Итак, интервал сходимости ряда (33.19), т.е. $(-R, R)$, есть в то же время интервал сходимости ряда (33.21), полученного почленным дифференцированием ряда (33.19). Теорема доказана.

Теоремы 33.9 и 33.10 означают, что по отношению к дифференцированию и интегрированию степенные ряды (в пределах их интервала сходимости) ведут себя как обыкновенные многочлены.

Заметим, что, применяя теорему 33.10 повторно, легко убедиться, что функция, которая разлагается в степенной ряд, бесконечно дифференцируема на интервале сходимости этого ряда.

Степенные ряды с произвольным центром

Степенным рядом называется также функциональный ряд вида

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \ldots + a_n(x-x_0)^n + \ldots . \quad (33.23)$$

Это степенной ряд по степеням двучлена $x - x_0$. Очевидно, степенной ряд (33.14) является частным случаем ряда (33.23).

Для определения области сходимости ряда (33.23) сделаем замену:

$$x - x_0 = X.$$

После этой замены ряд (33.23) примет вид

$$a_0 + a_1 X + a_2 X^2 + \ldots + a_n X^n + \ldots . \quad (33.24)$$

Пусть $(-R,\ R)$ – интервал сходимости ряда (33.24). Тогда ряд (33.23) сходится при $|x - x_0| < R$ и расходится при $|x - x_0| > R$. Поэтому интервалом сходимости ряда (33.23) будет множество всех значений x, удовлетворяющих неравенству $-R < x - x_0 < R$, т.е. $x_0 - R < x < x_0 + R$.

Следовательно, *интервалом сходимости* ряда (33.23) будет интервал $(x_0 - R, x_0 + R)$ с центром в точке x_0. Все свойства степенного ряда (33.14) полностью сохраняются для степенного ряда (33.23).

Вопросы

1. Что такое область сходимости функционального ряда?
2. Какой функциональный ряд называется равномерно сходящимся?
3. Какой функциональный ряд называется мажорируемым? Всякий ли мажорируемый ряд является равномерно сходящимся?
4. Всегда ли сумма функционального ряда, состоящего из непрерывных функций, есть непрерывная функция?
5. Какой функциональный ряд называется степенным рядом?
6. Что такое радиус сходимости степенного ряда? Может ли радиус сходимости быть равным нулю или бесконечности?
7. Каковы основные свойства степенного ряда?

Глава 34
РЯДЫ ТЕЙЛОРА И МАКЛОРЕНА

34.1. РАЗЛОЖЕНИЕ ФУНКЦИИ В СТЕПЕННОЙ РЯД

Пусть функция $f(x)$ на некотором интервале $(-R, R)$ может быть разложена в степенной ряд

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots . \tag{34.1}$$

Из теоремы 33.4 следует, что этот ряд можно дифференцировать почленно любое число раз. Тогда, дифференцируя n раз равенство (34.1), получаем

$$f'(x) = a_1 + 2a_2 x + \dots + n a_n x^{n-1} + \dots ,$$

$$f''(x) = 2a_2 + \dots + n(n-1)a_n x^{n-2} + \dots ,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$f^{(n)}(x) = n! a_n + (n+1)n \cdots 3 \cdot 2 \cdot a_{n+1} x + \dots .$$

Полагая в этих равенствах, а также в разложении (34.1) $x = 0$, имеем $f(0) = a_0$, $f'(0) = a_1$, $f''(0) = 2!a_2$, ..., $f^{(n)}(0) = n!a_n$. Отсюда

$$a_n = \frac{f^{(n)}(0)}{n!} \qquad (n = 0, 1, 2, \dots). \tag{34.2}$$

Подставив коэффициенты (34.2) в (34.1), получаем разложение функции $f(x)$ в степенной ряд:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \dots . \tag{34.3}$$

Ряд (34.3) называется **рядом Маклорена** для функции $f(x)$.

Мы доказали следующую теорему.

Теорема **34.1.** Если функция $f(x)$ может быть разложена на интервале $(-R, R)$ в степенной ряд, то этот ряд является ее рядом Маклорена.

Теорема 34.1 означает, что разложение функции в степенной ряд единственно. Коэффициенты этого разложения однозначно определяются формулами (34.2).

Известно, что для любой функции, имеющей производные до $(n+1)$-го порядка включительно, справедлива формула Маклорена [см. формулу (17.18)]:

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+\frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1},$$

где ξ – точка, расположенная между 0 и x $(\xi=\theta x,\ 0<\theta<1)$.

