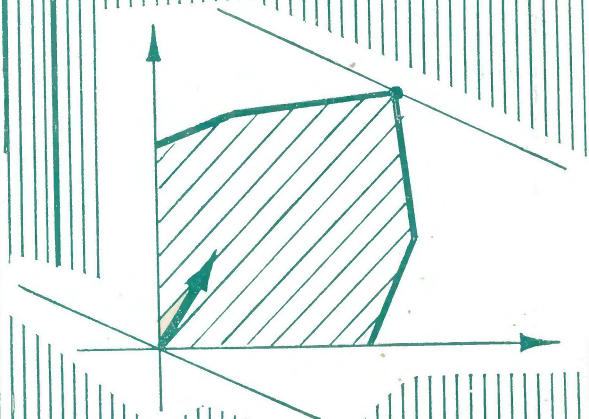
СПРАВОЧНИК по математике для экономистов

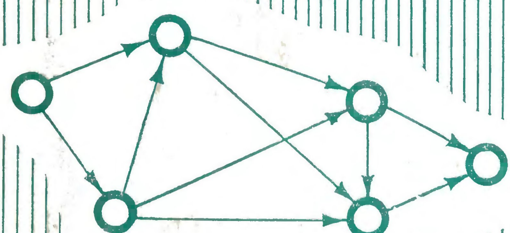
$$\mathrm{grad}\, F(\bar{x}) = \left\{ \frac{\partial F}{\partial x_1} \right.$$



$$x_k - \sum_{i=1}^{n} a_{ik} x_i = y_n \qquad k = 1, 2, \ldots, n$$

$$\frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n}\Big\}$$

$$F=\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$



$$1(X)=\int_{-\infty}^{\infty} x f(x)\,dx$$

СПРАВОЧНИК по математике для экономистов

Под редакцией проф. В. И. Ермакова

Москва «Высшая школа» 1987

ББК 22.1
С 74
УДК 51

Рекомендовано Министерством высшего и среднего специального образования СССР для использования в учебном процессе студентами экономических специальностей вузов

Авторы:
В. Е. Барбаумов, В. И. Ермаков, Н. Н. Кривенцова, А. С. Лебедев, В. И. Матвеев, Е. А. Силаева, О. К. Смагина

Рецензенты:
кафедра исследования операций (Московский государственный университет им. М. В. Ломоносова) и проф. В. А. Треногин (Московский институт стали и сплавов).

Справочник по математике для экономистов/
С 74 В. Е. Барбаумов, В. И. Ермаков, Н. Н. Кривенцова и др.; Под ред. В. И. Ермакова.— М.: Высш. шк., 1987.— 336 с.: ил.

В справочник включены следующие разделы: линейная алгебра, математическое программирование, теория графов и ее приложение, теория вероятностей, статистический анализ и др.,— широко использующиеся в экономических исследованиях и позволяющие проводить различного рода расчеты при решении экономических задач.

$\mathbf{С\,\frac{1702000000-392}{001(01)-87}\,115-87}$ **ББК 22.1**
51

Справочное издание

Барбаумов Виктор Ефимович
Ермаков Валерий Иванович
Кривенцова Наталья Николаевна и др.

СПРАВОЧНИК ПО МАТЕМАТИКЕ ДЛЯ ЭКОНОМИСТОВ

Зав. редакцией Е. С. Гридасова. Редактор Ж. И. Яковлева. Младшие редакторы Н. П. Майкова, Г. В. Вятоха. Обложка художника В. И. Казаковой. Художественный редактор В. И. Пономаренко. Технический редактор Т. Н. Полунина. Корректор Г. И. Кострикова.

ИБ № 6172

Изд. № ФМ-866. Сдано в набор 27.01.87. Подп. в печать 15.07.87. Формат 84×108/32 Бум. тип. № 2. Гарнитура литературная. Печать высокая. Объем 17,64 усл. печ. л. +0,21 усл. печ. л. форзац. 17,96 усл. кр.-отт. 15,99 уч.-изд. л. +0,35 уч.-изд. л. форзац. Тираж 102000 экз. Зак. № 924. Цена 1 р.

Издательство «Высшая школа», 101430, Москва, ГСП-4, Неглинная ул., д. 29/14.

Отпечатано с матриц ордена Октябрьской Революции и ордена Трудового Красного знамени МПО «Первая Образцовая типография» имени А. А. Жданова «Союзполиграфпрома» при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 113054, Москва, Валовая, 28 во Владимирской типографии «Союзполиграфпрома» при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли, 600000, г, Владимир, Октябрьский проспект, д. 7.

ПРЕДИСЛОВИЕ

Современный уровень требований, предъявляемых к экономической теории и практике, обязывает специалистов этого профиля постоянно знакомиться с передовыми идеями модельной структуризации и анализа. В последние годы значительный вес в экономических исследованиях приобрели математические методы.

Большую роль в экономическом моделировании играют методы математического программирования и сетевого планирования, опирающиеся на линейную алгебру, анализ функций одной и многих переменных и некоторые другие разделы математики. Растет роль теории игр при решении экономических задач. Значительное место в экономических исследованиях занимают статистические методы. По всем этим вопросам имеется много специальной литературы, изучение которой требует серьезной математической подготовки. В то же время ощущается потребность в литературе, содержащей информацию из указанных разделов в достаточно обозримой и доступной форме.

В данном справочнике авторы стремились охватить те основные разделы математики, которые в настоящее время применяются при анализе экономических систем. Изложение материала характеризуется достаточной простотой, что позволяет привлечь к математическим методам широкий круг экономистов.

Справочник предназначен для студентов экономических вузов; может быть использован аспирантами и преподавателями вузов, а также экономистами различных специальностей в практической работе.

Авторы

РАЗДЕЛ 1. ОБЩИЕ СВЕДЕНИЯ

1.1. Постоянные величины

Величина	Числовое значение	Величина	Числовое значение
π	3,141 592 654	e	2,718 281 828
$\pi/2$	1,570 796 327	$\sqrt{e}$	1,648 721 271
π^2	9,869 604 401	$1/e$	0,367 879 441
$\sqrt{\pi}$	1,772 453 851	$1/\sqrt{e}$	0,606 530 660
$\sqrt{2\pi}$	2,506 628 275	e^{π}	23,140 692 633
$1/\pi$	0,318 309 886	$M=\lg e$	0,434 294 482
$\ln \pi$	1,144 729 886	$1/M=\ln 10$	2,302 585 093
$\sqrt{2}$	1,414 213 562	$\sqrt[3]{2}$	1,259 921 050
$\sqrt{3}$	1,732 050 808	$\sqrt[3]{3}$	1,442 249 570
$\sqrt{5}$	2,236 067 977	$\sqrt[3]{4}$	1,587 401 052
$\sqrt{6}$	2,449 489 743	$\sqrt[3]{5}$	1,709 975 947
$\sqrt{7}$	2,645 751 311	$\sqrt[3]{6}$	1,817 120 593
$\sqrt{8}$	2,828 427 125	$\sqrt[3]{10}$	2,154 434 690
$\sqrt{10}$	3,162 277 660	$\sqrt[3]{100}$	4,641 588 834
1 радиан	$57°17'44,8''$	$1°$	0,017 453 293 радиан

1.2. Основные алгебраические формулы

$(a+b)^2=a^2+2ab+b^2$; $(a-b)^2=a^2-2ab+b^2$;
$(a+b)^3=a^3+3a^2b+3ab^2+b^3$; $(a-b)^3=a^3-3a^2b+3ab^2-b^3$;
$(a+b)^4=a^4+4a^3b+6a^2b^2+4ab^3+b^4$;
$(a-b)^4=a^4-4a^3b+6a^2b^2-4ab^3+b^4$;
$a^2-b^2=(a-b)(a+b)$;
$a^3+b^3=(a+b)(a^2-ab+b^2)$;
$a^3-b^3=(a-b)(a^2+ab+b^2)$;
$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+a^{n-3}b^2+\ldots+ab^{n-2}+b^{n-1})$;
$(a_1+a_2+\ldots+a_n)^2=a_1^2+a_2^2+\ldots+a_n^2+2(a_1a_2+a_1a_3+\ldots+$
$+a_1a_n+a_2a_3+\ldots+a_{n-1}a_n)$.

В частности,

$$(a_1+a_2+a_3)^2=a_1^2+a_2^2+a_3^2+2a_1a_2+2a_1a_3+2a_2a_3.$$

1.3. Натуральные числа. Разложение на простые множители.

Числа 1, 2, 3, 4, 5, ... называют *натуральными*. *Делителем* натурального числа a называют всякое натуральное число, на которое a делится без остатка (нацело). Натуральное число a называется *простым*, если оно имеет лишь два делителя: 1 и a. Натуральное число, имеющее более двух делителей, называют *составным*. Например, число 17 — простое, число 28 — составное, так как имеет делители 1, 2, 4, 7, 14, 28. Всякое составное число единственным образом представляется в виде произведения простых чисел. Так, $28=2\cdot2\cdot7=2^2\cdot7$; $156=$ $=2\cdot2\cdot3\cdot13=2^2\cdot3\cdot13$.

1.4. Наибольший общий делитель и наименьшее общее кратное

Наибольшим общим делителем (НОД) нескольких натуральных чисел называют наибольшее натуральное число, на которое делится без остатка каждое из данных чисел. Для отыскания НОД нескольких чисел необходимо разложить их на простые множители, а затем составить произведение из общих множителей в наименьших степенях. Например, НОД чисел 54 и 180 равен 18. Действительно, $54=2\cdot3^3$, $180=2^2\cdot3^2\cdot5$. Следовательно, НОД$(54,\ 180)=2\cdot3^2=18$. Понятие НОД используют при сокращении обыкновенных дробей.

Два числа a_1 и a_2 называют *взаимно простыми*, если НОД $(a_1,\ a_2)=1$.

Наименьшим общим кратным (НОК) нескольких натуральных чисел называют наименьшее натуральное число, которое делится на каждое из данных чисел. Для отыскания НОК нескольких чисел необходимо разложить их на простые множители, в полученных разложениях выделить наибольшие степени каждого простого множителя и затем выделенные степени перемножить. Например, НОК чисел 12 и 90 равно 180. В самом деле, $12=2^2\cdot3$, $90=2\cdot3^2\cdot5$ и НОК $(12,\ 90)=2^2\cdot3^2\cdot5=180$.

Понятие НОК используют при сложении и вычитании обыкновенных дробей.

1.5. Обыкновенные и десятичные дроби

Числа 0, ±1, ±2, ±3, ±4, ±5, ... называют *целыми*. Отношение двух целых чисел p и q принято называть *обыкновенной дробью* $\frac{p}{q}$ (используют также запись $p:q$ и p/q). При этом целое число p называют *числителем*, а целое число $q \neq 0$ — *знаменателем* дроби.

Если числа p и q имеют общий делитель, отличный от единицы, то дробь можно сократить. Сокращение дроби производится делением числителя и знаменателя на их общий делитель. Результатом сокращения является дробь, тождественно равная данной дроби. Например, дробь $34/51$ можно сократить на НОД $(34, 51) = 17$, так что $34/51 = 2/3$.

Если $|p| < |q|$ (см. п. 1.8), то дробь называют *правильной*; если $|p| \geqslant |q|$, то *неправильной*. Неправильная дробь может быть представлена в виде суммы целого числа и правильной дроби, т. е. в виде смешанного числа. Например, $\frac{8}{3} = 2 + \frac{2}{3} = 2\frac{2}{3}$.

При сложении (вычитании) обыкновенных дробей a/b и c/d поступают следующим образом:

а) находят НОК (b, d);

б) определяют дополнительные множители для каждой из данных дробей, т. е. находят такие числа r и t, что $br = dt =$ НОК (b, d);

в) строят искомую дробь в виде

$$\frac{ar \pm ct}{\text{НОК}\,(b, d)};$$

г) сокращают полученную дробь.

Например,

$$\frac{\overset{2|}{5}}{12} + \frac{\overset{3|}{3}}{8} = \frac{5\cdot2 + 3\cdot3}{24} = \frac{19}{24};$$

$$\frac{\overset{3|}{7}}{40} - \frac{\overset{5|}{1}}{24} = \frac{7\cdot3 - 1\cdot5}{120} = \frac{16}{120} = \frac{2}{15}.$$

Умножение и деление обыкновенных дробей осуществляют по следующим правилам:

$$\frac{a}{b}\cdot\frac{c}{d} = \frac{a\cdot c}{b\cdot d}; \quad \frac{a}{b} : \frac{c}{d} = \frac{a}{b}\cdot\frac{d}{c} = \frac{a\cdot d}{b\cdot c};$$

при этом полученные результаты необходимо сократить, если это возможно. Например,

$$\frac{2}{5}\cdot\frac{3}{10}=\frac{2\cdot 3}{5\cdot 10}=\frac{3}{25};\quad \frac{3}{7}:\frac{9}{14}=\frac{3}{7}\cdot\frac{14}{9}=\frac{3\cdot 14}{7\cdot 9}=\frac{2}{3}.$$

Числа, представимые обыкновенными дробями, называют *рациональными*. Все целые числа входят в множество рациональных чисел.

Конечной десятичной дробью называют дробь, знаменатель которой является целой положительной степенью числа 10. В этом случае дробь принято записывать без знаменателя, отделяя в числителе запятой (справа налево) столько знаков, сколько нулей в знаменателе. Например, $\frac{3}{10}=0{,}3$; $\frac{1721}{100}=17{,}21$; $\frac{13}{1000}=0{,}013$.

Бесконечная десятичная дробь имеет вид $x_0,\ x_1x_2x_3\ldots$ $\ldots x_n\ldots$, где x_0 — целое число, а каждая из величин $x_1,\ x_2,\ \ldots,\ x_n,\ \ldots$ принимает одно из значений 0, 1, 2,...,9.

Бесконечную десятичную дробь называют *периодической*, если в ее записи начиная с некоторого места бесконечно повторяется одна и та же группа цифр. Эту повторяющуюся группу цифр называют *периодом* дроби. В записи дроби период принято заключать в скобки. Например, дробь 1,6234234234 ... записывают в виде 1,6(234).

Если бесконечная десятичная дробь не содержит периода, то ее называют *непериодической*.

В тех случаях, когда период дроби равен 0 или 9, дробь рассматривают как конечную. Здесь имеют место следующие правила:

$$x_0,\ 000\ldots=x_0;\quad x_0-1,\ 999\ldots=x_0;$$

$$x_0,\ x_1x_2\ldots x_n000\ldots=x_0,\ x_1x_2\ldots x_n\ (x_n\neq 0,\ n=1,\ 2,\ 3,\ldots);$$

$$x_0,\ x_1x_2\ldots(x_n-1)\,999\ldots=x_0,\ x_1x_2\ldots x_n\ (x_n\neq 0,\ n=1,2,3\ldots).$$

Например, $0{,}37\,(9)=0{,}38\,(0)=0{,}38$.

Числа, представимые всевозможными десятичными дробями, называют *действительными (вещественными)*.

Всякое рациональное число представимо либо в виде конечной, либо в виде бесконечной периодической десятичной дроби. Например, $7/22=0{,}3\,(18)$; $3/16=0{,}1875$. Все рациональные числа входят в множество действительных чисел.

Действительные числа, не являющиеся рациональными, принято называть *иррациональными*. Всякое иррациональное число представимо в виде бесконечной непериодической десятичной дроби. Например, $\sqrt{2}=1{,}414213\ldots$; $\pi=3{,}141592\ldots$; $e=2{,}718281\ldots$ — иррациональные числа.

Для любого действительного числа x и для любого сколь угодно малого положительного рационального числа ε найдутся два рациональных числа α_1 и α_2 такие, что $\alpha_1 \leqslant x \leqslant \alpha_2$ и $\alpha_2 - \alpha_1 < \varepsilon$. Числа α_1 и α_2 называют *приближенными значениями* числа x по недостатку и по избытку соответственно при заданной степени точности ε. Например, $1{,}414 \leqslant \sqrt{2} \leqslant 1{,}415$ с точностью до $0{,}001$.

1.6. Проценты

Процентом (%) называют сотую часть числа. Например, 20% от числа 35 составляют $\frac{20}{100}$ его частей и, следовательно, равны $35 \cdot \frac{20}{100} = 35 \cdot \frac{1}{5} = 7$.

Если α % некоторого числа x равны a, то само число $x = \frac{a \cdot 100}{\alpha}$. Например, если 30% некоторого числа x равны 15, то само число $x = \frac{15 \cdot 100}{30} = 50$.

Вклад a при p% годовых от величины вклада через t лет будет равен $a \cdot \left(1 + \frac{pt}{100}\right)$.

Например, вклад в 300 руб. при 2% годовых через 10 лет составит 360 руб. Действительно, 2% годовых от величины вклада дадут ежегодный прирост $300 \cdot \frac{2}{100} = 6$ руб., так что за 10 лет «нарастает» 60 руб.

Сложные проценты — это проценты, начисляемые в определенные сроки как на основной вклад, так и на наращенные за предыдущий срок проценты. В этом случае вклад a при p% годовых через t лет составит $a \cdot \left(1 + \frac{p}{100}\right)^t$. Так, в предыдущем примере за первый год нарастает 6 руб. и вклад станет равным 306 руб. За второй год к 306 руб. добавится еще $306 \cdot \frac{2}{100} = 6$ руб. 12 коп. и вклад окажется равным 312 руб. 12 коп. и т. д.

1.7. Пропорции

Пропорцией называют равенство двух отношений.

Если $\frac{a}{b} = \frac{c}{d}$, то $a \cdot d = b \cdot c$ (основное свойство пропорции).

Кроме того, при любых числах k, l, m и n имеют место равенства

$$\frac{ka+lb}{ma+nb}=\frac{kc+ld}{mc+nd}$$

(производные пропорции).

В частности, $\frac{a\pm b}{b}=\frac{c\pm d}{d}$; $\frac{a-b}{a+b}=\frac{c-d}{c+d}$.

1.8. Абсолютная величина (модуль) действительного числа

Абсолютной величиной (модулем) действительного числа x называют число

$$|x|=\begin{cases} x, & \text{если } x\geqslant 0, \\ -x, & \text{если } x<0. \end{cases}$$

Например, $|7|=7$; $|-3|=-(-3)=3$.

Основные свойства абсолютных величин

1°. $|x|\geqslant 0$; $|x|=0$ тогда и только тогда, когда $x=0$.
2°. $-|x|\leqslant x\leqslant |x|$.
3°. $|x\pm y|\leqslant |x|+|y|$.
4°. $||x|-|y||\leqslant |x-y|$.
5°. $|x\cdot y|=|x|\cdot|y|$.
6°. $\left|\frac{x}{y}\right|=\frac{|x|}{|y|}$ $(y\neq 0)$.
7°. $\sqrt{x^2}=|x|$ (см. п. 1.13).

1.9. Средние величины

Средним арифметическим n чисел $x_1, x_2, \ldots, x_n$ называют величину $\frac{x_1+x_2+\ldots+x_n}{n}$.

Среднее геометрическое n чисел $x_1, x_2, \ldots, x_n$ — это величина $\sqrt[n]{x_1x_2\ldots x_n}$.

Средним квадратическим n чисел $x_1, x_2, \ldots, x_n$ является величина $\sqrt{x_1^2+x_2^2+\ldots+x_n^2}$.

Если $x_1, x_2, \ldots, x_n$ — положительные числа, то всегда справедливо неравенство

$$\sqrt[n]{x_1x_2\ldots x_n}\leqslant\frac{x_1+x_2+\ldots+x_n}{n}\leqslant\sqrt{x_1^2+x_2^2+\ldots+x_n^2}.$$

1.10. Прогрессии и конечные суммы

Арифметической прогрессией называют такую последовательность чисел $a_1, a_2, a_3, \ldots, a_n, \ldots$ — *членов прогресии*, в которой каждое последующее число получается из предыдущего прибавлением некоторого числа d—разности прогрессии.

Например, —1, 3, 7, 11, ...—арифметическая прогрессия с разностью $d=4$.

Формулы арифметической прогрессии

$$a_n=a_1+(n-1)d; \quad \frac{a_{n-k}+a_{n+k}}{2}=a_n;$$

$S_n=\frac{a_1+a_n}{2}\cdot n$ (S_n—сумма n членов арифметической прогрессии).

Геометрической прогрессией называют такую последовательность чисел $a_1, a_2, \ldots, a_n, \ldots$ (членов прогрессии), в которой каждое последующее число получается из предыдущего умножением его на определенное число q (знаменатель геометрической прогрессии).

Например, 2, 8, 32, 128, ...—геометрическая прогрессия со знаменателем $q=4$.

Формулы геометрической прогрессии

$$a_n=a_1q^{n-1}; \quad a_{n-k}a_{n+k}=a_n^2;$$

$$S_n=\frac{a_1(q^n-1)}{q-1}=\frac{a_1(1-q^n)}{1-q}, \quad q\neq 1$$

(S_n—сумма n членов геометрической прогрессии).

Если $q=1$, то $S_n=na_1$.

Если в геометрической прогрессии $|q|<1$, то

$$\lim_{n\to\infty} S_n=\frac{a_1}{1-q}.$$

В этом случае число $S=\frac{a_1}{1-q}$ называют *суммой бесконечно убывающей геометрической прогрессии.*

Например, $1+\frac{1}{3}+\frac{1}{3^2}+\ldots+\frac{1}{3^n}+\ldots=\frac{1}{1-\frac{1}{3}}=\frac{3}{2}$.