Если обозначить в этой формуле остаточный член через $R_n(x)$:

$$R_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1},$$

а частичную сумму ряда Маклорена через $S_n(x)$, то формулу Маклорена можно записать в следующем виде:

$$f(x)=S_n(x)+R_n(x). \tag{34.4}$$

Из равенства (34.4) легко следует **критерий разложимости функции в ряд Маклорена**.

***Теорема* 34.2.** Для того чтобы бесконечно дифференцируемая функция $f(x)$ разлагалась в ряд Маклорена на интервале $(-R, R)$, необходимо и достаточно, чтобы остаточный член формулы Маклорена для этой функции стремился к нулю на указанном интервале при $n\to\infty$.

Доказательство. 1. *Необходимость*. Пусть $\lim\limits_{n\to\infty}S_n(x)=f(x)$ для всех $x\in(-R, R)$. Тогда для всех $x\in(-R, R)$ в силу (34.4)

$$\lim_{n\to\infty}R_n(x)=\lim_{n\to\infty}[f(x)-S_n(x)]=f(x)-f(x)=0.$$

2. *Достаточность*. Пусть теперь дано, что $\lim\limits_{n\to\infty}R_n(x)=0$ для всех $x\in(-R, R)$. Тогда из (34.4) получаем

$$\lim_{n\to\infty}S_n(x)=\lim_{n\to\infty}[f(x)-R_n(x)]=f(x)-0=f(x),$$

т.е. ряд Маклорена сходится к функции $f(x)$. Доказательство закончено.

Заметим, что если $\lim\limits_{n\to\infty}R_n(x)\neq 0$, то ряд Маклорена не представляет данной функции, хотя и может сходиться (к некоторой другой функции).

Как известно, формула Маклорена является частным случаем формулы Тейлора [см. формулы (17.7) и (17.12)]:

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+R_n(x),$$

где $R_n(x)=\dfrac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$.

Если функция $f(x)$ имеет производные любого порядка в окрестности точки $x = a$, то получим бесконечный ряд, который называется **рядом Тейлора**:

$$f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\ldots .$$

Ряд Маклорена является частным случаем ряда Тейлора при $a = 0$.

Ряд Тейлора для функции $f(x)$ сходится к этой функции тогда и только тогда, когда остаточный член в формуле Тейлора для этой функции стремится к нулю: $\lim\limits_{n\to\infty} R_n(x)=0$.

Легко убедиться в том, что если все производные ограничены: $f^{(n)}(x) \le M$ $(n = 1, 2, \ldots)$, то $\lim\limits_{n\to\infty} R_n(x)=0$.

34.2. РАЗЛОЖЕНИЕ НЕКОТОРЫХ ЭЛЕМЕНТАРНЫХ ФУНКЦИЙ В РЯД МАКЛОРЕНА

1. Пусть $\boldsymbol{f(x) = \mathrm{e}^x}$. По формуле Тейлора

$$\mathrm{e}^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+R_n(x),$$

где $R_n(x)=\dfrac{x^{n+1}}{(n+1)!}\mathrm{e}^{\theta}$, $0<\theta<1$.

Так как при любом фиксированном x величина e^x ограничена, то $\lim\limits_{n\to\infty} R_n(x)=0$. Следовательно, при всех $x\in(-\infty, \infty)$

$$\mathrm{e}^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+\ldots . \tag{34.5}$$

2. Аналогично получаем разложение в ряд Маклорена функций $\boldsymbol{f(x) = \sin x}$ и $\boldsymbol{f(x) = \cos x}$ (см. § 17.3; здесь также $\lim\limits_{n\to\infty} R_n(x)=0$ при всех x):

$$\sin x=x-\frac{x^3}{3!}+\ldots+(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\ldots , \tag{34.6}$$

$$\cos x=1-\frac{x^2}{2!}+\ldots+(-1)^n\frac{x^{2n}}{(2n)!}+\ldots . \tag{34.7}$$

3. Рассмотрим функцию $\boldsymbol{f(x)=(1+x)^m}$, где m – произвольное постоянное число. Имеем

$$f(x)=(1+x)^m,\ f'(x)=m(1+x)^{m-1},\ f''(x)=m(m-1)(1+x)^{m-2},\ \dots,\ f^{(n)}(x)=m(m-1)\cdots(m-n+1)(1+x)^{m-n}.$$

При $x=0$:

$$f(0)=1,\quad f'(0)=m,\quad f''(0)=m(m-1),\quad \dots,$$

$$f^{(n)}(0)=m(m-1)\cdots(m-n+1).$$

Получаем ряд, который называется **биномиальным**:

$$1+\frac{m}{1!}x+\frac{m(m-1)}{2!}x^2+\dots+\frac{m(m-1)\cdots(m-n+1)}{n!}x^n+\dots .$$

Определим радиус сходимости этого ряда $R=\lim\limits_{n\to\infty}\left|\dfrac{a_n}{a_{n+1}}\right|$. Так как

$$a_n=\frac{m(m-1)\cdots(m-n+1)}{n!},\quad a_{n+1}=\frac{m(m-1)\cdots(m-n+1)(m-n)}{(n+1)!},$$

то $R=\lim\limits_{n\to\infty}\left|\dfrac{n+1}{m-n}\right|=|-1|=1.$

Таким образом, биномиальный ряд сходится при $x\in(-1,1)$ и расходится вне отрезка $[-1,1]$.

Оценка остаточного члена этого ряда связана с определенными сложностями, поэтому примем без доказательства, что $\lim\limits_{n\to\infty}R_n(x)=0$ при $x\in(-1,1)$.

Итак, при $x\in(-1,1)$

$$(1+x)^m=1+\frac{m}{1!}x+\frac{m(m-1)}{2!}x^2+\dots+\frac{m(m-1)\cdots(m-n+1)}{n!}x^n+\dots . \quad (34.8)$$

Заметим, что если m – целое положительное число, то начиная с члена, содержащего x^{m+1}, все коэффициенты равны нулю и ряд превращается в конечный многочлен.

4. При $m=-1$ биномиальный ряд имеет вид

$$\frac{1}{1+x}=1-x+x^2-x^3+\dots+(-1)^n x^n+\dots . \quad (34.9)$$

Интегрируем это равенство от 0 до x, где $|x|<1$:

$$\int_0^x\frac{dt}{1+t}=\int_0^x(1-t+t^2-t^3+\dots+(-1)^n t^n+\dots)dt.$$

Отсюда получаем разложение функции $f(x) = \ln(1 + x)$:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + \frac{(-1)^{n+1}x^n}{n} + \ldots . \qquad (34.10)$$

Это равенство справедливо в интервале $(-1, 1)$. Можно показать, что оно справедливо и при $x = 1$. Таким образом, область сходимости ряда (34.10) есть $(-1, 1]$.

34.3. ПРИЛОЖЕНИЕ СТЕПЕННЫХ РЯДОВ К ПРИБЛИЖЕННЫМ ВЫЧИСЛЕНИЯМ

Используя разложения в ряд Маклорена элементарных функций, можно получить значения этих функций с любой точностью. Для этого надо взять достаточное число членов ряда. Точность вычисления определяется остаточным членом формулы Маклорена.

1. Рассмотрим **разложение показательной функции $f(x) = e^x$.**

Как уже отмечалось,

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + R_n(x),$$

где $R_n(x) = \frac{x^{n+1}}{(n+1)!}e^{\theta}$, и при всех x показательная функция разлагается в ряд

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots . \qquad (34.5)$$

Радиус сходимости этого ряда бесконечен, т.е. ряд сходится на всей числовой прямой.

Ряд (34.5) пригоден для вычислений при небольших значениях x. Если же x велико по абсолютной величине, то его представляют в виде суммы целой и дробной частей:

$$x = E(x) + q,$$

где $E(x)$ – целая часть x (т.е. наибольшее целое число, не превосходящее x), а q – дробная его часть, $0 \le q < 1$. Тогда

$$e^x = e^{E(x)}e^q.$$

Первый множитель $e^{E(x)}$, являющийся целой степенью числа e, находят с помощью умножения, второй, т.е. e^q, – с помощью разложения (34.5).

Остаток ряда оценивается следующим образом:

$$0 < R_n(x) < \frac{x^{n+1}}{n!n}.$$

Пример **34.1.** Найти $\sqrt{e}$ с точностью до 10^{-6}.

Решение. Так как члены разложения (34.5) имеют вид $u_0 = 1$, ..., $u_{k-1} = \dfrac{1}{2^{k-1}(k-1)!}$, $u_k = \dfrac{1}{2^k k!}$ $(k = 1, 2, ..., n)$, то $u_k = \dfrac{u_{k-1}}{2k}$. Имеем

$$e^{\frac{1}{2}} = \sum_{k=0}^{n} u_k + R_n\left(\frac{1}{2}\right).$$

Подсчитаем слагаемые с двумя запасными знаками:

$$u_0 = 1, \qquad u_5 = \frac{u_4}{10} = 0{,}00026042,$$

$$u_1 = \frac{u_0}{2} = 0{,}50000000, \qquad u_6 = \frac{u_5}{12} = 0{,}00002170,$$