Некоторые конечные суммы

$$1^2+2^2+3^2+\ldots+(n-1)^2+n^2=\frac{n(n+1)(2n+1)}{6}.$$

$$1^3+2^3+\ldots+(n-1)^3+n^3=\frac{n^2(n+1)^2}{4}.$$

$$1^2+3^2+5^2+\ldots+(2n-1)^2=\frac{n(4n^2-1)}{3}.$$

$$1^3+3^3+\ldots+(2n-1)^3=n^2(2n^2-1).$$

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{(n-1)n}=1-\frac{1}{n}.$$

1.11. Факториал

Факториалом целого положительного числа n называют произведение $1 \cdot 2 \cdot 3 \ \dots \ (n-1)n$.

Факториал числа n обозначают символом $n!$ По определению, $0!=1$.

Например, $2!=1 \cdot 2=2$; $3!=1 \cdot 2 \cdot 3=6$; $4!=1 \cdot 2 \cdot 3 \cdot 4=24$.

Свойства факториала

1^0. $\frac{n!}{(n-1)!}=n$; $\frac{(2n)!}{(2n-2)!}=(2n-1)\,2n$; $\frac{(2n+1)!}{(2n-1)!}=2n\,(2n+1)$.

2^0. При возрастании n факториал $n!$ растет очень быстро: $\lim\limits_{n \to \infty} \frac{n^k}{n!}=0$ (k — любое натуральное число); $\lim\limits_{n \to \infty} \frac{a^n}{n!}=0$ (a — любое положительное число).

3^0. Факториалы больших чисел можно приближенно оценить по формуле Стирлинга:

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \cdot e^{\frac{\theta}{12n}}, \text{ где } 0<\theta<1.$$

1.12. Размещения, перестановки, сочетания

Всевозможные группировки из данных n элементов по m в каждой, отличающиеся друг от друга либо самими элементами, либо порядком расположения элементов, называют *размещениями* из n элементов по m.

Например, размещения из трех элементов a, b, c по 2: ab, ac, bc, ba, ca, cb.

Число всех размещений из n различных элементов по m обозначают A_n^m и вычисляют по формуле

$$A_n^m=\underbrace{n \cdot (n-1) \cdot (n-2) \dots (n-m+1)}_{m \text{ множителей}}=\frac{n!}{(n-m)!}.$$

Например, $A_5^3=5 \cdot 4 \cdot 3=60$.

Перестановками из n элементов называют их группировки, отличающиеся друг от друга только порядком входящих в них элементов.

Например, перестановки из трех элементов a, b, c: abc, acb, cab, cba, bca, bac.

Число всех различных перестановок из n различных элементов обозначают P_n и вычисляют по формуле

$$P_n=1 \cdot 2 \cdot 3 \dots (n-1) \cdot n=n!$$

Например, $P_5=1 \cdot 2 \cdot 3 \cdot 4 \cdot 5=120$.

Всевозможные группировки из данных n элементов по m в каждой, отличающиеся друг от друга хотя бы одним элементом, называют *сочетаниями* из n элементов по m.

Например, сочетания из четырех элементов a, b, c, d по 2: ab, ac, ad, bc, bd, cd.

Число всех сочетаний из n различных элементов по m обозначают C_n^m или $\binom{m}{n}$ и вычисляют по формуле

$$C_n^m=\frac{n\cdot(n-1)\cdot(n-2)\ldots(n-m+1)}{1\cdot2\cdot3\ldots m}=\frac{n!}{m!\cdot(n-m)!}.$$

Например, $C_5^3=\frac{5\cdot4\cdot3}{1\cdot2\cdot3}=10.$

Свойства сочетаний

1°. $C_n^1=n$; $C_n^0=C_n^n=1$.

2°. $C_n^m=C_n^{n-m}$ $(m=0,\ 1,\ 2,\ \ldots,\ n)$.

3°. $C_n^0+C_n^1+C_n^2+\ldots+C_n^n=2^n$.

4°. $C_n^0-C_n^1+C_n^2-\ldots+(-1)^n C_n^n=0$.

1.13. Степени и корни

Если n—натуральное число, то n-я степень (a^n) некоторого действительного числа a (основания степени) определяется как произведение n сомножителей, равных a $(a^n=a\cdot a\ldots a)$. При этом число n называют *показателем степени*.

По определению, при любом $a\neq0$ считают $a^0=1$, $a^1=a$, $a^{-n}=1:a^n$. Например, $2^3=2\cdot2\cdot2=8$, $2^{-3}=1:2^3=$ $=1/8$.

При любых натуральных показателях m и n справедливы следующие равенства:

$$a^m\cdot a^n=a^{m+n};\ a^m:a^n=a^{m-n}\ (a\neq0);\ (a^m)^n=a^{m\cdot n};$$

$$(a\cdot b)^n=a^n\cdot b^n;\left(\frac{a}{b}\right)^n=\frac{a^n}{b^n}\ (b\neq0).$$

Приведенные соотношения верны и при любых действительных показателях при $a>0$, $b>0$.

Если $a>0$ и n—натуральное число, то *арифметическим корнем* n-й степени из a называют единственное положительное число x такое, что $x^n=a$. Обозначение корня $a^{1/n}$ или $\sqrt[n]{a}$.

Корень второй степени из a (квадратный корень) принято обозначать $\sqrt{a}$.

Если $a=0$, то $\sqrt[n]{a}=0$.

Если $a < 0$, то корень n-й степени из a определяется лишь для нечетных n. В этом случае $\sqrt[n]{a}$ есть единственное отрицательное число x такое, что $x^n = a$.

Например, $\sqrt[4]{16} = 2$, так как $2^4 = 16$ и $2 > 0$; $\sqrt[3]{27} = 3$; $\sqrt[3]{-64} = -4$.

Если $a \geqslant 0$, m и n — натуральные числа, то, по определению, считают

$$a^{m/n} = (a^{1/n})^m = (a^m)^{1/n}; \quad a^{-m/n} = 1 : a^{m/n} \ (a \neq 0).$$

При этом имеют место следующие равенства:

$$a^{m/n} = (\sqrt[n]{a})^m = \sqrt[n]{a^m}; \quad \sqrt[m]{\sqrt[n]{a}} = \sqrt[m \cdot n]{a};$$

$$\sqrt[n]{ab} = \sqrt[n]{a}\sqrt[n]{b}; \quad \sqrt[n]{\frac{a}{b}} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}} \ (b \neq 0).$$

Для преобразования корней из целых чисел полезно подкоренное число разложить на простые множители. Например, $\sqrt{1156} = \sqrt{4 \cdot 289} = \sqrt{2^2 \cdot 17^2} = 2 \cdot 17 = 34$; $\sqrt[3]{9261} = \sqrt[3]{27 \cdot 343} = \sqrt[3]{3^3 \cdot 7^3} = 3 \cdot 7 = 21$.

1.14. Бином Ньютона

При любых действительных a и b и любом натуральном n справедливы равенства

$$(a+b)^n = a^n + C_n^1 a^{n-1} b + \ldots + C_n^k a^{n-k} b^k + \ldots + b^n;$$
$$(a-b)^n = a^n - C_n^1 a^{n-1} b + \ldots + (-1)^k C_n^k a^{n-k} b^k + \ldots + (-1)^n b^n.$$

Например,

$$(a+b)^4 = a^4 + 4a^3 b + 6a^2 b^2 + 4ab^3 + b^4;$$
$$(a-b)^5 = a^5 - 5a^4 b + 10a^3 b^2 - 10a^2 b^3 + 5ab^4 - b^5.$$

1.15. Логарифмы

Логарифмом $\log_a N$ числа N $(N > 0)$ при основании a $(a > 0,\ a \neq 1)$ называют показатель степени, в которую нужно возвести основание a, чтобы получить N.

Например, $\log_2 32 = 5$, так как $2^5 = 32$; $\log_3 81 = 4$, так как $3^4 = 81$.

Свойства логарифмов

При любых $a > 0$, $a \neq 1$, $b > 0$, $b \neq 1$, любых $N > 0$, $N_1 > 0$, $N_2 > 0$ и любом α имеют место следующие равенства:

1°. $\log_a 1 = 0$; $\log_a a = 1$.

2°. $a^{\log_a N} = N$.

3°. $\log_a (N_1 N_2) = \log_a N_1 + \log_a N_2$.

4°. $\log_a \left(\frac{N_1}{N_2}\right) = \log_a N_1 - \log_a N_2$.

5°. $\log_a (N^\alpha) = \alpha \cdot \log_a N$.

6°. $\log_{a^\alpha} N = \frac{1}{\alpha} \cdot \log_a N$.

7°. $\log_a N = \frac{1}{\log_N a}$ $(N \neq 1)$.

8°. $\log_a N = \frac{\log_b N}{\log_b a}$.

Логарифмы по основанию 10 называют *десятичными* (обозначение: $\lg N$), а по основанию $e = 2{,}71828\ldots$ — натуральными (обозначение: $\ln N$). При этом

$$\ln N \approx 2{,}3 \cdot \lg N.$$

1.16. Многочлены

Многочленом от неизвестного x называют выражение вида

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n,$$

где $a_0, a_1, \ldots, a_{n-1}, a_n$ — некоторые числа.

Числа $a_0, a_1, \ldots, a_{n-1}$ — *коэффициенты* при степенях x, а a_n — *свободный член* многочлена.

Наивысшую степень x, входящую в многочлен с ненулевым коэффициентом, называют *степенью* многочлена. Степень многочлена $f(x)$ обозначают $\deg f(x)$.

Два многочлена считают равными, если при одинаковых степенях x стоят равные коэффициенты.

Многочлены можно складывать и перемножать. Например, если $f(x) = x^3 - 2x^2 - 3$, а $\varphi(x) = x^2 - x + 2$, то

$$f(x) + \varphi(x) = (x^3 - 2x^2 - 3) + (x^2 - x + 2) = x^3 - x^2 - x - 1;$$
$$f(x) \cdot \varphi(x) = (x^3 - 2x^2 - 3)(x^2 - x + 2) = x^5 - 2x^4 - 3x^2 -$$
$$- x^4 + 2x^3 + 3x + 2x^3 - 4x^2 - 6 = x^5 - 3x^4 + 4x^3 - 7x^2 + 3x - 6.$$

Для любых двух многочленов $f(x)$ и $\varphi(x)$ можно найти такие многочлены $q(x)$ и $r(x)$, что

$$f(x) = \varphi(x)\, q(x) + r(x),$$

где $\deg r(x) < \deg \varphi(x)$ или $r(x) \equiv 0$. Многочлены $q(x)$ и $r(x)$, удовлетворяющие этому условию, определены однозначно. Их соответственно называют *частным* и *остатком* от деления $f(x)$ на $\varphi(x)$.

○ **Пример.** Разделить многочлен $f(x)=2x^3-3x^2+5x-4$ на $\varphi(x)=x^2-4x+3$.

Имеем

$$\begin{array}{r|l}
-\,2x^3-3x^2+5x-4 & x^2-4x+3 \\ \cline{2-2}
2x^3-8x^2+6x & 2x+5 \\ \cline{1-1}
-\,5x^2-x-4 & \\
5x^2-20x+15 & \\ \cline{1-1}
19x-19 &
\end{array}$$

Здесь $q(x)=2x+5$ — частное, $r(x)=19x-19$ — остаток от деления $f(x)$ на $\varphi(x)$. ●

Если остаток от деления многочлена $f(x)$ на многочлен $\varphi(x)$ тождественно равен 0, т. е. $f(x)=\varphi(x)q(x)$, то говорят, что $f(x)$ *нацело делится* на $\varphi(x)$. Число α называют *корнем многочлена* $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_{n-1}x+a_n$, если

$$f(\alpha)=a_0\alpha^n+a_1\alpha^{n-1}+\ldots+a_{n-1}\alpha+a_n=0.$$

Например, число $x=2$ — корень многочлена $f(x)=3x^3-2x^2-10x+4$, так как $f(2)=3\cdot2^3-2\cdot2^2-10\cdot2+4=0$.

Свойства корней многочлена

1°. Многочлен степени n не может иметь более чем n различных корней.

2°. Если α — корень многочлена $f(x)$, то $f(x)$ нацело делится на $(x-\alpha)$, т. е. $f(x)=(x-\alpha)q(x)$.

3°. Корни многочлена ax^2+bx+c (второй степени) можно найти по формуле

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}\quad(\text{если } b^2-4ac\geqslant 0).$$

1.17. Рациональные дроби

Выражение вида $\frac{f(x)}{\varphi(x)}$, где $f(x)$ и $\varphi(x)$ — многочлены, называют *рациональной дробью*.

Две рациональные дроби $\frac{f_1(x)}{\varphi_1(x)}$ и $\frac{f_2(x)}{\varphi_2(x)}$ считают *равными*, если $f_1(x)\varphi_2(x)=\varphi_1(x)f_2(x)$.

Рациональные дроби можно складывать и перемножать. Например,

$$\frac{^{x-3|}x-1}{x^2+x-2}+\frac{^{x^2+x-2|}x^2}{x-3}=\frac{(x-1)(x-3)+x^2(x^2+x-2)}{(x^2+x-2)(x-3)}=$$

$$=\frac{x^2-x-3x+3+x^4+x^3-2x^2}{x^3+x^2-2x-3x^2-3x+6}=\frac{x^4+x^3-x^2-4x+3}{x^3-2x^2-5x+6};$$

$$\frac{x-1}{x^2+x-2}\cdot\frac{x^2}{x-3}=\frac{(x-1)x^2}{(x^2+x-2)(x-3)}=\frac{x^3-x^2}{x^3-2x^2-5x+6}.$$

Рациональная дробь $\frac{f(x)}{\varphi(x)}$ называется *правильной*, если степень числителя $f(x)$ меньше степени знаменателя $\varphi(x)$.

Всякая рациональная дробь представима, притом единственным способом, в виде суммы многочлена и правильной дроби.

Например, $\frac{x^2+3}{x-1}=x+1+\frac{4}{x-1}$.

1.18. Графики элементарных функций

Ц е л а я р а ц и о н а л ь н а я ф у н к ц и я (м н о г о ч л е н)

а) Линейная функция $y=ax+b$ (рис. 1.1). Графиком этой функции является прямая линия. Функция возрас-

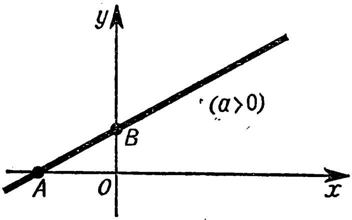


Рис. 1.1

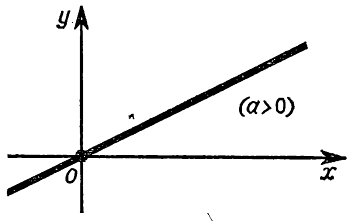


Рис. 1.2

тает при $a>0$ и убывает при $a<0$. Оси координат пересекаются прямой в точках $A\left(-\frac{b}{a};\ 0\right)$ и $B(0;\ b)$. В случае $b=0$ получаем прямую пропорциональность $y=ax$

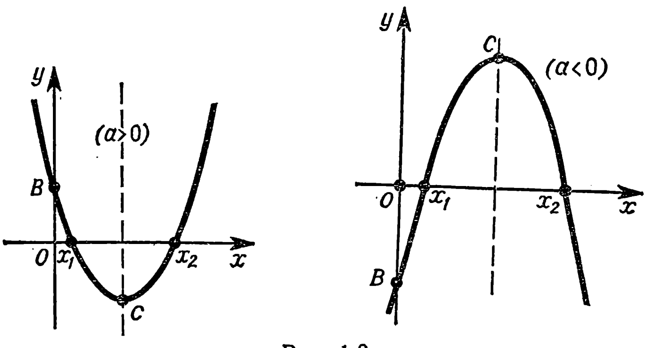
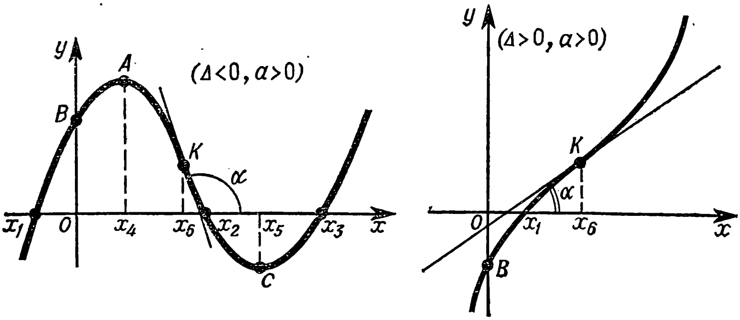


Рис. 1.3

(рис. 1.2). График функции проходит через начало координат.

б) Квадратичная функция $y=ax^2+bx+c$ (рис. 1.3). Графиком функции является парабола с осью симметрии, параллельной оси ординат. При $a>0$ ветви параболы направлены вверх, при $a<0$ — вниз. Ось ординат пересекается в точке $B(0;\ c)$. Вершина параболы C имеет координаты $\left(-\frac{b}{2a};\ \frac{4ac-b^2}{4a}\right)$. Абсциссы x_1, x_2 точек пересечения параболы с осью Ox определяют по формуле $x_{1,\,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$. Величины x_1 и x_2 являются корнями квадратного уравнения $ax^2+bx+c=0$ в том случае, когда



оно имеет решения на множестве действительных чисел.

в) Многочлен третьей степени $y=ax^3+bx^2+cx+d$ (рис. 1.4). Графиком функции является кубическая парабола. Поведение функции зависит от знаков a и $\Delta=3ac-b^2$. В случае $\Delta\geqslant 0$ функция возрастает при $a>0$ и убывает при $a<0$. Если же $\Delta<0$, то функция имеет одну точку

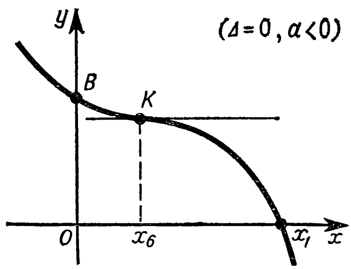


Рис. 1.4

максимума и одну точку минимума. Кубическая парабола имеет одну точку перегиба K. Ось ординат пересекается кривой в точке $B(0;\ d)$. Абсциссы точек максимума и минимума x_4 и x_5 определяют по формуле $x=\frac{-b\pm\sqrt{-\Delta}}{3a}$. Абсцисса точки перегиба x_6 равна $\left(-\frac{b}{3a}\right)$. Касательная к графику в точке перегиба наклонена к оси Ox под углом α таким, что $\operatorname{tg}\alpha=\frac{\Delta}{3a}$.

г) Степенна́я функция $y=ax^n$ ($n>1$ — целое) (рис. 1.5).

Графиком функции является парабола n-го порядка, которая проходит через точки $O(0,\ 0)$ и $A(1,\ a)$ и касается оси Ox в начале координат. При n четном график функции симметричен относительно оси Oy и в начале координат имеет минимум при $a > 0$ и максимум при

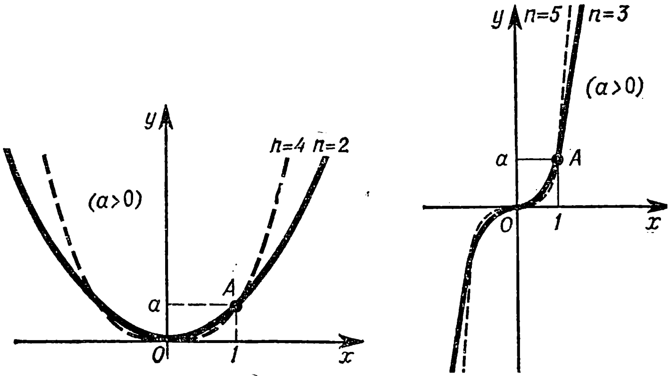


Рис. 1.5

$a < 0$. При n нечетном график функции симметричен относительно начала координат, которое является точкой перегиба графика.

Дробно-рациональная функция

а) Обратная пропорциональность $y = \frac{a}{x}$ (рис. 1.6). Графиком функции является равносторонняя гипербола, ветви

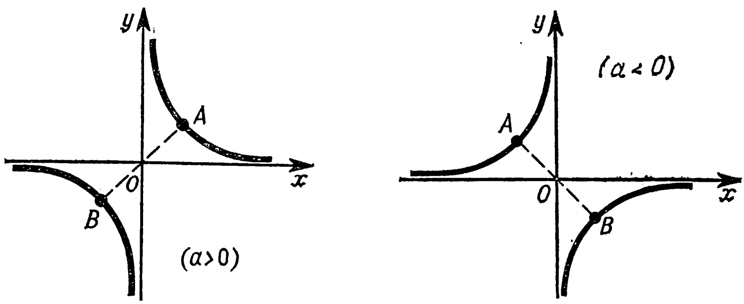


Рис. 1.6

которой симметричны относительно начала координат. Оси координат служат асимптотами графика. В случае $a > 0$ гипербола имеет вершины в точках $A(\sqrt{a};\ \sqrt{a})$, $B(-\sqrt{a};\ -\sqrt{a})$. В случае $a < 0$ вершины имеют координаты $A(-\sqrt{|a|};\ \sqrt{|a|})$, $B(\sqrt{|a|};\ -\sqrt{|a|})$.

б) Дробно-линейная функция $y=\frac{bx+b}{cx+d}$ (рис. 1.7). Графиком функции является равносторонняя гипербола. Асимптотами служат прямая $x=-\frac{d}{c}$, параллельная оси Oy, и прямая $y=\frac{a}{c}$, параллельная оси Ox. Расположение ветвей гиперболы зависит от знака величины $\Delta=\frac{bc-ad}{c^2}$.

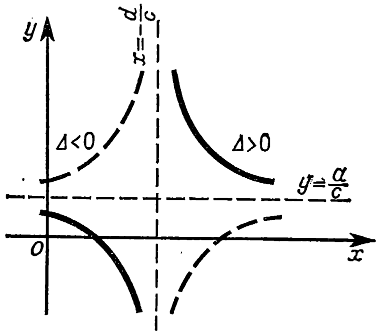


Рис. 1.7

в) Степенна́я функция $y=\frac{a}{x^n}$ ($n>1$ — целое) (рис. 1.8). Графиком функции является кривая гиперболического типа, проходящая через точку (1; a). Оси координат служат асимптотами графика. При n четном график симметричен относительно оси Oy, при n нечетном имеет место симметрия относительно начала координат.

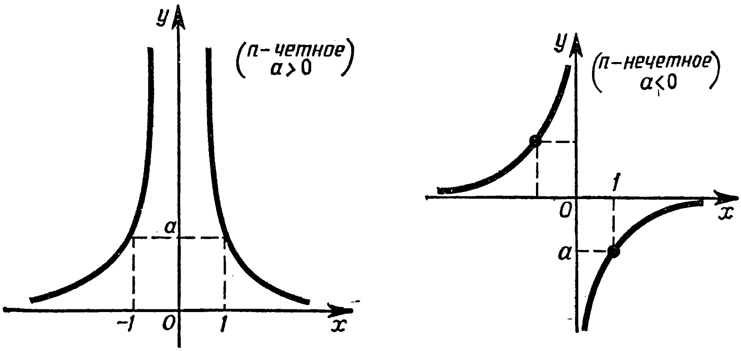


Рис. 1.8

Некоторые иррациональные функции (рис. 1.9).

Показательные и логарифмические функции

а) Показательная функция $y=a^x$ ($a>0$, $a\neq1$) (рис. 1.10).

График функции при любом a проходит через точку

(0; 1) и асимптотически приближается к оси Ox. Функция принимает только положительные значения.

б) Логарифмическая функция $y=\log_a x$ ($a>0$, $a\neq 1$) (рис. 1.11). График функции при любом a проходит через

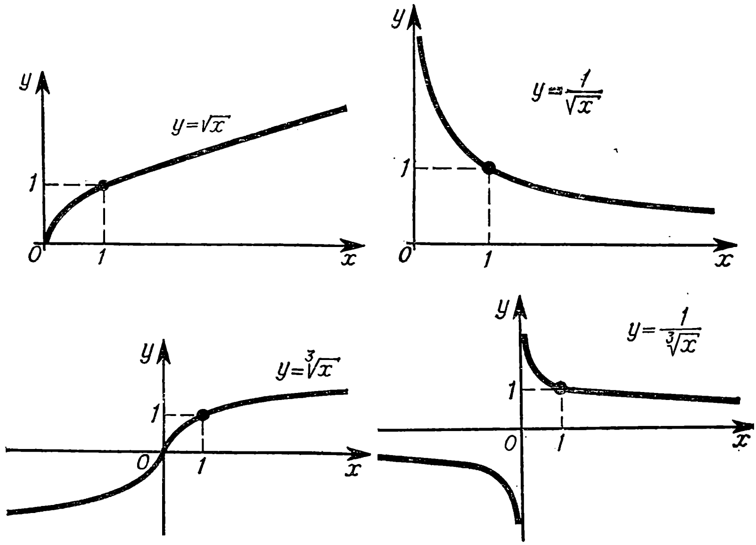


Рис. 1.9

точку (1; 0) и асимптотически приближается к оси Oy. Функция определена только для положительных значений аргумента x.

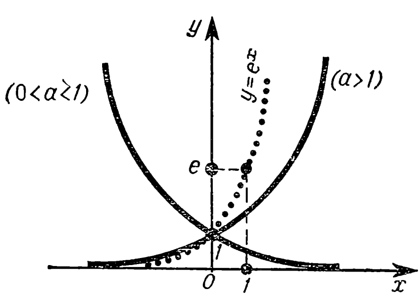


Рис. 1.10

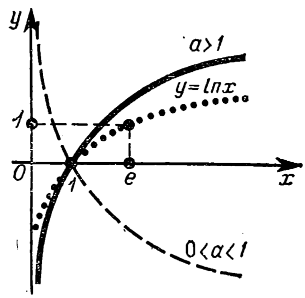


Рис. 1.11

Замечание. Важное место в исследованиях многих явлений (в частности, экономических) занимают показательная функция $y=e^x$ и логарифмическая функция $y=\ln x$ ($y=\log_e x$). Число e — иррациональное ($e\approx 2{,}72$).

в) Кривая Гаусса $y=e^{-x^2}$ (рис. 1.12). График функции имеет одну точку максимума $A(0;\ 1)$, две точки перегиба $B\left(\frac{1}{\sqrt{2}};\ \frac{1}{\sqrt{e}}\right)$ и $C\left(-\frac{1}{\sqrt{2}};\ \frac{1}{\sqrt{e}}\right)$, симметричен относительно оси ординат и асимптотически приближается к оси абсцисс.

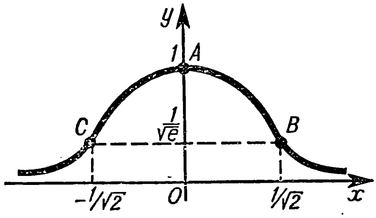


Рис. 1.12

Тригонометрические функции

а) Синус и косинус $y=\sin x$ и $y=\cos x$ (рис. 1.13). Функции $\sin x$ и $\cos x$ периодические с периодом 2π.

б) Тангенс и котангенс $y=\operatorname{tg} x$ и $y=\operatorname{ctg} x$ (рис. 1.14). Функции $\operatorname{tg} x$ и $\operatorname{ctg} x$ периодические с периодом π.

в) Секанс, косеканс $y=\sec x$ и $y=\operatorname{cosec} x$ (рис. 1.15); $\sec x=\frac{1}{\cos x}$, $\operatorname{cosec} x=\frac{1}{\sin x}$. Функции $\sec x$ и $\operatorname{cosec} x$ периодические с периодом 2π.

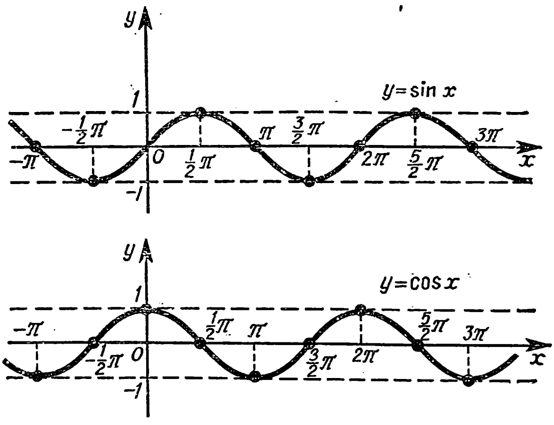


Рис. 1.13

Обратные тригонометрические функции

а) Арксинус и арккосинус $y=\arcsin x$ и $y=\arccos x$ (рис. 1.16).

Функция $y=\arcsin x$ каждому действительному числу $x\in[-1,\ 1]$ ставит в соответствие угол $y\in\left[-\frac{\pi}{2},\ \frac{\pi}{2}\right]$ такой, что $\sin y=x$. Например, $\arcsin\frac{1}{2}=\frac{\pi}{6}$, так как $\sin\frac{\pi}{6}=\frac{1}{2}$.

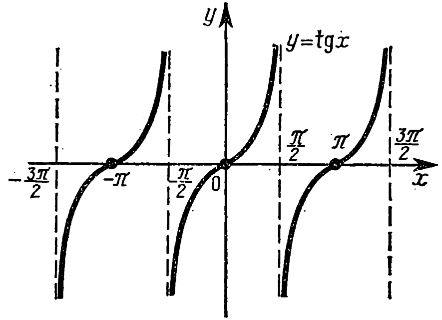


Рис. 1.14

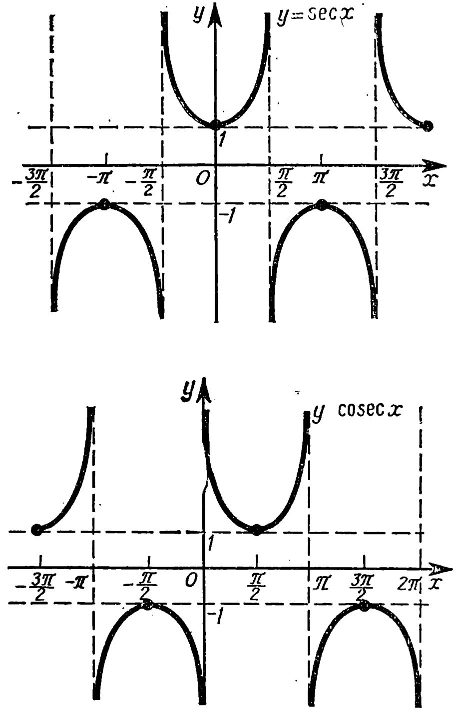
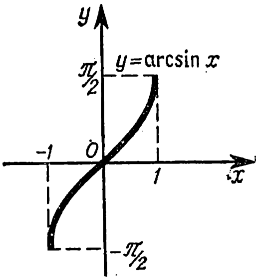


Рис. 1.15

Функция $y = \arccos x$ каждому действительному числу $x \in [-1, 1]$ ставит в соответствие угол $y \in [0, \pi]$ такой, что $\cos y = x$. Например, $\arccos \frac{1}{2} = \frac{\pi}{3}$, так как $\cos \frac{\pi}{3} = \frac{1}{2}$.



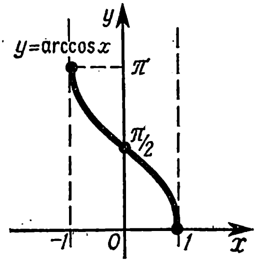
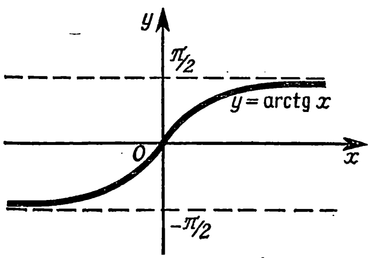


Рис. 1.16

б) Арктангенс и арккотангенс $y = \operatorname{arctg} x$ и $y = \operatorname{arcctg} x$ (рис. 1.17). Функция $y = \operatorname{arctg} x$ каждому действительному числу $x \in]-\infty, +\infty[$ ставит в соответствие угол $y \in]-\pi/2, \pi/2[$ такой, что $\operatorname{tg} y = x$. Например, $\operatorname{arctg} \sqrt{3} = \pi/3$, так как $\operatorname{tg}(\pi/3) = \sqrt{3}$.



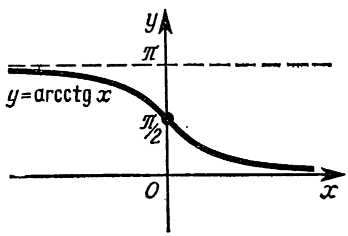


Рис. 1.17

Функция $y = \operatorname{arcctg} x$ каждому действительному числу $x \in]-\infty, +\infty[$ ставит в соответствие угол $y \in]0, \pi[$ такой, что $\operatorname{ctg} y = x$. Например, $\operatorname{arcctg} 1 = \pi/4$, так как $\operatorname{ctg}(\pi/4) = 1$.

1.19. Примеры неэлементарных функций и важнейших кривых

Неэлементарные функции

а) $y = [x]$ (читается: «y равно антье x») — целая часть x. Определяется как наибольшее целое число, не превосходящее x (рис. 1.18). Например, $[3{,}24]=3$; $[0{,}7]=0$; $[-5{,}4]=-6$.

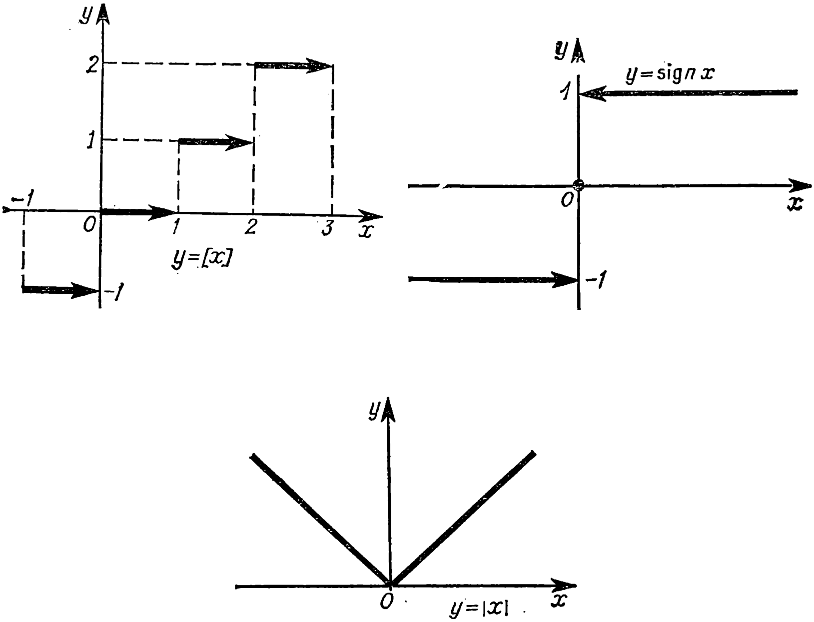


Рис. 1.18

б) $y=\operatorname{sign} x$ (читается: «y равно сигнум x») — знак числа x (рис. 1.18):

$$\operatorname{sign} x=\begin{cases}-1, & \text{если } x<0,\\ 0, & \text{если } x=0,\\ 1, & \text{если } x>0.\end{cases}$$

в) $y=|x|$ — абсолютная величина (модуль) x (рис. 1.18):

$$|x|=\begin{cases}x, & \text{если } x\geqslant 0,\\ -x, & \text{если } x<0.\end{cases}$$

Важнейшие кривые

а) Эллипс $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ (рис. 1.19).

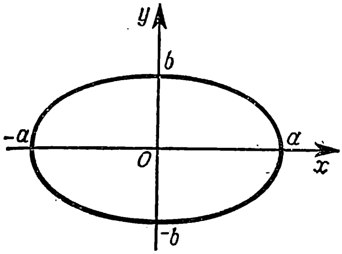


Рис. 1.19

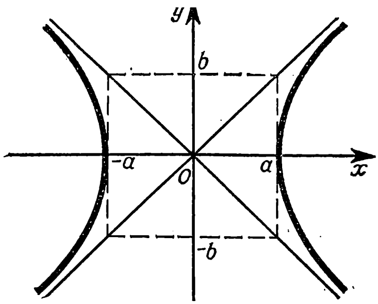


Рис. 1.20

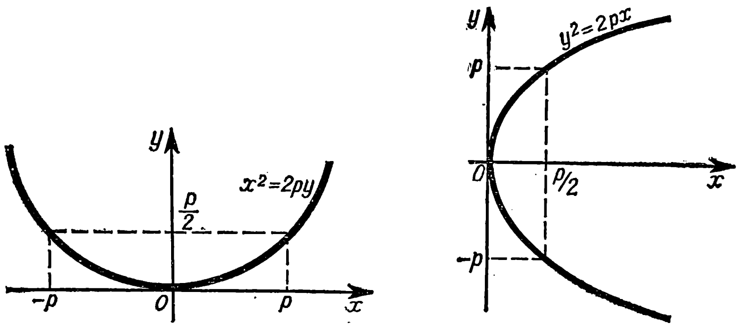


Рис. 1.21

б) Гипербола $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$ (рис. 1.20).

в) Парабола $x^2=2py$ или $y^2=2px$ (рис. 1.21).

1.20. Понятие множества

Понятие «множество» — одно из первичных (неопределяемых) понятий математики. Описательно термин «множество» объясняется как совокупность, коллекция, набор некоторых объектов произвольной природы, объединенных по каким-то общим для них признакам. Объекты, из которых состоит множество, называют его *элементами* (*точками*). Символическая запись $a \in A$ означает принадлежность элемента a множеству A. Запись $a \notin A$ означает, что элемент a не принадлежит множеству A.

Множество A называют *подмножеством* другого множества B, если каждый элемент множества A является одновременно элементом множества B. В этом случае пишут $A \subset B$ (читается: «A включается или содержится в B»).

Множества A и B равны ($A=B$) тогда и только тогда, когда $A\subset B$ и $B\subset A$, т. е. если эти множества состоят из одних и тех же элементов.

Множество, не содержащее ни одного элемента, называется *пустым* и обозначается символом $\varnothing$. Любое множество содержит $\varnothing$ в качестве подмножества. Очевидно, $A\subset A$; A и $\varnothing$ называют *несобственными подмножествами* множества A. Все остальные подмножества множества A называют *собственными*.

Множество A элементов x, обладающих свойством $P(x)$, символически записывают в виде $A=\{x\,|\,P(x)\}$. Например, $A=\{x\,|\,x=2k,\ k=1,\ 2,\ \ldots\}$ означает, что множество A состоит из четных положительных целых чисел 2, 4, 6, 8, ...

1.21. Операции над множествами

Объединением (суммой) множеств A и B называют множество $A\cup B$ всех элементов, принадлежащих хотя бы одному из множеств A и B; $C=A\cup B=\{x\,|\,x\in A$ или $x\in B\}$.

Пересечением множеств A и B называют множество $A\cap B$ всех элементов, принадлежащих одновременно и множеству A и множеству B; $C=A\cap B=\{x\,|\,x\in A$ и $x\in B\}$.

Замечание. Понятия объединения и пересечения могут быть обобщены на случай любого числа множеств (конечного или бесконечного). Если даны множества A_1, A_2, ..., A_k, ..., то символическая запись $\bigcup\limits_{k=1}^{\infty} A_k$ означает объединение данных множеств, т. е. определяет множество, каждый элемент которого принадлежит хотя бы одному из данных множеств. Символическая запись $\bigcap\limits_{k=1}^{\infty} A_k$ означает пересечение данных множеств, т. е. определяет множество, каждый элемент которого принадлежит всем данным множествам.

Разностью множеств A и B называют множество $A\setminus B$ тех элементов множества A, которые не содержатся во множестве B; $C=A\setminus B=\{x\,|\,x\in A$ и $x\notin B\}$.

Если $B\subset A$, то разность $A\setminus B$ называют *дополнением* множества B до множества A и обозначают C_AB.

Декартовым произведением множеств A и B называют множество $A\times B$ всех упорядоченных пар элементов $(a,\ b)$, где $a\in A$, $b\in B$. Элементы a и b называют при этом *компонентами* (*координатами*) пары $(a,\ b)$.

Декартово произведение $A_1 \times A_2 \times \ldots \times A_n$ множеств $A_1, A_2, \ldots, A_n$ представляет собой множество всех упорядоченных n-ок (энок) элементов $(a_1, a_2, \ldots, a_n)$, где $a_1 \in A_1, \ldots, a_n \in A_n$. В частности, декартово произведение $\underbrace{R \times R \times \ldots \times R}_{n}$, где R — множество действительных чисел, определяет n-мерное арифметическое пространство R^n (см. п. 3.1).

Примеры.

1. Если A — множество четных положительных чисел, а B — множество нечетных положительных чисел, то $A \cup B$ определяет множество натуральных чисел, т. е. множество $N = \{1, 2, 3, \ldots, n, \ldots\}$.

2. Если A — множество всех чисел, делящихся на 2, а B — множество всех чисел, делящихся на 5, то $A \cap B$ определяет множество всех чисел, делящихся и на 2 и на 5, т. е. делящихся на 10.

3. Если $A = \{1, 2, 3, 4, 5\}$, а $B = \{3, 5\}$, то $C_A B = A \setminus B = \{1, 2, 4\}$, а $B \setminus A = \varnothing$.

4. Если $A = \{1, 2\}$, а $B = \{3, -1, 0\}$, то $A \times B = \{(1, 3), (1, -1), (1, 0), (2, 3), (2, -1), (2, 0)\}$.

Введенные операции обладают следующими свойствами:

1°. $A \cup \varnothing = A$. 2°. $A \cap \varnothing = \varnothing$.

3°. $A \cup A = A$; $A \cap A = A$ (идемпотентность).

4°. $A \cup B = B \cup A$; $A \cap B = B \cap A$ (коммутативность).

5°. $A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$ (ассоциативность).

6°. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$; $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ — (дистрибутивность).

Если $A \subset E$ и $B \subset E$, то:

7°. $C_E(A \cup B) = C_E A \cap C_E B$

или $E \setminus (A \cup B) = (E \setminus A) \cap (E \setminus B)$;

$C_E(A \cap B) = C_E A \cup C_E B$ или

$E \setminus (A \cap B) = (E \setminus A) \cup (E \setminus B)$ — законы двойственности.