$$u_2 = \frac{u_1}{4} = 0{,}12500000, \qquad u_7 = \frac{u_6}{14} = 0{,}00000155,$$

$$u_3 = \frac{u_2}{6} = 0{,}02083333, \qquad u_8 = \frac{u_7}{16} = 0{,}00000010,$$

$$u_4 = \frac{u_3}{8} = 0{,}00260417, \qquad S_8 = 1{,}64872117.$$

Округляя сумму до шести десятичных знаков после запятой, получаем

$$\sqrt{e} = 1{,}648721.$$

2. Перейдем к **разложению логарифма** и **вычислению значений логарифмической функции**. Логарифмическая функция $\boldsymbol{f(x) = \ln(1 + x)}$ разлагается на $(-1, 1]$ в степенной ряд:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^n \frac{x^{n+1}}{n+1} + \ldots . \qquad (34.11)$$

Непосредственное применение этого ряда затруднительно, в частности, из-за его медленной сходимости. Заменим в (34.11) аргумент x на $-x$:

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \ldots - \frac{x^{n+1}}{n+1} - \ldots . \qquad (34.12)$$

Вычтем (34.12) из (34.11):

$$\ln\frac{1+x}{1-x}=2\left(x+\frac{x^3}{3}+\frac{x^5}{5}+\ldots\right), \tag{34.13}$$

где $|x|<1$.

Ряд (34.13) сходится довольно быстро – быстрее геометрической прогрессии со знаменателем $q=x^2$. Кроме того, выражение $\dfrac{1+x}{1-x}$ при указанных значениях x может принимать любые положительные значения. Поэтому формула (34.13) очень удобна для вычисления логарифмов.

Так как члены ряда (34.13) меньше членов геометрической прогрессии $2(x+x^3+x^5+\ldots)$, то остаток ряда оценивается следующим образом:

$$R_n(x)<\frac{2x^{2n+1}}{(2n+1)(1-x^2)}. \tag{34.14}$$

Пример **34.2.** Вычислить $\ln 8$ с точностью до 10^{-6}.

Решение. Из равенства $\dfrac{1+x}{1-x}=8$ можем найти $x=0{,}777\ldots$, но для ускорения сходимости ряда лучше представить $8=\mathrm{e}^2\dfrac{8}{\mathrm{e}^2}$. Тогда $\ln 8=\ln\mathrm{e}^2+\ln\dfrac{8}{\mathrm{e}^2}=2+\ln\dfrac{8}{\mathrm{e}^2}$. Полагая $\dfrac{1+x}{1-x}=\dfrac{8}{\mathrm{e}^2}$, находим $x=\dfrac{8-\mathrm{e}^2}{8+\mathrm{e}^2}=0{,}03969989$. Из (34.14) следует, что при таком x остаток приблизительно равен первому из отброшенных членов. Как и в предыдущем примере, будем проводить вычисления с двумя запасными знаками:

$$u_1=2x=0{,}07939978,$$

$$u_2=2\cdot\frac{x^3}{3}=0{,}00004171,$$

$$u_3=2\cdot\frac{x^5}{5}=0{,}00000004.$$

Получаем $\ln 8=2{,}079441$.

(Аналогичным образом вычисляют значения функций $\sin x$ и $\cos x$, используя их разложения в ряд Маклорена.)

3. Рассмотрим **вычисление значений «неберущихся» интегралов** с помощью рядов Маклорена. Известно, в частности, что первообразная от e^{-x^2} не является элементарной функцией.

Пример **34.3.** Вычислить интеграл

$$\int_0^a e^{-x^2}\,dx.$$

Решение. Разложим подынтегральную функцию в ряд, заменяя x в разложении показательной функции (34.5) на $-x^2$:

$$e^{-x^2}=1-\frac{x^2}{1!}+\frac{x^4}{2!}-\frac{x^6}{3!}+\ldots+(-1)^n\frac{x^{2n}}{n!}+\ldots .$$

Интегрируя обе части этого неравенства, получим

$$\int_0^a e^{-x^2}\,dx=\left.\left(\frac{x}{1}-\frac{x^3}{1!\cdot 3}+\frac{x^5}{2!\cdot 5}-\frac{x^7}{3!\cdot 7}+\ldots\right)\right|_0^a=$$

$$=\frac{a}{1}-\frac{a^3}{1!\cdot 3}+\frac{a^5}{2!\cdot 5}-\frac{a^7}{3!\cdot 7}+\ldots .$$

С помощью этого равенства можно при любом a вычислить данный интеграл с любой степенью точности. В частности, для вычисления интеграла $\int_0^1 e^{-x^2}\,dx$ с точностью до 10^{-4} достаточно взять семь членов разложения. Получаем

$$\int_0^1 e^{-x^2}\,dx\approx 0{,}7468.$$

Заметим, что здесь остаток ряда не превышает по модулю первого из отброшенных членов, так как ряд знакочередующийся.