1.22. Отображение. Функция

Отображение — одно из основных понятий математики. Пусть A и B — два непустых множества. Если каждому элементу $x \in A$ ставится в соответствие по правилу f один вполне определенный элемент $y \in B$, то говорят, что задано *отображение* множества A в множество B. Обозначение: $f\colon A \to B$. При этом $y = f(x)$ называют *образом элемента* x, а x — *прообразом элемента* y. Множество всех $y \in B$, в которые переходят различные $x \in A$, называют *множе-*

ством значений отображения f и обозначают $f(A)$. Очевидно, $f(A) \subseteq B$.

Если при отображении f каждый элемент $y \in B$ соответствует некоторому элементу $x \in A$, то говорят об отображении множества A на множество B.

○ **Примеры.**

1. Поставим в соответствие каждому слову некоторого словаря (на русском языке) его заглавную букву. Такое соответствие определяет отображение множества слов словаря в множество букв русского алфавита.

2. Поставим в соответствие каждому трехзначному числу цифру его десятков. Такое соответствие определяет отображение множества трехзначных чисел на множество $B=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. ●

Отображение f называют *обратимым*, если из условия $x_1 \neq x_2$ $(x_1, x_2 \in A)$ вытекает $y_1 \neq y_2$ $(y_1, y_2 \in B)$, т. е. разным прообразам соответствуют разные образы. В этом случае каждый образ y имеет единственный прообраз x и можно

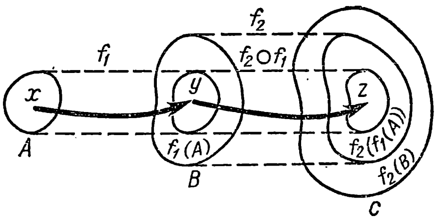


Рис. 1.22

определить отображение $f^{-1}: f(A) \to A$, называемое *обратным* к отображению f. Обратное отображение устанавливает *взаимно однозначное соответствие* между множествами A и $f(A)$, т. е. такое соответствие, при котором каждому элементу множества A соответствует единственный элемент множества $f(A)$ и каждому элементу множества $f(A)$ соответствует единственный элемент множества A.

Если заданы отображения $f_1: A \to B$ и $f_2: B \to C$, то отображение $f_2 \circ f_1$, сопоставляющее каждому элементу $x \in A$ определенный элемент $z \in C$ такой, что $z=f_2(y)$, где $y=f_1(x)$, называют *суперпозицией* отображений f_1 и f_2 (рис. 1.22).

Отображение f называют *функционалом*, если множество B является множеством действительных чисел $(B=R)$. Если же и множество A—числовое, то отображение f называют *функцией*. В частности, если $A \subseteq R$, то говорят о функции $y=f(x)$ одной переменной x. Множество $f(A)=\{y\in R \mid y=f(x),\ x\in A\}$ принято обозначать $E(f)$ и называть *областью значений* функции. Если $A\subseteq R^n$ (см. n-мерное арифметическое пространство), то говорят о функции $y=f(x_1, x_2, \ldots, x_n)$ n переменных $x_1, x_2, \ldots, x_n$ (см. п. 4.1.) В этом случае

$$E(f)=\{y\in R \mid y=f(x_1, x_2, \ldots, x_n),\ (x_1, x_2, \ldots, x_n)\in A\}.$$

1.23. Мощность множества

Множества A и B называют *эквивалентными* или *равномощными* ($A\sim B$), если между ними можно установить взаимно однозначное соответствие (см. п. 1.22).

Множество A является бесконечным, если оно эквивалентно некоторому своему собственному подмножеству; в противном случае множество A — конечное.

Мощность конечного множества совпадает с количеством его элементов.

Всякое бесконечное множество, эквивалентное множеству N натуральных чисел, называется *счетным*.

Из любого бесконечного множества можно выделить счетное подмножество. Всякое бесконечное подмножество счетного множества является счетным множеством.

Объединение конечного или счетного множества счетных множеств есть счетное множество. Декартово произведение конечного множества счетных множеств есть счетное множество.

Множества $\mathbf{Z}$ (целых чисел) и $\mathbf{Q}$ (рациональных чисел) есть счетные множества.

Множество $\mathbf{R}$ (действительных чисел) несчетно.

Всякое бесконечное множество, эквивалентное множеству действительных чисел, называют *множеством мощности континуума*.

1.24. Числовые множества. Грани числового множества

Множество натуральных чисел

$$\mathbf{N}=\{n\}=\{1, 2, \ldots, n, \ldots\}.$$

Множество целых чисел

$$\mathbf{Z}=\{n\}\cup\{0\}\cup\{-n\}=\{0, \pm 1, \pm 2, \pm 3, \ldots, \pm n, \ldots\}.$$

Множество рациональных чисел

$$\mathbf{Q}=\left\{\frac{p}{q}\right\}, \text{ где } p\in\mathbf{Z},\ q\in\mathbf{Z},\ q\neq 0.$$

Множество действительных (вещественных) чисел $\mathbf{R}=\{x\}$. Имеет место такое последовательное включение: $\mathbf{N}\subset\mathbf{Z}\subset\mathbf{Q}\subset\mathbf{R}$.

Все указанные числовые множества обладают свойством упорядоченности, т. е. для любых двух различных элементов a и b любого из данных множеств можно сказать, что либо $a>b$, либо $a<b$. Кроме того, выполняется свойство транзитивности: из $a>b$ и $b>c$ следует, что $a>c$.

Множества $\mathbf{Q}$ и $\mathbf{R}$ являются всюду плотными множествами. Это означает, что между любыми двумя различными элементами a и b любого из указанных множеств найдется хотя бы один элемент этого же множества. Таким элементом является, например, элемент $c=\frac{a+b}{2}$.

Множество $\mathbf{R}$ обладает важным свойством непрерывности. Оно постулирует возможность установления взаимно однозначного соответствия (см. п. 1.22) между множеством действительных чисел и множеством точек на прямой линии.

Пусть $A=\{x\}$ — некоторое непустое множество действительных чисел.

Множество A называют *ограниченным сверху* (*снизу*), если существует действительное число K такое, что для всех $x\in A$ выполняется неравенство $x\leqslant K$ ($x\geqslant K$).

Всякое число K с указанным свойством называют *верхней* (*нижней*) гранью множества A.

Множество называют *ограниченным*, если оно ограничено и сверху и снизу.

Наименьшую из верхних граней множества A называют *точной верхней гранью* этого множества и обозначают символом $\sup A$ (супремум A).

Наибольшую из нижних граней множества A называют *точной нижней гранью* этого множества и обозначают символом $\inf A$ (инфимум A).

Свойства точной верхней и точной нижней граней.

1°. Для любого элемента $x\in \mathbf{A}$ выполняется неравенство $x\leqslant \sup \mathbf{A}$ ($x\geqslant \inf \mathbf{A}$).

2°. Для любого числа $\varepsilon>0$ найдется элемент $\bar{x}\in\mathbf{A}$ такой, что $\bar{x}>\sup A-\varepsilon$ ($\bar{x}<\inf \mathbf{A}+\varepsilon$).

Всякое непустое ограниченное сверху (снизу) множество действительных чисел имеет точную верхнюю (нижнюю) грань.

○ **Примеры.**

1. $A =]a, b[= \{x \mid a < x < b\}$ — ограниченный открытый интервал. Здесь $\sup A = b$, $\inf A = a$ не принадлежат данному множеству.

2. $A = [a, b] = \{x \mid a \leqslant x \leqslant b\}$ — ограниченный замкнутый интервал или отрезок. Здесь $\sup A = b$, $\inf A = a$ принадлежат данному множеству.

3. $A =]-\infty, a[= \{x \mid -\infty < x < a\}$; $B =]a, +\infty[= \{x \mid a < x < +\infty\}$; $R =]-\infty; +\infty[$ — неограниченные открытые интервалы. Здесь $\sup A = a$, $\inf B = a$ не принадлежат указанным множествам.

4. $A = [a, b[= \{x \mid a \leqslant x < b\}$; $B =]a, b] = \{x \mid a < x \leqslant b\}$; $C =]-\infty; a] = \{x \mid -\infty < x \leqslant a\}$; $D = [a, +\infty[= \{x \mid a \leqslant x < +\infty\}$ — полуоткрытые интервалы. Здесь $\inf A = a$, $\sup B = b$, $\sup C = a$, $\inf D = a$ принадлежат указанным множествам; $\sup A = b$, $\inf B = a$ — не принадлежат им. ●

1.25. Комплексные числа

Мнимую единицу i определяют как число, квадрат которого равен (-1). Таким образом, $i^2 = -1$.

Всякое комплексное число представляют в виде $z = a + bi$ (алгебраическая форма записи комплексного числа). Здесь a и b — действительные числа. При этом a называют *действительной частью* комплексного числа z ($a = \operatorname{Re} z$), b — его мнимой частью ($b = \operatorname{Im} z$).

Если $a = 0$, то $z = bi$ называют *чисто мнимым* числом.

Если $b = 0$, то $z = a$, т. е. комплексное число z равно действительному числу a. Множество действительных чисел, таким образом, есть подмножество множества комплексных чисел.

Два комплексных числа $z_1 = a_1 + b_1 i$ и $z_2 = a_2 + b_2 i$ считают равными ($z_1 = z_2$), если и только если $a_1 = a_2$ и $b_1 = b_2$. В противном случае $z_1 \neq z_2$. Отношений «больше», «меньше» для комплексных чисел не существует.

Всякое комплексное число $z = a + bi$ удобно изображать точкой (a, b) или соответствующим радиусом-вектором на комплексной плоскости (рис. 1.23). Оси Ox и Oy прямоугольной декартовой системы координат называют при этом соответственно *действительной* и *мнимой* осью. Величину ρ — длину радиуса-вектора точки z называют *модулем* комплексного числа z и обозначают $|z|$. Угол φ (в радианах) называют *аргументом* комплексного числа z и обозначают $\varphi = \arg z$. Имеют место соотношения $a = \rho\cos\varphi$, $b = \rho\sin\varphi$. Тогда $z = \rho(\cos\varphi + i\sin\varphi)$ — *тригонометрическая* форма записи комплексного числа; с другой стороны $\rho = \sqrt{a^2 + b^2}$, $\operatorname{tg}\varphi = \frac{b}{a}$, при этом $0 \leqslant \rho < +\infty$, а $-\infty < \varphi < +\infty$. Более того, для данного комплексного числа z аргумент φ имеет бесконечное множество значений, отличающихся друг от друга на величину $2\pi k$ (k — целое число). Главное значение аргумента заключено в промежутке $-\pi < \varphi \leqslant \pi$.

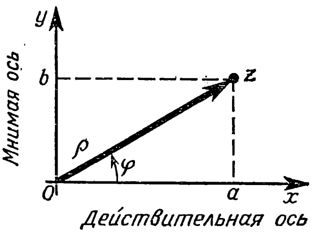


Рис. 1.23

Для числа $z = 0$ $(a = b = 0)$ аргумент не определяется, а $|z| = 0$.

Имеет место показательная форма записи комплексного числа $a + bi = \rho e^{\varphi i}$. При этом $e^{\varphi i} = \cos\varphi + i\sin\varphi$ (*формула Эйлера*).

Два комплексных числа z и $\bar{z}$ называют *взаимно сопряженными*, если они имеют равные действительные части и отличающиеся лишь знаком мнимые части ($\operatorname{Re}\bar{z} = \operatorname{Re} z$, $\operatorname{Im}\bar{z} = -\operatorname{Im} z$).

Очевидно, что $|\bar{z}| = |z|$, а $\arg\bar{z} = -\arg z$, так что на комплексной плоскости точки z и $\bar{z}$ симметричны относительно действительной оси Ox.

При этом

$$z = a + bi = \rho(\cos\varphi + i\sin\varphi) = \rho e^{\varphi i};$$
$$\bar{z} = a - bi = \rho(\cos\varphi - i\sin\varphi) = \rho e^{-\varphi i}.$$

Пусть даны два комплексных числа $z_1 = a_1 + b_1 i = \rho_1(\cos\varphi_1 + i\sin\varphi_1)$ и $z_2 = a_2 + b_2 i = \rho_2(\cos\varphi_2 + i\sin\varphi_2)$; тогда:

1°. $z_1 \pm z_2 = (a_1 \pm a_2) + (b_1 + b_2)\cdot i$;

2°. $z_1 \cdot z_2 = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)\, i = \rho_1\rho_2[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)]$.

3°. $\frac{z_1}{z_2} = \frac{(a_1 a_2 + b_1 b_2) + (a_2 b_1 - a_1 b_2)\, i}{a_2^2 + b_2^2} = \frac{\rho_1}{\rho_2}\cdot[\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)]$, $(z_2 \neq 0)$.

4°. $z_1\cdot\bar{z}_1=a_1^2+b_1^2=\rho_1^2$; $\overline{z_1+z_2}=\bar{z}_1+\bar{z}_2$; $\overline{z_1\cdot z_2}=\bar{z}_1\cdot\bar{z}_2$; $\overline{z_1/z_2}=\dfrac{\bar{z}_1}{\bar{z}_2}$ $(\bar{z}_2\neq 0)$.

5°. $z_1^n=\rho_1^n(\cos n\varphi_1+i\sin n\varphi_1)$, где n — целое число (формула Муавра).

В частности, $i^3=-i$, $i^4=1$, …, $i^{4n+m}=i^m$.

6°. $\sqrt[n]{z_1}=\sqrt[n]{\rho_1}\left(\cos\dfrac{\varphi_1+2\pi k}{n}+i\sin\dfrac{\varphi_1+2\pi k}{n}\right)$, где n — натуральное число, а $k=0,\ 1,\ 2,\ \ldots,\ n-1$.

РАЗДЕЛ II. ЛИНЕЙНАЯ АЛГЕБРА

2.1. Линейные уравнения

Линейным уравнением относительно неизвестных $x_1, x_2, \ldots, x_n$ называют выражение вида

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,$$

где $a_1, a_2, a_3, \ldots, a_n, b$ — числа.

Последовательность n чисел $k_1, k_2, \ldots, k_n$ называют *решением* линейного уравнения с n неизвестными $x_1, x_2, \ldots, x_n$, если после подстановки $x_1 = k_1, x_2 = k_2, \ldots, x_n = k_n$ в данное уравнение оно превращается в верное числовое соотношение.

○ **Пример.** Уравнение $2x_1 + 3x_2 - 5x_3 + x_4 = 4$ имеет решение 2, 1, 1, 2, так как после подстановки $x_1 = 2$, $x_2 = 1$, $x_3 = 1$, $x_4 = 2$ получаем верное числовое соотношение $2 \cdot 2 + 3 \cdot 1 - 5 \cdot 1 + 2 = 4$. Последовательность же чисел 3, 2, 0, 1 не является решением данного уравнения, так как после подстановки $x_1 = 3$, $x_2 = 2$, $x_3 = 0$, $x_4 = 1$ получим числовое соотношение $2 \cdot 3 + 3 \cdot 2 - 5 \cdot 0 + 1 = 4$, которое неверно. ●

Линейное уравнение

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

не имеет решений. Оно называется *противоречивым.*

Линейное уравнение $0x_1 + 0x_2 + \ldots + 0x_n = 0$ называют *тривиальным*. Каждая последовательность чисел $k_1, k_2, \ldots, k_n$ является решением тривиального уравнения.

2.2. Системы линейных уравнений

Конечную совокупность линейных уравнений относительно неизвестных $x_1, x_2, \ldots, x_n$ называют *системой линейных уравнений*. Если перенумеровать уравнения системы, то система линейных уравнений запишется в следующем общем виде:

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

где a_{ij} — коэффициент при неизвестном x_j из i-го уравнения, b_i — свободный член i-го уравнения системы.

Числа a_{11}, a_{12}, ..., a_{mn} называют *коэффициентами*, а b_1, b_2, ..., b_m — *свободными членами* системы линейных уравнений.

Системы линейных уравнений часто записывают в виде таблицы

x_1	x_2	...	x_n	
a_{11}	a_{12}	...	a_{1n}	b_1
a_{21}	a_{22}	...	a_{2n}	b_2
...	...	...	...	...
a_{m1}	a_{m2}	...	a_{mn}	b_m

в i-й строке которой, $1 \leqslant i \leqslant m$, записаны коэффициенты при неизвестных и свободный член i-го уравнения системы.

Решением системы уравнений называют такой упорядоченный набор чисел k_1, k_2, ..., k_n, который является решением каждого уравнения системы. Решить систему уравнений — значит найти все ее решения или убедиться в том, что их нет.

Система уравнений либо несовместна (не имеет ни одного решения), либо является определенной (имеет единственное решение), либо является неопределенной (имеет бесконечное множество решений).

Систему уравнений, которая имеет хотя бы одно решение, называют совместной.

Если одно из уравнений системы является противоречивым, то система несовместна.

Две системы линейных уравнений называют *равносильными*, если они имеют одни и те же решения.

2.3. Разрешенные системы линейных уравнений

Неизвестное x_i называют *разрешенным*, если какое-нибудь уравнение системы содержит неизвестное x_i с коэффициентом единица, а во все остальные уравнения системы неизвестное x_i не входит.

○ **Пример.** Система уравнений

$$
\begin{cases}
x_1 \quad + 3x_3 - \qquad 3x_5 = 5, \\
\quad -7x_3 + x_4 + \ x_5 = 8, \\
x_2 + 2x_3 \qquad - \ x_5 = 1
\end{cases}
\tag{2.1}
$$

содержит разрешенные неизвестные x_1, x_2, x_4. Неизвестные же x_3 и x_5 не являются разрешенными. ●

Если каждое уравнение системы содержит разрешенное неизвестное, то такую систему называют *разрешенной.*

Совокупность неизвестных $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$ называют *набором разрешенных неизвестных* данной системы линейных уравнений, если каждое неизвестное x_{i_k}, $1 \leqslant k \leqslant r$, является разрешенным и каждое уравнение данной системы содержит ровно одно неизвестное из набора $x_{i_1}, x_{i_2}, \ldots, x_{i_r}$.

Разрешенная система уравнений обладает набором разрешенных неизвестных.

Все неизвестные разрешенной системы уравнений, которые не входят в данный набор разрешенных неизвестных, называют *свободными.*

Для отыскания решения разрешенной системы уравнений надо свободным неизвестным придать какие-либо значения, подставить их в систему уравнений и найти значения разрешенных неизвестных. Полученная совокупность значений неизвестных является решением разрешенной системы уравнений.

Все решения разрешенной системы уравнений могут быть получены указанным способом.

○ **Пример.** Найти решение разрешенной системы линейных уравнений (2.1).

○ Из каждого уравнения системы выберем разрешенные неизвестные x_1, x_2, x_4. Тогда неизвестные x_3, x_5 являются свободными. Придадим свободным неизвестным x_3, x_5 значения $x_3 = 1$, $x_5 = 2$ и подставим их в систему уравнений

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

Из полученной системы находим: $x_1 = 8$, $x_2 = 1$, $x_4 = 13$, т. е. упорядоченный набор чисел 8, 1, 1, 13, 2 является решением рассматриваемой системы уравнений. ●

Разрешенная система уравнений всегда совместна. Если все неизвестные разрешенной системы уравнений образуют набор разрешенных неизвестных, то она имеет единственное решение. В противном случае разрешенная система уравнений имеет бесчисленное множество решений.

2.4. Метод Гаусса построения общего решения системы линейных уравнений

Общим решением совместной системы линейных уравнений называют равносильную ей разрешенную систему линейных уравнений.

Для отыскания всех решений совместной системы линейных уравнений достаточно найти ее общее решение.

Метод построения общего решения совместной системы линейных уравнений называется методом Гаусса.

Общее решение строят из исходной системы уравнений с помощью элементарных преобразований, под которыми понимается любое из следующих действий:

1) вычеркивание уравнения, у которого все коэффициенты при неизвестных и свободный член равны нулю;

2) умножение обеих частей какого-либо уравнения системы на отличное от нуля число;

3) замена i-го уравнения системы уравнением, которое получается путем прибавления к i-му уравнению системы ее j-го уравнения, умноженного на число.

Элементарные преобразования переводят данную систему уравнений в равносильную ей систему.

Пусть дана система линейных уравнений, записанная в табличной форме:

x_1	...	x_s	...	x_n	
a_{11}	...	a_{1s}	...	a_{1n}	b_1
...	...	...	...	...	...
a_{r1}	...	$\boxed{a_{rs}}$	...	a_{rn}	b_r
...	...	...	...	...	...
a_{m1}	...	a_{ms}	...	a_{mn}	b_n

(2.2)

Возьмем любой отличный от нуля коэффициент a_{rs} системы уравнений. *Жордановым преобразованием* системы с ведущим элементом $a_{rs} \neq 0$ называется совокупность следующих преобразований:

1) умножение r-й строки таблицы (2.2) на число $1/a_{rs}$

x_1	...	x_s	...	x_n	
a_{11}	...	a_{1s}	...	a_{1n}	b_1
...	...	...	...	...	...
a'_{r1}	...	1	...	a'_{rn}	b'_r
...	...	...	...	...	...
a_{m1}	...	a_{ms}	...	a_{mn}	b_m

(2.3)

2) прибавление к первой строке таблицы (2.3) ее r-й строки, умноженной на $-a_{1s}$, прибавление ко второй строке r-й строки, умноженной на $-a_{2s}$, и т. д. После этих преобразований система уравнений (2.3) принимает вид

x_1	...	x_s	...	x_n	
a'_{11}	...	0	...	a'_{1n}	b'_1
...	...	...	...	...	...
a'_{21}	...	1	...	a'_{2n}	b'_r
...	...	...	...	...	...
a'_{m1}	...	0	...	a'_{mn}	b'_m

(2.4)

В результате жорданова преобразования с ведущим элементом a_{rs} получим систему (2.4), у которой неизвестное x_s является разрешенным.

Если проделать одно или несколько жордановых преобразований над данной системой, то получим систему, равносильную исходной.

○ **Пример.** Выполнить жорданово преобразование системы уравнений

$$\begin{cases} 2x_1+7x_2+4x_3+\ \ x_4=6, \\ 3x_1+5x_2+2x_3+2x_4=4, \\ 4x_1+4x_2+\ \ x_3+7x_4=2 \end{cases}$$

с ведущим элементом $a_{23}=2$.

Запишем данную систему в виде таблицы

x_1	x_2	x_3	x_4	
2	7	4	1	6
3	5	$\boxed{2}$	2	4
4	4	1	7	2

(2.5)

После умножения 2-й строки таблицы (2.5) на элемент $1/a_{23}=1/2$ получим

x_1	x_2	x_3	x_4	
2	7	4	1	6
3/2	5/2	1	1	2
4	4	1	7	2

(2.6)

Теперь из первого и третьего уравнений системы (2.6) исключим неизвестное x_3. Для этого к первой строке прибавим вторую строку, умноженную на —4, а к третьей строке прибавим вторую строку, умноженную на —1. После выполнения этих преобразований получим систему уравнений

x_1	x_2	x_3	x_4	
-4	-3	0	-3	-2
$3/2$	$5/2$	1	1	2
$5/2$	$3/2$	0	6	0

(2.7)

Таким образом, в результате жорданова преобразования с ведущим элементом $a_{23}=2$ система (2.5) преобразовалась в систему уравнений (2.7). ●

Преобразование совместной системы уравнений

x_1	x_2	$\ldots$	x_n	
a_{11}	a_{12}	$\ldots$	a_{1n}	b_1
a_{21}	a_{22}	$\ldots$	a_{2n}	b_2
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
a_{m1}	a_{m2}	$\ldots$	a_{mn}	b_m

(2.8)

в общее решение методом Гаусса состоит из выполнения ряда последовательных шагов, причем перед выполнением очередного шага надо в системе уравнений вычеркнуть все тривиальные уравнения.

1-й шаг. Выберем в первом уравнении любой отличный от нуля коэффициент при неизвестном и выполним жорданово преобразование системы (2.8) с этим ведущим элементом.

На k-м шаге, $k=2, 3, \ldots$, выполняем жорданово преобразование системы, полученной после выполнения предыдущего шага, с любым ненулевым коэффициентом k-го уравнения этой системы. После выполнения k-го шага получим систему, содержащую не менее k уравнений, причем каждое из первых k уравнений будет содержать разрешенное неизвестное.

Если полученная после выполнения k-го шага система содержит ровно k нетривиальных уравнений, то процесс решения прекращают. Если же эта система содержит более чем k нетривиальных уравнений, то необходимо выполнить $(k+1)$-й шаг. Не более чем через m шагов (m — число уравнений в системе (2.8)) получим общее решение системы (2.8).

○ **Примеры. 1.** Найти общее решение системы уравнений

$$\begin{cases} 2x_1 - 3x_2 + x_3 = 3, \\ x_1 - x_3 = 3, \\ 3x_1 + x_2 = 8, \\ 13x_2 - 3x_3 = 8. \end{cases}$$

Запишем эту систему в виде таблицы и будем выполнять шаги до тех пор, пока процесс преобразования не закончится:

x_1	x_2	x_3	
2	−3	$\boxed{1}$	3
1	0	−1	3
3	1	0	8
0	13	−3	8

→

x_1	x_2	x_3	
2	−3	1	3
$\boxed{3}$	−3	0	6
3	1	0	8
6	4	0	17

→

x_1	x_2	x_3	
0	−1	1	−1
1	−1	0	2
0	$\boxed{4}$	0	2
0	10	0	5

→

x_1	x_2	x_3	
0	0	1	−1/2
1	0	0	5/2
0	1	0	1/2
0	0	0	0

Приходим к общему решению: $x_3=-1/2$, $x_1=5/2$, $x_2=1/2$. Эта система обладает единственным решением; следовательно, исходная система оказалась определенной.

2. Найти общее решение системы уравнений

$$\begin{cases} 2x_1+7x_2+3x_3+\ \ x_4=6, \\ 3x_1+5x_2+2x_3+2x_4=4, \\ 9x_1+4x_2+\ \ x_3+7x_4=2. \end{cases}$$

○ Имеем:

x_1	x_2	x_3	x_4	
2	7	3	$\boxed{1}$	6
3	5	2	2	4
9	4	1	7	2

→

x_1	x_2	x_3	x_4	
2	7	3	1	6
$\boxed{-1}$	−9	−4	0	−8
−5	−45	−20	0	−40

→

x_1	x_2	x_3	x_4	
0	−11	−5	1	−10
1	9	4	0	8
0	0	0	0	0

Общее решение исходной системы имеет вид

$$\begin{cases} \qquad -11x_2-5x_3+x_4=-10, \\ x_1+\ \ 9x_2+4x_3 \qquad\ =8. \end{cases}$$ ●

Методом Гаусса можно не только построить общее решение совместной системы, но и установить, является ли исходная система уравнений совместной.

○ **Пример.** Установить, является ли система уравнений

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

совместной.

Преобразуем систему уравнений методом Гаусса:

x_1	x_2	x_3	x_4	x_5	
—2	1	—3	2	—4	1
4	—2	5	1	7	2
2	—1	1	8	2	1

$\longrightarrow$

x_1	x_2	x_3	x_4	x_5	
2	1	—3	2	—4	1
0	0	—1	5	—1	4
0	0	—2	10	—2	2

$\longrightarrow$

x_1	x_2	x_3	x_4	x_5	
2	1	0	—13	—1	—11
0	0	1	— 5	1	—4
0	0	0	0	0	—6

Получена система уравнений, которая содержит противоречивое уравнение $0x_1+0x_2+0x_3+0x_4+0x_5=-6$. Следовательно, исходная система уравнений несовместна. ●

2.5. Векторы. Действия с n-мерными векторами

Последовательность n чисел $a_1, a_2, \ldots, a_n$ называют n-мерным вектором x:

$$x=(a_1,\ a_2,\ \ldots,\ a_n).$$

Число a_1 называют *первой координатой* вектора x, a_2 — *второй координатой* и т. д. Количество координат у вектора x называют его *размерностью*.

Если у n-мерных векторов $x=(a_1,\ a_2,\ \ldots,\ a_n)$ и $y=(b_1,\ b_2,\ \ldots,\ b_n)$, имеющих одну и ту же размерность, одноименные координаты равны, т. е. если $a_1=b_1$, $a_2=b_2$, $\ldots$, $a_n=b_n$, то такие векторы называют *равными* и пишут $x=y$. Если же хотя бы одна пара одноименных координат у векторов x и y различна, то $x\neq y$.

Вектор, у которого все координаты равны нулю, называют *нулевым*:

$$\theta=(0,\ 0,\ \ldots,\ 0).$$

Суммой n-мерных векторов $x=(a_1,\ a_2,\ \ldots,\ a_n)$, $y=(b_1, b_2, \ldots, b_n)$ называется n-мерный вектор

$$x+y=(a_1+b_1,\ a_2+b_2,\ \ldots,\ a_n+b_n).$$

Для каждого n-мерного вектора x

$$x+\theta=x.$$

Умножение вектора $x=(a_1, a_2, \ldots, a_n)$ на число определено следующим образом:

$$xk=kx=(a_1k, a_2k, \ldots, a_nk).$$

Вектор $(-1)x$ называют вектором, *противоположным* x, и обозначают $-x$. Вместо $x+(-1)y$ пишут $x-y$. Вектор $x-y$ называют *разностью* векторов x и y.

Свойства операций сложения векторов и умножения вектора на число (x, y, z — n-мерные векторы, k_1, k_2, k — числа):

1°. $x+y=y+x$,

2°. $(x+y)+z=x+(y+z)$,

3°. $(x+y)k=xk+yk$,

4°. $x(k_1+k_2)=xk_1+xk_2$,

5°. $x(k_1k_2)=(xk_1)k_2$.

Скалярным произведением n-мерных векторов $x=(a_1, a_2, \ldots, a_n)$ и $y=(b_1, b_2, \ldots, b_n)$ называют число, обозначаемое xy и равное сумме парных произведений соответственных координат векторов x и y:

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

Свойства скалярного произведения (x, y, z — n-мерные векторы, k — число):

1°. $xy=yx$.

2°. $x(y+z)=xy+xz$.

3°. $k(xy)=(kx)y=x(ky)$.

4°. $xx\geqslant 0$, причем $xx=0$ тогда и только тогда, когда $x=\theta$.

2.6. Длина вектора. Угол между n-мерными векторами

Длиной n-мерного вектора $x=(a_1, a_2, \ldots, a_n)$ называют число $|x|$, равное

$$|x|=\sqrt{xx}=\sqrt{x_1^2+x_2^2+\ldots+x_n^2}.$$

○ **Пример.** Найти длину вектора $x=(-12, 3, -4)$.

Имеем $|x|=\sqrt{xx}=\sqrt{(-12)^2+3^2+(-4)^2}=13$. ●

Каждый n-мерный вектор имеет длину, причем нулевой вектор является единственным вектором, длина которого равна нулю.

Скалярное произведение xx называют скалярным квадратом вектора x и обозначают x^2. Квадрат длины вектора равен его скалярному квадрату, т. е. $|x|^2 = x^2$.

Если x и y — произвольные n-мерные векторы, то их длины $|x|$ и $|y|$ связаны со скалярным произведением xy соотношением

$$|xy| \leqslant |x| \cdot |y|,$$

которое называется *неравенством Коши — Буняковского*. Это неравенство в координатной форме имеет вид

$$|a_1b_1 + a_2b_2 + \ldots + a_nb_n| \leqslant$$
$$\leqslant \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}\ \sqrt{b_1^2 + b_2^2 + \ldots + b_n^2},$$

где $x = (a_1, a_2, \ldots, a_n)$, $y = (b_1, b_2, \ldots, b_n)$.

Для каждой пары n-мерных векторов x, y справедливо соотношение

$$|x + y| \leqslant |x| + |y|,$$

которое называется *неравенством треугольника*.

Углом φ между ненулевыми n-мерными векторами x и y называют угол (от 0 до π), косинус которого равен

$$\cos\varphi = \frac{xy}{|x||y|}.$$

Откуда $xy = |x||y|\cos\varphi$, т. е. скалярное произведение векторов x и y равно произведению длин этих векторов на косинус угла между ними.

2.7. Линейные комбинации векторов и векторная форма записи систем линейных уравнений

Вектор $A_1k_1 + A_2k_2 + \ldots + A_mk_m$ называется *линейной комбинацией* векторов A_1, A_2, $\ldots$, A_m с коэффициентами k_1, k_2, $\ldots$, k_m.

○ **Пример.** Дана система векторов $A_1 = (1, 2, 5, -9)$, $A_2 = (-1, 3, 1, -5)$, $A_3 = (0, 7, -2, 4)$, $A_4 = (1, -2, -2, 3)$. Найти координаты линейной комбинации $2A_1 - 3A_2 + A_3 - 0A_4$.

Выполняя указанные операции над векторами, получим

$$2A_1-3A_2+A_3-0A_4=2\begin{pmatrix}1\\2\\5\\-9\end{pmatrix}-3\begin{pmatrix}-1\\3\\1\\-5\end{pmatrix}+\begin{pmatrix}0\\7\\-2\\4\end{pmatrix}-$$

$$-0\begin{pmatrix}1\\-2\\-2\\3\end{pmatrix}=\begin{pmatrix}2\\4\\10\\-18\end{pmatrix}+\begin{pmatrix}3\\-9\\-3\\15\end{pmatrix}+\begin{pmatrix}0\\7\\-2\\4\end{pmatrix}+\begin{pmatrix}0\\0\\0\\0\end{pmatrix}=\begin{pmatrix}5\\2\\5\\1\end{pmatrix}. \quad ●$$

Используя введенные операции над векторами, запишем систему линейных уравнений

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=b_1,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=b_2,\\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=b_m \end{cases} \tag{2.9}$$

в векторной форме.

Обозначим через $A_1, A_2, \ldots, A_n$ столбцы коэффициентов при неизвестных $x_1, x_2, \ldots, x_n$, а через B — столбец свободных членов системы уравнений, т. е.

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

Теперь систему линейных уравнений (2.9) можно записать в виде

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

или

$$A_1x_1+A_2x_2+\ldots+A_nx_n=B.$$

Последовательность чисел $k_1, k_2, \ldots, k_n$ является решением системы (2.9) тогда и только тогда, когда справедливо соотношение

$$A_1k_1+A_2k_2+\ldots+A_nk_n=B.$$

2.8. Разложение вектора по системе векторов

По определению, n-мерный вектор B разлагается по системе n-мерных векторов $A_1, A_2, \ldots, A_n$, если можно подобрать такие числа $k_1, k_2, \ldots, k_n$, что векторы B и $A_1k_1 + A_2k_2 + \ldots + A_nk_n$ равны, т. е. $B = A_1k_1 + A_2k_2 + \ldots$ $\ldots + A_nk_n$.

Числа $k_1, k_2, \ldots, k_n$ называются *коэффициентами разложения*.

Чтобы найти разложение вектора B по системе векторов $A_1, A_2, \ldots, A_n$, достаточно найти какое-нибудь решение системы линейных уравнений

$$A_1x_1 + A_2x_2 + \ldots + A_nx_n = B.$$

○ **Пример.** Даны система векторов A_1, A_2, A_3 и вектор B:

$$A_1 = \begin{pmatrix} 2 \\ 4 \\ 3 \\ 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \\ 3 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 \\ -1 \\ 10 \\ -3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 \\ 7 \\ 17 \\ 0 \end{pmatrix}.$$

Выяснить, разлагается ли вектор B по системе векторов A_1, A_2, A_3.

Найдем общее решение системы линейных уравнений

$$A_1x_1 + A_2x_2 + A_3x_3 = B,$$

Имеем

x_1	x_2	x_3	
2	−3	1	2
4	0	−1	7
3	1	10	17
0	3	−3	0

$\longrightarrow$

x_1	x_2	x_3	
2	−3	1	2
6	−3	0	9
−17	31	0	−3
6	−6	0	6

$\longrightarrow$

x_1	x_2	x_3	
−4	0	1	−7
−2	1	0	−3
$\boxed{45}$	0	0	90
−6	0	0	−12

$\longrightarrow$

x_1	x_2	x_3	
0	0	1	1
0	1	0	1
1	0	0	2
0	0	0	0

Исходная система уравнений равносильна системе $x_3 = 1$, $x_2 = 1$, $x_1 = 2$, которая имеет единственное решение 2, 1, 1. Следовательно, $B = 2A_1 + A_2 + A_3$, т. е. вектор B разлагается по системе векторов A_1, A_2, A_3. ●

Разложения вектора B по системе $A_1, A_2, \ldots, A_n$:

$$B = k_1A_1 + k_2A_2 + \ldots + k_nA_n,$$
$$B = l_1A_1 + l_2A_2 + \ldots + l_nA_n$$

называются *различными*, если $k_i \neq l_i$ хотя бы при одном значении i, $1 \leqslant i \leqslant n$.

2.9. Линейная зависимость векторов

Система векторов $A_1, A_2, \ldots, A_n$ называется *линейно зависимой*, если можно подобрать такие числа $k_1, k_2, \ldots, k_n$, которые не все равны нулю, что

$$A_1k_1 + A_2k_2 + \ldots + A_nk_n = \theta.$$

Если же каждая линейная комбинация векторов A_1, $A_2, \ldots, A_n$ с коэффициентами $k_1, k_2, \ldots, k_n$, которые не все равны нулю, отлична от нулевого вектора, то система векторов $A_1, A_2, \ldots, A_n$ называется *линейно независимой*.

Система m-мерных векторов $A_1, A_2, \ldots, A_n$ является линейно зависимой, если система линейных уравнений

$$A_1x_1 + A_2x_2 + \ldots + A_nx_n = \theta \qquad (2.10)$$

имеет ненулевое решение. Если же система уравнений (2.10) не имеет ненулевых решений, то система векторов $A_1, A_2, \ldots, A_n$ линейно независимая.

○ **Примеры.**

1. Выяснить, является ли система векторов

$$A_1 = \begin{pmatrix} 3 \\ 5 \\ 1 \\ 4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -2 \\ 1 \\ -5 \\ -7 \end{pmatrix}, \quad A_3 = \begin{pmatrix} -1 \\ -2 \\ 0 \\ -1 \end{pmatrix}$$

линейно зависимой или линейно независимой.

Преобразуем систему линейных уравнений $A_1x_1 + A_2x_2 + A_3x_3 = \theta$ методом Гаусса:

x_1	x_2	x_3	
3	−2	$\boxed{-1}$	0
5	1	−2	0
1	−5	0	0
4	−7	−1	0

$\longrightarrow$

x_1	x_2	x_3	
−3	2	1	0
$\boxed{-1}$	5	0	0
1	−5	0	0
1	−5	0	0

$\longrightarrow$

x_1	x_2	x_3	
0	−13	1	0
1	−5	0	0
0	0	0	0
0	0	0	0

Общее решение исходной системы имеет вид

$$\begin{cases} \quad -13x_2 + x_3 = 0, \\ x_1 - 5x_2 \qquad = 0. \end{cases}$$

Эта система имеет ненулевое решение 5, 1, 13. Следовательно, векторы A_1, A_2, A_3 линейно зависимы.

2. Выяснить, является ли система векторов

$$A_1 = \begin{pmatrix} -20 \\ -15 \\ -4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -7 \\ -2 \\ -4 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix}$$

линейно зависимой или линейно независимой.

Преобразуем систему линейных уравнений $A_1x_1+A_2x_2+$ $+A_3x_3=\theta$ методом Гаусса:

$$\begin{array}{r|r|r|l}
x_1 & x_2 & x_3 & \\
\hline
-20 & -7 & 3 & 0 \\
-15 & -2 & -1 & 0 \\
-4 & -4 & \boxed{-2} & 0
\end{array}
\longrightarrow
\begin{array}{r|r|r|l}
x_1 & x_2 & x_3 & \\
\hline
-26 & -13 & 0 & 0 \\
-13 & 0 & 0 & 0 \\
2 & 2 & 1 & 0
\end{array}
\longrightarrow$$

$$\longrightarrow
\begin{array}{r|r|r|l}
x_1 & x_1 & x_3 & \\
\hline
2 & 1 & 0 & 0 \\
\boxed{-13} & 0 & 0 & 0 \\
-2 & 0 & 1 & 0
\end{array}
\longrightarrow
\begin{array}{r|r|r|l}
x_1 & x_2 & x_3 & \\
\hline
0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{array}$$

Общее решение исходной системы имеет вид $x_1=0$, $x_2=0$, $x_3=0$. Эта система, а следовательно исходная система уравнений, не имеет ненулевых решений. Таким образом, векторы A_1, A_2, A_3 линейно независимы. ●

Если каждый из векторов $B_1, B_2, \ldots, B_n$ разлагается по системе векторов $A_1, A_2, \ldots, A_m$, $m<n$, то система векторов $B_1, B_2, \ldots, B_n$ линейно зависима.

2.10. Базис и ранг системы векторов

Линейно независимая часть $B_1, B_2, \ldots, B_r$ системы векторов $A_1, A_2, \ldots, A_n$ называется *базисом* этой системы, если каждый вектор системы $A_1, A_2, \ldots, A_n$ разлагается по векторам $B_1, B_2, \ldots, B_r$.

Каждую линейно независимую часть системы векторов можно дополнить до базиса этой системы.

Векторы системы $A_1, A_2, \ldots, A_n$ разлагаются по базису этой системы единственным образом.

Для отыскания базиса системы векторов $A_1, A_2, \ldots, A_n$ находят общее решение системы линейных уравнений

$$A_1x_1+A_2x_2+\ldots+A_nx_n=\theta. \qquad (2.11)$$

Тогда векторы-коэффициенты уравнения (2.11) при неизвестных, составляющих набор разрешенных неизвестных общего решения, образуют базис системы векторов $A_1, A_2, \ldots, A_n$.

○ **Пример.** Найти базис системы векторов $A_1=(5,2,-3,1)$, $A_2=(4,1-2,3)$, $A_3=(1,1-1,-2)$, $A_4=(3,4-1,2)$.

Найдем общее решение системы уравнений

$$A_1x_1+A_2x_2+A_3x_3+A_4x_4=\theta \qquad (2.12)$$

Имеем:

x_1	x_2	x_3	x_4	
5	4	$\boxed{1}$	3	0
2	1	1	4	0
—3	—2	—1	—1	0
1	3	—2	2	0

$\longrightarrow$

x_1	x_2	x_3	x_4	
5	4	1	3	0
—3	—3	0	1	0
2	2	0	$\boxed{2}$	0
11	11	0	8	0

$\longrightarrow$

x_1	x_2	x_3	x_4	
0	—1	1	—2	0
0	0	0	$\boxed{4}$	0
1	1	0	1	0
0	0	0	—3	0

$\longrightarrow$

x_1	x_2	x_3	x_4	
0	—1	1	0	0
0	0	0	1	0
1	1	0	0	0
0	0	0	0	0

Из последней таблицы следует, что неизвестные x_1, x_3, x_4 образуют набор разрешенных неизвестных общего решения системы уравнений (2.12). Следовательно, векторы A_1, A_3, A_4 образуют базис системы векторов A_1, A_2, A_3, A_4. ●

Все базисы данной системы векторов состоят из одного и того же числа векторов.