Вопросы

1. Что такое ряд Маклорена? Для какой функции его можно определить?
2. Обязательно ли ряд Маклорена функции $f(x)$ сходится к этой функции?

3. В чем заключается критерий разложимости функции в ряд Маклорена?
4. Что такое ряд Тейлора?
5. Что такое биномиальный ряд? Какова его область сходимости?

Список литературы

1. *Архипов Г.И., Садовничий В.А., Чубариков В.Н.* Лекции по математическому анализу. – М.: Высшая школа, 2000.
2. *Беклемишев Д.В.* Курс аналитической геометрии и линейной алгебры: Учебник. – М.: Высшая школа, 1998.
3. *Бугров Я.С., Никольский С.М.* Элементы линейной алгебры и аналитической геометрии: Учебник для вузов. – Ростов-на-Дону: Феникс, 1997.
4. Высшая математика для экономистов: Учебник / Н.Ш. Кремер, Б.А. Путко, И.М. Тришин, М.Н. Фридман; Под ред. проф. Н.Ш. Кремера. – 2-е изд. – М.: Банки и биржи; ЮНИТИ, 2004.
5. *Замков О.О., Черемных Ю.Н., Толстопятенко А.В.* Математические методы в экономике: Учебник. – М.: Дело и Сервис, 1999.
6. *Красс М.С.* Математика для экономических специальностей: Учебник. – М.: ИНФРА-М, 1999.
7. *Малыхин В.И.* Математика в экономике: Учеб. пособие. – М.: ИНФРА-М, 2001.
8. Общий курс высшей математики для экономистов: Учебник / Под ред. В.И. Ермакова. – М.: ИНФРА-М, 2003.
9. Сборник задач по высшей математике для экономистов: Учеб. пособие / Под ред. В.И. Ермакова. – М.: ИНФРА-М, 2005.
10. *Солодовников А.С., Бабайцев В.А., Браилов А.В., Шандра И.Г.* Математика в экономике: Учебник: В 2 ч. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2003.
11. Справочник по математике для экономистов / Под ред. В.И. Ермакова. – М.: Высшая школа, 1997.
12. *Фихтенгольц Г.М.* Основы математического анализа. – СПб.: Лань, 2001.
13. *Шикин Е.В., Чхартишвили А.Г.* Математические методы и модели в управлении: Учеб. пособие. – М.: Дело, 2000.

14. *Anthony Martin, Biggs Norman.* Mathematics for Economics and Finance. Methods and Modeling: 2nd ed. – UK: Cambridge University Press, 1998.

15. *Fuente de la Angel.* Mathematics Methods and Models for Economics. – UK: Cambridge University Press, 2000.

ОГЛАВЛЕНИЕ

Раздел III. ЛИНЕЙНЫЕ МОДЕЛИ В ЭКОНОМИКЕ

Раздел IV. ВВЕДЕНИЕ В АНАЛИЗ

Раздел VI. ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ

Раздел IX. РЯДЫ

Учебное издание

Владимир Леонидович Клюшин

ВЫСШАЯ МАТЕМАТИКА ДЛЯ ЭКОНОМИСТОВ

Учебное пособие

Редактор *И.В. Мартынова*
Корректор *М.В. Литвинова*
Компьютерная верстка *О.В. Савостиной*, *Г.А. Волковой*

Оригинал-макет подготовлен
в Издательском Доме «ИНФРА-М»

ЛР № 070824 от 21.01.93

Сдано в набор 11.01.2006. Подписано в печать 20.04.2006.
Формат 60×90/16. Печать офсетная. Бумага офсетная.
Гарнитура «Newton». Усл. печ. л. 28,0. Уч.-изд. л. 29,56.
Тираж 3000 экз. Заказ №

Издательский Дом «ИНФРА-М»
127282, Москва, Полярная ул., д. 31в.
Тел.: (495) 380-05-40, 380-05-43. Факс: (495) 363-92-12.
E-mail: books@infra-m.ru; http://www.infra-m.ru

Отдел «Книга — почтой»:
(495) 363-42-60 (доб. 246, 247)