Рангом системы векторов называется число векторов в любом ее базисе.

Если ранг системы векторов $A_1, A_2, \ldots, A_n$ равен r, то каждая линейно независимая часть этой системы, состоящая из r векторов, является ее базисом.

2.11. Условия совместности и определенности системы линейных уравнений

Система линейных уравнений

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

записанная в векторной форме, совместна тогда и только тогда, когда ранги систем векторов $A_1, A_2, \ldots, A_n$ и $A_1, A_2, \ldots, A_n, B$ совпадают.

Совместная система линейных уравнений имеет единственное решение, если ранг системы векторов $A_1, A_2, \ldots, A_n$ равен числу неизвестных в системе. Если же ранг этой системы векторов меньше числа неизвестных, то совместная система уравнений имеет бесконечно много решений.

2.12. Однородные системы линейных уравнений

Система линейных уравнений называется *однородной*, если все свободные члены системы равны нулю. Такая система в векторной форме имеет следующий вид:

$$A_1x_1 + A_2x_2 + \ldots + A_nx_n = \theta.$$

Каждая однородная система линейных уравнений имеет нулевое решение $x_1=x_2=\ldots=x_n=0$ и, значит, совместна.

Всякая однородная система линейных уравнений, у которой число уравнений меньше числа неизвестных, имеет ненулевое решение.

Любое решение $x_1=k_1$, $x_2=k_2$, ..., $x_n=k_n$ системы уравнений с n неизвестными можно рассматривать как n-мерный вектор с координатами k_1, k_2, ..., k_n, а поэтому имеют смысл такие понятия, как линейная комбинация, линейная зависимость и линейная независимость решений. Произвольная линейная комбинация решений однородной системы уравнений является решением этой системы.

Линейно независимые решения F_1, F_2, ..., F_k однородной системы уравнений называются *фундаментальной системой решений*, если каждое решение системы является линейной комбинацией решений F_1, F_2, ..., F_k.

Если ранг r системы векторов A_1, A_2, ..., A_n меньше числа неизвестных n в однородной системе уравнений

$$A_1x_1+A_2x_2+\ldots+A_nx_n=\theta,$$

то эта система уравнений имеет фундаментальную систему решений и любая ее фундаментальная система решений состоит из $n-r$ решений.

Построение фундаментальной системы решений

1. Находят общее решение однородной системы уравнений.

2. Берут систему $n-r$ линейно независимых $(n-r)$-мерных векторов. Например, $e_1=(1, 0, \ldots, 0)$, $e_2=(0, 1, \ldots, 0)$, ..., $e_{n-r}=(0, 0, \ldots, 1)$.

3. Подставляют в общее решение вместо свободных неизвестных координаты вектора e_1, а затем находят значения разрешенных неизвестных. Полученная совокупность значений неизвестных является решением F_1. Аналогично, с помощью векторов e_2, ..., e_{n-r} находят решения F_2, ..., F_{n-r}.

Полученные решения F_1, F_2, ..., F_{n-r} составляют фундаментальную систему решений. Варьируя координаты линейно независимых векторов, получают все фундаментальные системы решений.

○ **Пример.** Найти фундаментальную систему решений однородной системы уравнений

$$\begin{cases} 3x_1-2x_2+2x_3-\ x_4+4x_5=0, \\ 5x_1+\ x_2+4x_3-2x_4+7x_5=0, \\ x_1-5x_2 \qquad\qquad\quad +\ x_5=0, \\ 4x_1-7x_2+2x_3-\ x_4+5x_5=0. \end{cases}$$

Общее решение данной системы имеет вид

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

Выбирая для свободных неизвестных x_2, x_3, x_5 значения, равные координатам векторов $e_1=(1, 0, 0)$, $e_2=(0, 1, 0)$, $e_3=(0, 0, 1)$, найдем фундаментальную систему решений: $F_1=(5, 1, 0, 13, 0)$, $F_2=(0, 0, 1, 2, 0)$, $F_3=(-1, 0, 0, 1, 1)$. ●

2.13. Общее решение системы линейных уравнений в векторной форме

Рассмотрим систему линейных уравнений, записанную в векторной форме:

$$A_1x_1+A_2x_2+\ldots+A_nx_n=B. \tag{2.13}$$

Если в системе (2.13) заменить все свободные члены нулями, то получим однородную систему

$$A_1x_1+A_2x_2+\ldots+A_nx_n=\theta. \tag{2.14}$$

Систему (2.14) называют *приведенной* для исходной системы уравнений (2.15).

Произвольное решение X совместной системы уравнений (2.13) определяется формулой

$$Y=F_0+\lambda_1F_1+\lambda_2F_2+\ldots+\lambda_kF_k, \tag{2.15}$$

где F_0—какое-нибудь решение системы (2.13); F_1, F_2, ..., F_k—фундаментальная система решений системы уравнений (2.14); λ_1 ... λ_k—произвольные действительные числа.

Формула (2.15) называется *общим решением в векторной форме* системы уравнений (2.13).

○ **Пример.** Найти общее решение в векторной форме системы линейных уравнений

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

Общее решение данной системы, найденное методом Гаусса, имеет вид

$$\begin{cases} x_1-(5/2)\,x_2+(7/2)\,x_4=6, \\ \qquad (3/2)\,x_2+x_3-(3/2)\,x_4=-2. \end{cases}$$

Вектор (6, 0, —2, 0) является решением этой системы. Система уравнений

$$x_1-(5/2)\,x_2+(7/2)\,x_4=0,$$
$$(3/2)\,x_2+x_3-(3/2)\,x_4=0$$

является общим решением приведенной системы. Выбирая для свободных неизвестных x_2 и x_4 значения, равные координатам векторов $e_1=(1,\ 0)$, $e_2=(0,\ 1)$, найдем фундаментальную систему решений приведенной системы уравнений: $F_1=(5/2,\ 1,\ -3/2,\ 0)$, $F_2=(-7/2,\ 0,\ 3/2,\ 1)$. Следовательно, общее решение в векторной форме данной системы уравнений имеет вид

$$X=\begin{pmatrix} 6 \\ 0 \\ -2 \\ 0 \end{pmatrix}+\lambda_1\begin{pmatrix} 5/2 \\ 1 \\ -3/2 \\ 0 \end{pmatrix}+\lambda_2\begin{pmatrix} -7/2 \\ 0 \\ 3/2 \\ 1 \end{pmatrix}. \quad ●$$

2.14. Ортогональные системы векторов

Два вектора называются *ортогональными*, если их скалярное произведение равно нулю. Система векторов называется ортогональной, если векторы этой системы попарно ортогональны.

○ **Пример.** Система векторов $e_1=(1,\ 0,\ \ldots,\ 0)$, $e_2=(0,\ 1,\ \ldots,\ 0)$, $\ldots$, $e_n=(0,\ 0,\ \ldots,\ 1)$ ортогональна. ●

Ортогональная система ненулевых векторов линейно независима.

Исходя из линейно независимой системы векторов $x_1,\ \ldots,\ x_{m+1}$ можно построить ортогональную систему ненулевых векторов $y_1,\ \ldots,\ y_{m+1}$ по следующим формулам:

$$y_1=x_1,$$

$$y_2=-\frac{y_1x_2}{y_1y_1}\,y_1+x_2,$$

$$\cdots\cdots\cdots$$

$$y_{m+1}=-\frac{y_1x_{m+1}}{y_1y_1}\,y_1-\frac{y_2x_{m+1}}{y_2y_2}\,y_2-\frac{y_mx_{m+1}}{y_my_m}\,y_m+x_{m+1}.$$

Приведенный способ построения ортогональной системы векторов $y_1,\ y_2,\ \ldots,\ y_{m+1}$ по заданной линейно независимой системе $x_1, x_2, \ldots, x_{m+1}$ называется *процессом ортогонализации системы векторов* $x_1, x_2, \ldots, x_{m+1}$.

○ **Пример.** Построить ортогональную систему векторов путем ортогонализации линейно независимой системы $x_1=(1, 1, 1, 0)$, $x_2=(0, 1, 1, 1)$, $x_3=(0, 0, 1, 1)$.

○ Строим систему векторов y_1, y_2, y_3:

$$y_1=x_1=(1, 1, 1, 0),$$

$$y_2=-\frac{y_1x_2}{y_1y_1}y_1+x_2=(-2/3)(1, 1, 1, 0)+(0, 1, 1, 1)=$$
$$=(-2/3, 1/3, 1/3, 1),$$

$$y_3=-\frac{y_1x_3}{y_1y_1}y_1-\frac{y_2x_3}{y_2y_2}y_2+x_3=(-1/3)(1, 1, 1, 0)-(4/5)\times$$
$$\times(-2/3, 1/3, 1/3, 1)+(0, 0, 1, 1)=(1/5, -3/5, 2/5, 1/5). \text{ ●}$$

Система векторов называется *ортонормированной*, если векторы этой системы попарно ортогональны и имеют длину, равную единице. Если $x_1, x_2, \ldots, x_n$ — ортогональная система ненулевых векторов, то $\frac{1}{|x_1|}x_1, \frac{1}{|x_2|}x_2, \ldots, \frac{1}{|x_n|}x_n$ — ортонормированная система векторов.

2.15. Матрицы

Прямоугольная таблица чисел

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

состоящая из m строк и n столбцов, называется матрицей размера $m\times n$. Числа $a_{11}, a_{12}, \ldots, a_{mn}$ называются ее *элементами*. Часто вместо подробной записи используют сокращенную: $A=(a_{ij})$.

Если число строк матрицы равно числу ее столбцов, то матрица называется *квадратной*, а число ее строк, равное числу столбцов, — *порядком* квадратной матрицы.

Множество всех элементов квадратной матрицы, которые лежат на отрезке, соединяющем левый верхний угол с правым нижним, называется *главной диагональю*, а на отрезке, соединяющем правый верхний угол с левым нижним, — *побочной диагональю*.

Квадратная матрица называется *диагональной*, если все ее элементы, не лежащие на главной диагонали, равны нулю. Диагональная матрица обозначается символом

$\{a_1, a_2, \ldots, a_n\}$, где в скобках указаны элементы, находящиеся на главной диагонали.

Две матрицы $A=(a_{ij})$ и $B=(b_{ij})$ называются *равными*, если число их строк и столбцов равны и если равны элементы, стоящие на соответственных местах этих матриц: $a_{ij}=b_{ij}$ при любых i и j.

2.16. Умножение матрицы на число и сложение матриц

По определению, чтобы умножить матрицу A на число k, нужно каждый элемент матрицы A умножить на k.

Например,

$$\begin{pmatrix} 3 & 1 & -2 & 4 \\ 7 & 0 & 3 & 21 \\ -1 & 2 & 1 & 7 \end{pmatrix} \cdot 3 = \begin{pmatrix} 9 & 3 & -6 & 12 \\ 21 & 0 & 9 & 63 \\ -3 & 6 & 3 & 21 \end{pmatrix}.$$

Складывать можно только матрицы с одинаковым числом строк и столбцов. Суммой матриц $A=(a_{ij})$ и $B=(b_{ij})$ называется матрица $C=(c_{ij})$, элементы которой равны суммам соответственных элементов матриц A и B: $c_{ij}=a_{ij}+b_{ij}$ при любых i и j.

Например,

$$\begin{pmatrix} 3 & 2 & 1 \\ 8 & -1 & 5 \end{pmatrix} + \begin{pmatrix} 2 & 5 & 7 \\ 7 & 0 & -40 \end{pmatrix} = \begin{pmatrix} 5 & 7 & 8 \\ 15 & -1 & -35 \end{pmatrix}.$$

Матрица, все элементы которой равны нулю, называется *нулевой* и обозначается через $\mathbf{0}$. Для любой матрицы A имеем $A+\mathbf{0}=A$.

Матрица $A(-1)$ называется *противоположной* A и обозначается через $-A$. Вместо $A+(-B)$ пишут $A-B$.

Свойства умножения матрицы на число и сложения матриц (A, B, C — матрицы, k, l — числа)

1°. $A(kl)=(Ak)l$.

2°. $A+B=B+A$.

3°. $(A+B)+C=A+(B+C)$.

4°. $A(k+l)=Ak+Al$.

5°. $(A+B)k=Ak+Bk$.

2.17. Умножение матриц

Произведение матрицы A на матрицу B определено только в том случае, когда число столбцов матрицы A равно числу строк матрицы B. В результате умножения получится матрица AB, у которой столько же строк, сколько их в матрице A, и столько же столбцов, сколько их в матрице B:

	A	B	AB
число строк	m	n	m
число столбцов	n	l	l

Запишем матрицы A и B в виде

$$A=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & b_{12} & \dots & b_{1j} & \dots & b_{1l} \\ b_{21} & b_{22} & \dots & b_{2j} & \dots & b_{2l} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{n1} & b_{n2} & \dots & b_{nj} & \dots & b_{nl} \end{pmatrix}.$$

Обозначим элементы матрицы AB через c_{ij}, $1\leqslant i\leqslant m$, $1\leqslant j\leqslant l$. Тогда

$$AB=\begin{pmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1l} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{i1} & \dots & c_{ij} & \dots & c_{il} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{m1} & \dots & c_{mj} & \dots & c_{ml} \end{pmatrix}.$$

По определению, элемент c_{ij} матрицы AB равен скалярному произведению i-й строки матрицы A (i — первый индекс элемента c_{ij}) на j-й столбец матрицы B (j — второй индекс элемента c_{ij}), т. е.

$$c_{ij}=(a_{i1},\ a_{i2},\ \dots,\ a_{in})(b_{1j},\ b_{2j}\ \dots,\ b_{nj})=$$
$$=a_{i1}b_{1j}+a_{i2}b_{2j}+\dots+a_{in}b_{nj}.$$

○ **Пример.** Найти произведение AB, если

$$A=\begin{pmatrix} 2 & 3 & 4 & 5 \\ 9 & 2 & -3 & 4 \\ -1 & -5 & 3 & 11 \end{pmatrix}, \quad B=\begin{pmatrix} 3 & 2 \\ 4 & -1 \\ 1 & -3 \\ 2 & 5 \end{pmatrix}.$$

Матрица AB является матрицей размера 3×2. Вычисляем элементы c_{ij} матрицы AB. Имеем:

$c_{11}=(2,\ 3,\ 4,\ 5)(3,\ 4,\ 1,\ 2)=2\cdot3+3\cdot4+4\cdot1+5\cdot2=32;$

$c_{12}=(2,\ 3,\ 4,\ 5)(2,\ -1,\ -3,5)=2\cdot2+3\cdot(-1)+4\cdot(-3)+$ $+5\cdot5=14;$

$c_{21}=(9,\ 2,\ -3,\ 4)(3,\ 4,\ 1,\ 2)=9\cdot3+2\cdot4+(-3)\cdot1+4\cdot2=$ $=40;$

$c_{22}=(9,2-3,4)(2,-1,-3,5)=9\cdot2+2\cdot(-1)+(-3)\times$ $\times(-3)+4\cdot5=45;$

$c_{31}=(-1,\ -5,\ 3,\ 11)(3,\ 4,\ 1,\ 2)=(-1)\cdot3+(-5)4+$ $+3\cdot1+11\cdot2=2;$

$c_{32}=(-1,\ -5,\ 3,\ 11)(2,\ -1,\ -3,5)=(-1)\cdot2+(-5)\times$ $\times(-1)+3(-3)+11\cdot5=49.$

Итак, $AB = \begin{pmatrix} 32 & 14 \\ 40 & 45 \\ 2 & 49 \end{pmatrix}$. ●

Свойства умножения матриц

1°. $(AB)k = (Ak)B = A(Bk)$, k — число.
2°. $(A+B)C = AC + BC$.
3°. $C(A+B) = CA + CB$.
4°. $(AB)C = A(BC)$.

Произведение матриц зависит от порядка множителей. Если

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ то } AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Матрицы A, B называются *перестановочными*, если $AB = BA$.

2.18. Блочные матрицы и действия с ними

Пусть некоторая матрица A разбита на клетки горизонтальными и вертикальными прямыми. Например, матрица

$$A = \left(\begin{array}{cc|ccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{array}\right)$$

разбита на четыре клетки. Каждая клетка является матрицей. Обозначим клетки матрицы A через $A_{11}, A_{12}, A_{21}, A_{22}$, где

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}, \quad A_{12} = \begin{pmatrix} a_{13} & a_{14} & a_{15} \\ a_{23} & a_{24} & a_{25} \\ a_{33} & a_{34} & a_{35} \end{pmatrix},$$

$$A_{21} = (a_{41}, \ a_{42}), \quad A_{22} = (a_{43}, \ a_{44}, \ a_{45}).$$

Теперь матрицу A можно записать в виде

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

Матрица, которая некоторым образом разбита на клетки, называется *блочной* или *клеточной*. Каждую матрицу можно представить в блочной форме разными способами.

При умножении блочной матрицы на число следует все ее клетки умножить на это число.

Чтобы сложить две матрицы одинакового размера и одинаковым образом разбитых на клетки, достаточно сложить одноименные клетки этих матриц, т. е.

$$\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1n} \\ B_{21} & B_{22} & \dots & B_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ B_{m1} & B_{m2} & \dots & B_{mn} \end{pmatrix} =$$

$$= \begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} & \dots & A_{1n}+B_{1n} \\ A_{21}+B_{21} & A_{22}+B_{22} & \dots & A_{2n}+B_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ A_{m1}+B_{m1} & A_{m2}+B_{m2} & \dots & A_{mn}+B_{mn} \end{pmatrix}.$$

Пусть теперь даны матрица A размера $s\times t$ и матрица B размера $t\times l$, причем

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \cdot & \cdot & \cdot \\ A_{m1} & \dots & A_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & \dots & B_{1p} \\ \cdot & \cdot & \cdot \\ B_{n1} & \dots & B_{np} \end{pmatrix}$$

и число столбцов клетки A_{ij} равно числу строк клетки B_{jk} при всех $i=1, \dots, m$; $j=1, \dots, n$; $k=1, \dots, p$. Тогда

$$AB = \begin{pmatrix} c_{11} & \dots & c_{1p} \\ \cdot & \cdot & \cdot \\ c_{m1} & \dots & c_{mp} \end{pmatrix},$$

где $c_{ik} = A_{i1}B_{1k} + A_{i2}B_{2k} + \dots + A_{in}B_{nk}$.

2.19. Умножение матрицы на вектор

Каждый вектор можно рассматривать как одностолбцовую или однострочную матрицу. Одностолбцовую матрицу будем называть *вектор-столбцом*, а однострочную матрицу — *вектор-строкой*.

Если A — матрица размера $m\times n$, вектор-столбец x имеет размерность n, а вектор-строка y — размерность m, то определены произведения Ax и yA, причем Ax — вектор-столбец размерности m, а yA — вектор-строка размерности m.

Таким образом, чтобы умножить матрицу на вектор, надо рассматривать вектор как вектор-столбец. При умножении вектора на матрицу его нужно рассматривать как вектор-строку.

○ **Пример.** Даны матрица A и векторы x и y:

$$A=\begin{pmatrix}1 & 2 & -1 & 2\\ 3 & 0 & 4 & -2\\ 2 & 3 & 3 & 5\end{pmatrix},\qquad x=\begin{pmatrix}-1\\ 2\\ 1\\ 3\end{pmatrix},\quad y=(2,\ 1,\ -3).$$

Вычислить координаты векторов Ax и yA.

Имеем

$$Ax=\begin{pmatrix}1 & 2 & -1 & 2\\ 3 & 0 & 4 & -2\\ 2 & 3 & 3 & 5\end{pmatrix}\begin{pmatrix}-1\\ 2\\ 1\\ 3\end{pmatrix}=\begin{pmatrix}8\\ -5\\ 22\end{pmatrix},$$

$$yA=(2,\ 1,\ -3)\begin{pmatrix}1 & 2 & -1 & 2\\ 3 & 0 & 4 & -2\\ 2 & 3 & 3 & 5\end{pmatrix}=(-1,\ -5,\ -7,\ -13). \ ●$$

Свойства умножения матрицы на вектор (λ — число, A — матрица, x_1, x_2, x, y_1, y_2, y — векторы)

1°. $A(x_1+x_2)=Ax_1+Ax_2$. 2°. $A(\lambda x)=\lambda(Ax)$.

3°. $(y_1+y_2)A=y_1A+y_2A$. 4°. $(\lambda y)A=\lambda(yA)$.

5°. $y(Ax)=(yA)x$.

2.20. Матрично-векторная форма записи системы линейных уравнений

Рассмотрим систему линейных уравнений

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

и введем следующие обозначения:

$$A=\begin{pmatrix}a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ a_{m1} & a_{m2} & \ldots & a_{mn}\end{pmatrix},\qquad x=\begin{pmatrix}x_1\\ x_2\\ \cdot\\ \cdot\\ \cdot\\ x_n\end{pmatrix},\qquad b=\begin{pmatrix}b_1\\ b_2\\ \cdot\\ \cdot\\ \cdot\\ b_m\end{pmatrix}.$$

Матрицу A называют матрицей системы линейных уравнений; x — вектор-столбец неизвестных, а b — вектор-столбец свободных членов.

Так как столбцов у матрицы A ровно столько, сколько координат у вектор-столбца x, то определено произведение

$$Ax=\begin{pmatrix} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n \end{pmatrix}.$$

Теперь систему линейных уравнений можно записать в виде одного векторного равенства $Ax=b$.

2.21. Обратная матрица

Квадратная матрица

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

называется *единичной* и обозначается через E.

Квадратная матрица A называется *обратимой*, если можно подобрать такую матрицу B, что $AB=BA=E$.

Матрица B называется *обратной* для матрицы A.

Матрица называется *невырожденной*, если ее столбцы линейно независимы.

Квадратная матрица обратима тогда и только тогда, когда она невырожденная.

Обратимая матрица имеет только одну обратную матрицу, которую обозначают через A^{-1}.

Квадратная матрица A порядка n обратима тогда и только тогда, когда каждая из n систем линейных уравнений $AX=E^1$, $AX=E^2$, $\ldots$, $AX=E^n$ имеет единственное решение, где E^1, E^2, $\ldots$, E^n—столбцы единичной матрицы, а $X=\begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$—вектор-столбец, координатами которого являются неизвестные x_1, x_2, $\ldots$, x_n.

Если матрица A обратима, то единственное решение системы уравнений $AX=E^i$, $i=1, 2, \ldots, n$, совпадает с i-м столбцом матрицы A^{-1}.

Для определения элементов матрицы A^{-1} необходимо решить n систем линейных уравнений с n неизвестными. Так как эти системы отличаются только набором свободных членов, то их можно решать параллельно в одной таблице.

○ **Пример.** Найти обратную матрицу A^{-1}, если

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

Решаем параллельно системы уравнений

x_1	x_2	x_3	E^1	E^2	E^3
$\boxed{1}$	0	—1	1	0	0
2	3	2	0	1	0
—1	1	2	0	0	1

$\longrightarrow$

x_1	x_2	x_3	E^1	E^2	E^3
1	0	—1	1	0	0
0	3	4	—2	1	0
0	$\boxed{1}$	1	1	0	1

$\longrightarrow$

x_1	x_2	x_3			
1	0	—1	1	0	0
0	0	$\boxed{1}$	—5	1	—3
0	1	1	1	0	1

$\longrightarrow$

x_1	x_2	x_3			
1	0	0	—4	1	—3
0	0	1	—5	1	—3
0	1	0	6	—1	4

Из последней таблицы находим:

единственное решение системы уравнений $AX=E^1$

$$x_1=-4,\ x_2=6,\ x_3=-5;$$

единственное решение системы уравнений $AX=E^2$

$$x_1=1,\ x_2=-1,\ x_3=1;$$

единственное решение системы уравнений $AX=E^3$

$$x_1=-3,\ x_2=4,\ x_3=-3.$$

Таким образом

$$A^{-1}=\begin{pmatrix} -4 & 1 & -3 \\ 6 & -1 & 4 \\ -5 & 1 & -3 \end{pmatrix}. \quad \bullet$$

Свойства обратной матрицы

$$(A^{-1})^{-1}=A,\ (AB)^{-1}=B^{-1}A^{-1}.$$

2.22. Транспонирование матрицы

Наряду с матрицей A часто приходится рассматривать матрицу, столбцами которой являются строки матрицы A. Эту матрицу называют *транспонированной* к A и обозначают через A' или A^{T}.

Пример. Транспонированной к матрице

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

является матрица

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

Свойства операции транспонирования (k — число)

1°. $(Ak)^{\text{т}}=kA^{\text{т}}$.

2°. $(A+B)^{\text{т}}=A^{\text{т}}+B^{\text{т}}$.

3°. $(AB)^{\text{т}}=B^{\text{т}}A^{\text{т}}$.

4°. $(A^{\text{т}})^{\text{т}}=A$.

Если A — обратимая матрица, то

$$(A^{-1})^{\text{т}}=(A^{\text{т}})^{-1}.$$

2.23. Ранг матрицы

Ранг системы векторов — строк матрицы A — равен рангу системы ее вектор-столбцов. Число, равное рангу системы строк (или столбцов) матрицы, называется *рангом* этой матрицы.

Ранг матрицы не изменяется при транспонировании.

Если обозначить ранг матрицы A через $r(A)$, то

$$r(AB)\leqslant r(A),\quad r(AB)\leqslant r(A).$$

Если же матрица B обратима, то

$$r(AB)=r(A),\quad r(BA)=r(A).$$

Для ранга произведения матриц A и B справедливо неравенство

$$r(A)+r(B)-n\leqslant r(AB),$$

где n — число столбцов матрицы A и число строк матрицы B.

2.24. Симметрические и ортогональные матрицы

Квадратная матрица A называется *симметрической*, если $A=A^{\text{т}}$. Если же $A=-A^{\text{т}}$, то матрица A называется *кососимметрической*. Элементы a_{ik} и a_{ki}, расположенные симметрично относительно главной диагонали, у симметрической матрицы равны, а у кососимметрической — противоположны.

Если $A^{\mathrm{T}} = A^{-1}$, то квадратная матрица A называется ортогональной. Матрица является ортогональной тогда и только тогда, когда ее строки или столбцы образуют ортонормированную систему векторов.

2.25. Определители квадратных матриц

Назовем произведение n элементов квадратной матрицы *правильным*, если эти элементы расположены в ее различных строках и различных столбцах, т. е. по одному в каждой строке и каждом столбце.

Если

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

то произведение $a_{11}a_{22}\dots a_{nn}$ является правильным.

Каждое правильное произведение можно записать в виде

$$a_{\alpha_1 1}a_{\alpha_2 2}\dots a_{\alpha_n n}, \tag{2.16}$$

т. е. первый множитель содержится в первом столбце, второй — во втором столбце и т. д. Числа $\alpha_1, \alpha_2, \dots, \alpha_n$ — это номера строк, в которых расположены множители правильного произведения (2.16).

Назовем *инверсией* в последовательности $\alpha_1, \alpha_2, \dots, \alpha_n$ такое расположение индексов, когда больший индекс стоит левее меньшего. Число всех инверсий в последовательности $\alpha_1, \alpha_2, \dots, \alpha_n$ обозначим через $N(\alpha_1, \alpha_2, \dots, \alpha_n)$.

○ **Пример.** В последовательности 2, 4, 1, 3 имеется три инверсии (2 находится левее 1, 4 — левее 1, 4 — левее 3). Таким образом, $N(2, 4, 1, 3) = 3$. ●

Перед каждым правильным произведением вида (2.16) будем писать знак, определяемый выражением $(-1)^{N(\alpha_1, \alpha_2, \dots, \alpha_n)}$.

Определителем матрицы A называется алгебраическая сумма всех правильных произведений этой матрицы, имеющих знак плюс или минус в соответствии с приведенным выше правилом. Определитель матрицы A обозначают $\det A$ или $|A|$.

Применим это определение к матрицам второго и третьего порядка. Из элементов матрицы $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ можно составить только два правильных произведения: $a_{11}a_{22}$ и $a_{21}a_{12}$, причем первому из них приписывается знак плюс, а второму — знак минус.

Следовательно,

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

Правильные произведения матрицы

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

исчерпываются произведениями

$$a_{11}a_{22}a_{33}, \quad a_{31}a_{12}a_{23}, \quad a_{21}a_{32}a_{13}, \tag{2.17}$$

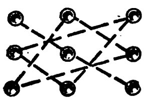
$$a_{31}a_{22}a_{13}, \quad a_{21}a_{12}a_{33}, \quad a_{11}a_{32}a_{23}, \tag{2.18}$$

причем произведениям (2.17) приписывается знак плюс, а произведению (2.18) — знак минус. Таким образом,

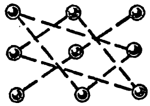
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{31}a_{12}a_{23} + a_{21}a_{32}a_{13} - \\ - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}. \tag{2.19}$$

Знаки, которые приписываются правильным произведениям в (2.19), можно запомнить следующим образом.

Соединим пунктирной линией каждые три элемента матрицы, произведение которых входит в (2.19) со знаком плюс. Тогда получим следующую легко запоминающуюся схему:



Аналогично, для произведений, входящих со знаком минус, имеем



2.26. Разложение определителя по строке и столбцу

Рассмотрим алгебраическую сумму всех правильных произведений матрицы A, содержащих множителем элемент a_{ik}, вынесем этот общий множитель за скобки и выражение, оставшееся в скобках, обозначим через A_{ik}. Выражение A_{ik} называется *алгебраическим дополнением* элемента a_{ik} в определителе матрицы A.

Вычеркнем в матрице A i-ю строку и j-й столбец. Определитель полученной матрицы $(n-1)$-го порядка называют *минором* элемента a_{ij} в определителе матрицы A и обозначают через M_{ij}.

Алгебраическое дополнение A_{ij} равно соответствующему минору M_{ij}, умноженному на $(-1)^{i+j}$, т. е.

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

Справедливы следующие равенства:

$$\det A=(-1)^{i+1} a_{i1}M_{i1}+\ldots+(-1)^{i+j} a_{ij}M_{ij}+\ldots \\ \ldots+(-1)^{i+n} a_{in}M_{in}, \tag{2.20}$$

$$\det A=(-1)^{1+j} a_{1j}M_{1j}+\ldots+(-1)^{i+j} a_{ij}M_{ij}+\ldots \\ \ldots+(-1)^{n+j} a_{nj}M_{nj}. \tag{2.21}$$

Равенство (2.20) называется разложением определителя матрицы A по элементам i-строки, а равенство (2.21) — разложением по элементам j-го столбца.

Формулы (2.20) и (2.21) можно использовать для вычисления определителей матриц.

○ **Пример.** Вычислить определитель матрицы $\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 3 & 4 \\ 1 & 1 & 0 & 2 \\ 5 & 2 & 1 & 0 \end{pmatrix}$.

Разлагая определитель по элементам третьего столбца, получаем

$$\begin{vmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 3 & 4 \\ 1 & 1 & 0 & 2 \\ 5 & 2 & 1 & 0 \end{vmatrix}=(-1)^{2+3}\cdot 3\begin{vmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \\ 5 & 2 & 0 \end{vmatrix}+(-1)^{4+3}\cdot 1\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 4 \\ 1 & 1 & 2 \end{vmatrix}=$$

$$=(-3)\cdot 13-1(-1)=-38. \bullet$$

2.27. Свойства определителей. Вычисление определителей

1°. Определитель матрицы не изменяется при транспонировании.

2°. Общий множитель элементов любой строки (столбца) можно выносить за знак определителя этой матрицы, т. е.

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ ka_{i1} & ka_{i2} & \ldots & ka_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}=k\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{i1} & a_{i2} & \ldots & a_{in} \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

3°. Если все элементы i-й строки матрицы n-го порядка представлены в виде суммы двух слагаемых $a_{ij} = b_j + c_j$, $j = 1, 2, \ldots, n$, то определитель этой матрицы равен сумме определителей матриц, у которых все строки, кроме i-й, такие же, как и в данной матрице, а i-я строка у одной из матриц состоит из элементов b_j, а у другой — из элементов c_j, т. е.

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ b_1+c_1 & b_2+c_2 & \ldots & b_n+c_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ b_1 & b_2 & \ldots & b_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ c_1 & c_2 & \ldots & c_n \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

Аналогичное свойство справедливо и в том случае, когда элементы некоторого столбца матрицы представлены в виде суммы двух слагаемых.

4°. Определитель матрицы, имеющей две одинаковые строки (столбца), равен нулю.

5°. Определитель матрицы не изменится, если к i-й строке (столбцу) матрицы A прибавить ее j-ю строку (столбец), умноженную на число.

Если в матрице порядка n имеется строка (столбец), все элементы которой равны нулю, кроме одного, то вычисление определителя матрицы n-го порядка сводится к вычислению единственного определителя матрицы порядка $(n-1)$.

Используя свойство 5° определителей матриц, можно, не изменяя величины определителя, преобразовать данную матрицу так, чтобы в выбранной строке (столбце) все элементы, кроме одного, обратились в нуль.

○ **Пример.** Вычислить определитель матрицы

$$A = \begin{pmatrix} -2 & -5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 18 & -7 & -10 \end{pmatrix}.$$

Прибавляя к первой строке удвоенную вторую, к третьей — вторую, умноженную на -3, а к четвертой строке — вторую, умноженную на —2, имеем

$$\begin{vmatrix} -2 & 5 & -1 & 3 \\ 1 & -9 & 13 & 7 \\ 3 & -1 & 5 & -5 \\ 2 & 18 & -7 & -10 \end{vmatrix} = \begin{vmatrix} 0 & -13 & 25 & 17 \\ 1 & -9 & 13 & 7 \\ 0 & 26 & -34 & -26 \\ 0 & 36 & -33 & -24 \end{vmatrix} =$$

$$= (-1)^{2+1} a_{21} M_{21} = -\begin{vmatrix} -13 & 25 & 17 \\ 26 & -34 & -26 \\ 36 & -33 & -24 \end{vmatrix}.$$

Получен определитель матрицы третьего порядка, который можно вычислить либо непосредственно, либо сведя его к вычислению определителя матрицы второго порядка. Имеем

$$\begin{vmatrix} -13 & 25 & 17 \\ 26 & -34 & -26 \\ 36 & -33 & -24 \end{vmatrix} = 2\begin{vmatrix} -13 & 25 & 17 \\ 13 & -17 & -13 \\ 36 & -33 & -24 \end{vmatrix} =$$

$$= 2\begin{vmatrix} -13 & 25 & 17 \\ 13 & -17 & -13 \\ 10 & 1 & 2 \end{vmatrix} = 2\begin{vmatrix} 0 & 8 & 4 \\ 13 & -17 & -13 \\ 10 & 1 & 2 \end{vmatrix} =$$

$$= 2\cdot 4\begin{vmatrix} 0 & 2 & 1 \\ 13 & -17 & -13 \\ 10 & 1 & 2 \end{vmatrix} = 8\begin{vmatrix} 0 & 0 & 1 \\ 13 & 9 & -13 \\ 10 & -3 & 2 \end{vmatrix} =$$

$$= 8(-39-90) = -1032.$$

Итак, $|A| = -1032$. ●

6°. Определитель матрицы A равен нулю тогда и только тогда, когда столбцы или строки матрицы A линейно зависимы.

7°. Определитель произведения матриц равен произведению их определителей, т. е.

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

2.28. Системы линейных уравнений с квадратной матрицей

Рассмотрим систему линейных уравнений, записанную в векторно-матричной форме:

$$Ax = b, \tag{2.22}$$

где A — квадратная матрица.

Если определитель матрицы A отличен от нуля, то система уравнений (2.22) имеет единственное решение, которое находят по формулам Крамера

$$x_1 = d_1/d, \; x_2 = d_2/d, \; \ldots, \; x_n = d_n/d,$$

где определитель d_j получен из определителя $d=|A|$ заменой j-го столбца на столбец b свободных членов системы уравнений.

○ **Пример.** Решить систему уравнений $Ax=b$, где

$$A=\begin{pmatrix}1 & -1\\ 1 & 2\end{pmatrix},\ x=\begin{pmatrix}x_1\\ x_2\end{pmatrix},\ b=\begin{pmatrix}3\\ 6\end{pmatrix}.$$

Определитель матрицы системы

$$d=|A|=\begin{vmatrix}1 & -1\\ 1 & 2\end{vmatrix}=3\neq 0$$

и, значит, можно найти решение системы по правилу Крамера. Имеем

$$d_1=\begin{vmatrix}3 & -1\\ 6 & 2\end{vmatrix}=12,\ d_2=\begin{vmatrix}1 & 3\\ 1 & 6\end{vmatrix}=3.$$

Отсюда $x_1=d_1/d=4$, $x_2=d_2/d=1$. ●

Если $|A|\neq 0$, то матрица A обратима. Умножая обе части уравнения (2.22) слева на матрицу A^{-1}, получаем

$$x=A^{-1}b. \qquad (2.23)$$

Формула (2.23) представляет собой матрично-векторную форму записи формул Крамера.

2.29. Собственные векторы и собственные значения матрицы

Число λ называется *собственным значением* (или *характеристическим числом*) квадратной матрицы A порядка n, если можно подобрать такой n-мерный ненулевой вектор x, что $Ax=\lambda x$.

Множество всех собственных значений матрицы A совпадает с множеством всех решений уравнения $|A-\lambda E|=0$, где λ—независимая переменная. Если раскрыть определитель $|A-\lambda E|$, то получится многочлен n-й степени относительно λ:

$$|A-\lambda E|=\begin{vmatrix}a_{11}-\lambda & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22}-\lambda & \ldots & a_{2n}\\ \cdot & \cdot & \cdot & \cdot\\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda\end{vmatrix}=$$
$$=a_n\lambda^n+a_{n-1}\lambda^{n-1}+\ldots+a_1\lambda+a_0.$$

Этот многочлен называется *характеристическим многочленом* матрицы A. Его коэффициенты $a_n, a_{n-1}, \ldots, a_0$ зависят от элементов матрицы A. Отметим, что $a_n=(-1)^n$,

$a_0 = |A|$. Уравнение $|A - \lambda E| = 0$ называется *характеристическим уравнением* матрицы A.

Ненулевой вектор x называется *собственным вектором* квадратной матрицы A, принадлежащим ее собственному значению λ, если $Ax = \lambda x$.

Множество всех собственных векторов матрицы A, принадлежащих ее собственному значению λ, совпадает с множеством всех ненулевых решений системы однородных уравнений $(A - \lambda E)x = 0$, записанной в векторно-матричной форме.

○ **Пример.** Найти собственные значения и собственные векторы матрицы

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

Запишем характеристическое уравнение матрицы

$$|A - \lambda E| = \begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} = \lambda^2 - 5\lambda + 6 = 0.$$

Его корни $\lambda_1 = 2$, $\lambda_2 = 3$ являются собственными значениями матрицы A. Найдем собственные векторы, принадлежащие найденным собственным значениям. Собственный вектор, принадлежащий собственному значению $\lambda_1 = 2$, является ненулевым решением системы

$$(A - 2E)x = \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0,$$

или

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

Теперь $x_1 = 2$ $x_2 = 1$ — ненулевое решение и, значит, $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ — искомый собственный вектор.

Аналогично находим собственный вектор $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ матрицы A, принадлежащий собственному значению $\lambda_2 = 3$. ●

Число различных собственных значений квадратной матрицы не превышает ее порядка.

Собственные векторы квадратной матрицы, принадлежащие ее различным собственным значениям, линейно независимы.

Ортогональная матрица может не иметь действительных собственных значений.

Симметрическая матрица всегда имеет действительное собственное значение.

Собственные векторы симметрической матрицы, принадлежащие различным собственным значениям, ортогональны.

2.30. Приведение квадратной матрицы к диагональному виду

Матрица A называется *подобной* матрице B, если найдется такая невырожденная матрица T, что $B = T^{-1}AT$. Характеристические многочлены подобных матриц совпадают и, значит, подобные матрицы имеют одни и те же собственные значения.

Если матрица A подобна диагональной матрице $B = T^{-1}AT$, то говорят, что матрица T *приводит матрицу A к диагональному виду*. Числа $\lambda_1, \lambda_2, \ldots, \lambda_n$, стоящие на главной диагонали матрицы B, являются собственными значениями матрицы A, а i-й столбец матрицы T — собственным вектором матрицы A, принадлежащим собственному значению λ_i, $i = 1, 2, \ldots, n$.

Квадратная матрица A порядка n тогда и только тогда приводится к диагональному виду, когда у матрицы A имеется n линейно независимых собственных векторов. Матрица T, столбцами которой служат координаты этих собственных векторов, приводит матрицу A к диагональному виду. Этот критерий, в частности, выполняется, когда у матрицы порядка n имеется n различных собственных значений.

Для каждой матрицы A можно построить такую матрицу B, у которой все собственные значения различны, а ее элементы отличаются по абсолютной величине от элементов матрицы A не более чем на ε, где ε — наперед заданное сколь угодно малое положительное число.

Правило построения матрицы T, приводящей матрицу A порядка n к диагональному виду B

1) Находят все собственные значения матрицы A.

2) Для каждого собственного значения λ_i ищут фундаментальную систему решений однородной системы линейных уравнений $(A - \lambda_i E)x = 0$.

3) Строят матрицу T, столбцами которой являются координаты решений всех найденных фундаментальных систем.

4) Если полученная матрица T является квадратной, то она приводит матрицу A к диагональному виду. Если

же матрица T не будет квадратной, то матрица A не может быть приведена к диагональному виду.

○ **Пример.** Выяснить, приводится ли к диагональному виду матрица

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

Вычислим характеристический многочлен матрицы:

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

Сначала из третьего столбца вычтем второй, а затем к третьей строке прибавим вторую:

$$|A-\lambda E|=\begin{vmatrix} 2-\lambda & -1 & 0 \\ -3 & 2-\lambda & \lambda-2 \\ 4 & 2 & 2-\lambda \end{vmatrix}=(\lambda-2)\begin{vmatrix} 2-\lambda & -1 & 0 \\ -3 & 2-\lambda & 1 \\ 4 & 2 & -1 \end{vmatrix}=$$

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

$$=(2-\lambda)(\lambda-3)^2.$$

Собственные значения матрицы A равны 2 и 3.

Теперь надо найти фундаментальные системы решений систем уравнений $(A-2E)x=0$ и $(A-3E)x=0$. Фундаментальная система решений первой системы состоит из одного решения $(0, -1, 1)$, а второй — из одного решения $(1, -3, 2)$. Следовательно, матрица T имеет вид

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

Эта матрица не является квадратной, поэтому матрица A не приводится к диагональному виду. ●

Для каждой симметрической матрицы существует такая ортогональная матрица Q, что $Q^{-1}AQ$ — диагональная матрица. Построение этой ортогональной матрицы осуществляется следующим образом:

1) строят невырожденную матрицу T, которая приводит матрицу A к диагональному виду;

2) подвергают столбцы найденной матрицы T процессу ортогонализации, а затем нормируют полученные векторы;

3) строят ортогональную матрицу Q, столбцами которой являются координаты полученной ортонормированной системы векторов.

2.31. Положительные матрицы

Вектор называется *положительным*, если все его координаты положительны.

Матрица называется *положительной*, если все ее элементы положительны.

Свойства собственных значений и собственных векторов положительной матрицы A

1°. Имеется такое собственное значение $\lambda^* > 0$ матрицы A, что $\lambda^* > |\lambda|$ для любого собственного значения.

2°. Для всех $\mu > \lambda^*$ матрица $\mu E - A$ невырожденна, а матрица $(\mu E - A)^{-1}$ положительна.

3°. Собственный вектор x^*, принадлежащий собственному значению λ^*, положителен.

4°. Кроме x^*, не существует собственных векторов с неотрицательными координатами.

2.32. Квадратичные формы

Переход от системы n неизвестных $x_1, x_2, \ldots, x_n$ к системе n неизвестных $y_1, y_2, \ldots, y_n$ по формулам

$$\begin{aligned} x_1 &= s_{11}y_1 + s_{12}y_2 + \ldots + s_{1n}y_n, \\ x_2 &= s_{21}y_1 + s_{22}y_2 + \ldots + s_{2n}y_n, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \quad (s_{ij} \text{ — числа при всех } i,\ j) \\ x_n &= s_{n1}y_1 + s_{n2}y_n + \ldots + s_{nn}y_n, \end{aligned}$$

называется линейным преобразованием неизвестных, которое в векторно-матричной форме имеет следующий вид:

$$x = Sy,$$

$$x = (x_1,\ x_2,\ \ldots,\ x_n),\ y = (y_1,\ y_2,\ \ldots,\ y_n),$$

$$S = \begin{pmatrix} s_{11} & s_{12} & \ldots & s_{1n} \\ s_{21} & s_{22} & \ldots & s_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ s_{n1} & s_{n2} & \ldots & s_{nn} \end{pmatrix}.$$

Матрица S называется матрицей линейного преобразования неизвестных. Если S невырожденна, то линейное преобразование неизвестных также называется невырожденным.

Квадратичной формой $Q(x_1, x_2, \ldots, x_n)$ от n неизвестных $x_1, x_2, \ldots, x_n$ называется сумма, каждое слагаемое которой является или квадратом одного из этих неизвестных, или произведением двух разных неизвестных.

Обозначим коэффициент при x_i^2 через a_{ii}, а коэффициент при произведении $x_i x_k = x_k x_i$ $(i \neq k)$ — через $a_{ik} + a_{ki}$, причем $a_{ik}=a_{ki}$. Член $(a_{ik}+a_{ki})x_ix_k$ запишем в виде $a_{ik}x_ix_k+a_{ki}x_kx_i$. Теперь квадратичную форму Q можно представить в следующем виде:

$$Q(x_1, x_2, \ldots, x_n)=a_{11}x_1^2+a_{12}x_1x_2+\ldots+a_{1n}x_1x_n+\ldots$$
$$\ldots+a_{n1}x_nx_1+a_{n2}x_nx_2+\ldots+a_{nn}x_n^2=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j.$$

Симметрическая матрица $A=(a_{ij})$ называется матрицей квадратичной формы Q.

○ **Пример.** Написать матрицу квадратичной формы

$$Q=2x_1^2-5x_2^2+8x_3^2+4x_1x_2-2x_1x_3+6x_2x_3.$$

Здесь $a_{11}=2$, $a_{22}=-5$, $a_{33}=8$, $a_{12}=a_{21}=2$, $a_{13}=a_{31}=-1$, $a_{23}=a_{32}=3$. Следовательно,

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

В векторно-матричной форме квадратичная форма имеет вид $Q(x)=xAx$, где $x=(x_1, x_2, \ldots, x_n)$. Если в квадратичной форме $Q=xAx$ неизвестные подвергнуть линейному преобразованию $x=Sy$, то получится квадратичная форма $Q=y(S^{\text{т}}AS)y$ с матрицей $S^{\text{т}}AS$.

Рангом квадратичной формы $Q=xAx$ называется ранг матрицы A. Ранг квадратичной формы не изменяется при невырожденных преобразованиях неизвестных.

Для каждой квадратичной формы $Q=xAx$ можно подобрать такое линейное преобразование неизвестных $x=Sy$ с ортогональной матрицей S, что матрица квадратичной формы $Q=y(S^{\text{т}}AS)y$ будет диагональной.

Если $Q(x)>0$ (<0) для всех $x\neq 0$, то квадратичная форма $Q(x)$ называется *положительно* (*отрицательно*) *определенной*.

Если квадратичная форма $Q(x)$ положительно определена, то форма $-Q(x)$ — отрицательно определенная.

Квадратичная форма $Q(x)=xAx$ положительно (отрицательно) определена тогда и только тогда, когда все собственные значения матрицы A положительны (отрицательны).

Если $A=(a_{ij})$ — квадратная матрица, то определители $\begin{vmatrix} a_{11} & \ldots a_{1k} \\ \cdot & \cdot\;\cdot\;\cdot\;\cdot \\ a_{k1} & \ldots a_{kk} \end{vmatrix}$, $k=1, 2, \ldots, n$, называются *главными* или *угловыми минорами* матрицы A.

Для того чтобы квадратичная форма была положительно определенной, необходимо и достаточно, чтобы все главные миноры матрицы этой формы были положительны.

Для того чтобы квадратичная форма была отрицательно определенной, необходимо и достаточно, чтобы все главные миноры нечетного порядка были отрицательны, а все главные миноры четного порядка — положительны.

2.33. Применение аппарата линейной алгебры для анализа балансовых моделей

Рассматривается экономическая система, состоящая из n отраслей. Обозначим через $x=(x_1, \ldots, x_n)$ вектор валовой продукции системы, а через $y=(y_1, y_2, \ldots, y_n)$ — вектор ее конечной продукции. Тогда система уравнений материального баланса при условии линейности функций производственных издержек имеет вид

$$x_i - \sum_{j=1}^{n} a_{ij}x_j = y_i, \quad i=1,\ 2,\ \ldots,\ n,$$

или в векторно-матричной форме

$$(E-A)\,x=y. \qquad (2.24)$$

Матрицу $A=(a_{ij})$ называют *матрицей затрат* или *технологической матрицей*; E — единичная матрица.

Коэффициенты a_{ik} называются *коэффициентами прямых затрат*; они представляют собой затраты продукции i-й отрасли на изготовление единицы валовой продукции k-й отрасли. Будем считать, что $a_{ik}=\text{const}$. Уравнение (2.24) называется *моделью Леонтьева.*

Одна из задач планирования состоит в том, чтобы при заданном векторе y конечного продукта определить необходимый вектор x валовой продукции.

Матрица A называется *продуктивной*, если существует неотрицательный вектор x^0, для которого $x^0 > Ax^0$.

Если матрица A продуктивна, то система уравнений $(E-A)\,x=y$ имеет единственное неотрицательное решение при любом $y \geqslant 0$, которое можно записать в виде $x = (E-A)^{-1}y$.

Элементы матрицы A определяют те количества промежуточного продукта, которые необходимы для производства единицы валового продукта каждой отрасли.

Элементы матрицы $A^2=(a_{ij}^{(1)})$ называют *косвенными затратами первого порядка*. Величина $(a_{ij}^{(1)})$ — это количество i-го промежуточного продукта, которое необходимо для производства всех материалов, используемых для производства единицы j-й продукции.

Аналогично, элементы матриц $A^3=(a_{ij}^{(2)})$, ..., $A^{k+1}=(a_{ij}^{(k)})$ называют *косвенными затратами второго* и следующих порядков.

Величину полных материальных затрат i-го продукта на производство единицы валовой продукции j-й отрасли определяют по формуле

$$c_{ij}=a_{ij}+a_{ij}^{(1)}+\ldots+a_{ij}^{(k)}+\ldots \quad (i,\ j=1,\ \ldots,\ n),$$

если, конечно, бесконечные ряды сходятся.

Элементы c_{ij} определяют матрицу полных затрат C, причем

$$C=A+A^2+\ldots+A^k+\ldots.$$

Отметим, что $x=Cy+y$. Отсюда следует, что величины c_{ij} представляют собой те количества промежуточного продукта i-го вида, которые необходимы для выпуска одной единицы конечной продукции j-й отрасли.

2.34. Динамическая модель планирования

В модели Леонтьева предполагается, что процесс производства совершается мгновенно. Временные лаги (задержки, отставания) в процессе производства учитывают с помощью моделей динамического межотраслевого баланса.

Разобьем промежуток планирования на T периодов (недель, месяцев, лет). Обозначим через $x^t=(x_1^t,\ x_2^t,\ \ldots,\ x_n^t)$ вектор валовой продукции, произведенный в конце t-го периода. С помощью этого набора продуктов в $(t+1)$-м периоде осуществляется производство вектора x^{t+1}. Последовательность $\{x^1,\ x^2,\ \ldots,\ x^T\}$ называют *траекторией развития производства*. Так как вектор x^{t+1} не определяется однозначно вектором x^t, то имеется много различных траекторий развития производства. Каждая траектория развития производства является решением системы неравенств:

$$Ax^{t+1}\leqslant x^t,\ t=0,\ 1,\ 2,\ \ldots,\ T-1,$$
$$x^t\geqslant 0,\ t=1,\ 2,\ \ldots,\ T,$$

где A — технологическая матрица.

2.35. Линейная модель производства

Производственный способ описывает производство продукции и расход ресурсов в единицу времени. Он может быть задан вектором (x, r), где x — вектор выпуска продукции, r — вектор затрат, соответствующий выпуску x.

Модель производства будет построена при следующих предположениях:

1) если вектор (x, r) технологически допустим, то вектор (kx, kr), $k > 0$, который описывает увеличенный в k раз выпуск продукции при одновременном увеличении затрат в такое же число раз, также технологически допустим;

2) если (x_1, r_1) и (x_2, r_2) — технологически допустимые векторы, то технологически допустимым является вектор $(x_1 + x_2, r_1 + r_2)$;

3) существует конечное число основных производственных процессов, описываемых векторами $A_1 = (x_1, r_1)$, $A_2 = (x_2, r_2)$, ..., $A_n = (x_n, r_n)$, причем каждый происходящий производственный процесс $u = (x, r)$ является линейной комбинацией этих векторов с неотрицательными коэффициентами, т. е.

$$u = \sum_{i=1}^{n} z_i A_i, \; z_i \geqslant 0.$$

Коэффициенты этого разложения называются *интенсивностями основных способов.*

Векторы $A_1, \ldots, A_n$ основных производственных способов служат параметрами производственной системы, а неотрицательный вектор их интенсивностей $z = (z_1, \ldots, z_n)$ является характеристикой внутреннего состояния системы.

Если в производственной системе используется m видов производственных ресурсов, имеющиеся запасы ресурса i ограничены объемом b_i, $i = 1, \ldots, m$, и a_{ij} — затраты ресурса i при использовании j-й технологии с единичной интенсивностью, то модель производственной системы имеет следующий вид:

$$\sum_{j=1}^{n} a_{ij} z_j \leqslant b_i \quad (i = 1, \ldots, m),$$

$$z_j \geqslant 0 \quad (j = 1, \ldots, n).$$

РАЗДЕЛ III. n-МЕРНОЕ ПРОСТРАНСТВО $\mathbf{R}^n$

3.1. Точки в n-мерном пространстве. Расстояние между точками

Последовательность n действительных чисел $x_1, x_2, \ldots, x_n$ называют n-мерной точкой $M(x_1; x_2; \ldots; x_n)$, а сами числа $x_1, x_2, \ldots, x_n$ — координатами точки M.

Множество всех n-мерных точек составляет n-мерное пространство $\mathbf{R}^n$.

Например, $M(1; 2; -3; 4)$ и $N(3; 10; -\sqrt{2}; 3)$ — точки 4-мерного пространства $\mathbf{R}^4$, т. е. $M \in \mathbf{R}^4$, $N \in \mathbf{R}^4$.

Расстоянием между точками $M(x_1; x_2; \ldots; x_n)$ и $N(y_1; y_2; \ldots; y_n)$ называют число

$$\rho(M, N) = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\ldots+(x_n-y_n)^2} = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}\,.$$

Например, если $M(2; 3; -4; 6)$, $N(1; 2; -1,1)$, то

$$\rho(M, N) = \sqrt{(2-1)^2+(3-2)^2+(-4+1)^2+(6-1)^2} = 6.$$

С в о й с т в а р а с с т о я н и я в n-м е р н о м п р о с т р а н с т в е

1°. $\rho(M, N) = \rho(N, M)$.

2°. $\rho(M, N) \geqslant 0$, $\rho(M, N) = 0 \Leftrightarrow M = N$.

3°. $\rho(M, N) \leqslant \rho(M, L) + \rho(L, N)$.

Если даны две точки $M(x_1; x_2; \ldots; x_n)$ и $N(y_1; y_2; \ldots; y_n)$ в пространстве $\mathbf{R}^n$, то можно рассмотреть вектор

$$\overrightarrow{MN} = (y_1-x_1, y_2-x_2, \ldots, y_n-x_n).$$

При этом длина вектора $\overrightarrow{MN}$ совпадает с расстоянием $\rho(M, N)$, т. е. $|\overrightarrow{MN}| = \rho(M, N)$.

Вектор $\overrightarrow{OM}$, где $O(0, 0, \ldots, 0)$, называют *радиусом-вектором* точки M.

3.2. Окрестность точки в n-мерном пространстве

Если r — некоторое положительное число, то r-окрестностью $S_r(M_0)$ точки M_0 в n-мерном пространстве $\mathbf{R}^n$ называют множество всех точек $M \in \mathbf{R}^n$ таких, что $\rho(M, M_0) < r$, т. е.

$$S_r(M_0) = \{M \in \mathbf{R}^n \mid \rho(M, M_0) < r\}.$$

Например, $M_1(2; 3; -1; 3) \in S_2(M_0)$, где $M_0(1; 2; -1; 2)$, так как $\rho(M_1, M_0) = \sqrt{3} < 2$, а $M_2(3; 3; -1; 3) \notin S_2(M_0)$, так как $\rho(M_2, M_0) = \sqrt{6} > 2$.

В пространстве $\mathbf{R}^1$ r-окрестность точки $M_0(a)$ — это интервал $]a-r, a+r[$.

В пространстве $\mathbf{R}^2$ r-окрестность точки $M_0(a; b)$ — это внутренность круга радиуса r с центром в точке $M_0(a, b)$.

В пространстве $\mathbf{R}^3$ r-окрестность точки $M_0(a; b; c)$ — это внутренность шара радиуса r с центром в точке $M_0(a, b, c)$.

3.3. Ограниченные множества в n-мерном пространстве

Множество V точек n-мерного пространства $\mathbf{R}^n$ называют *ограниченным*, если существует число $A > 0$ такое, что для любой точки $M(x_1; x_2; \ldots; x_n) \in V$ выполняются следующие соотношения: $|x_1| \leqslant A$, $|x_2| \leqslant A, \ldots; |x_n| \leqslant A$.

○ **Примеры ограниченных множеств.**

1. Множество V точек $M(x_1, x_2, \ldots, x_n)$ таких, что $|x_1| \leqslant A_1, |x_2| \leqslant A_2, \ldots, |x_n| \leqslant A_n$ ($A_1, A_2, A_3, \ldots, A_n$ — некоторые положительные числа), всегда ограничено.

2. r-окрестность любой точки в n-мерном пространстве всегда ограниченное множество.

3. Пересечение и объединение ограниченных множеств — множества ограниченные. ●

3.4. Внутренние и граничные точки множества в n-мерном пространстве

Точку M_0 называют *внутренней точкой множества* V (рис. 3.1) точек n-мерного пространства $\mathbf{R}^n$, если она входит в множество V вместе с некоторой окрестностью $S_r(M_0)$.

Точку M_0 называют *граничной точкой множества* V, если каждая окрестность точки M_0 содержит как точки из множества V, так и точки, не принадлежащие этому множеству (рис. 3.2).

Множество всех граничных точек множества V называют *границей* этого *множества*.

Например, если $V=[a, b]$, то все точки интервала $]a, b[$ являются внутренними точками множества V, а граница этого множества состоит из двух точек: a и b.

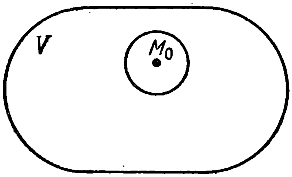


Рис. 3.1

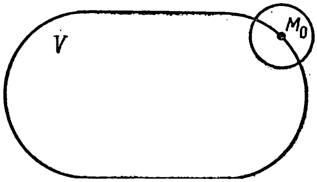


Рис. 3.2

Если же $V=\{M(x, y) \in \mathbf{R}^2 \mid x^2+y^2<1\}$, то все точки этого множества внутренние, а граница совпадает с окружностью $x^2+y^2=1$.

3.5. Предельные точки множества в n-мерном пространстве

Точку $M_0 \in \mathbf{R}^n$ называют *предельной точкой множества* V n-мерных точек, если каждая окрестность точки M_0 содержит бесконечно много точек множества V.

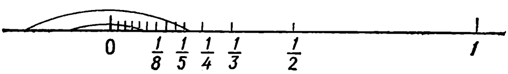


Рис. 3.3

Например, точка $O(0)$ — предельная точка множества $\left\{\frac{1}{k}\right\}$ $(k=1, 2, \ldots)$ (рис. 3.3).

Свойства предельных точек

1°. Любая внутренняя точка множества V является предельной точкой этого множества.

2°. Если предельная точка множества V не принадлежит этому множеству, то она является граничной точкой множества V.

3.6. Замкнутые и открытые множества в $\mathbf{R}^n$

Множество V в $\mathbf{R}^n$ называют *замкнутым*, если оно содержит все свои предельные точки.

Множество V в $\mathbf{R}^n$ называют *открытым*, если все точки множества V являются внутренними.

Например:

1) $[a, b]$ — замкнутое множество в $\mathbf{R}^1$;

2) $]a, b[$ — открытое множество в $\mathbf{R}^1$;

3) $\{M(x, y) \in \mathbf{R}^2 \mid x^2+y^2 \leqslant r^2\}$ — замкнутое множество в $\mathbf{R}^2$;

4) $\{M(x, y) \in \mathbf{R}^2 \mid x^2+y^2 < r^2\}$ — открытое множество в $\mathbf{R}^2$;

5) r-окрестность любой n-мерной точки — открытое множество в пространстве $\mathbf{R}^n$.

Свойства открытых и замкнутых множеств

1°. Если множество V содержит свою границу, то оно замкнуто.

2°. Пересечение любого числа замкнутых множеств — множество замкнутое.

3°. Объединение конечного числа замкнутых множеств — множество замкнутое.

4°. Пересечение конечного числа открытых множеств — множество открытое.

5°. Объединение любого числа открытых множеств — открытое множество.

6°. Множество V открыто тогда и только тогда, когда его дополнение замкнуто.

Ограниченное замкнутое множество в пространстве $\mathbf{R}^n$ называют *компактным*.

3.7. Последовательности n-мерных точек

Говорят, что задана бесконечная последовательность n-мерных точек, если указан закон, по которому каждому натуральному числу k ставится в соответствие определенная n-мерная точка M_k. В этом случае последовательность записывают в виде $M_1, M_2, \ldots, M_k, \ldots$ или, кратко, $\{M_k\}$. Точки $M_1, M_2, \ldots, M_k$ называют *членами последовательности*: M_1 — первым, M_2 — вторым, M_k — k-м членом последовательности.

Например, если каждому натуральному числу k ставится в соответствие точка $M_k(k, k^2)$, то задана последовательность двумерных точек: $M_1(1; 1)$, $M_2(2; 4)$, $M_3(3; 9)$, ..., $M_k(k; k^2)$, ...

Последовательность одномерных точек называют *числовой последовательностью*. Таким образом, числовую последовательность считают заданной, если указан закон, по которому каждому натуральному числу k ставится в соответствие определенное число x_k.

Например, если каждому натуральному числу k ставится в соответствие число $k/(k+1)$, то задана числовая последовательность $1/2, 2/3, 3/4, \ldots, k/(k+1), \ldots$.

Числовую последовательность часто задают с помощью рекуррентного соотношения (выражения последующих членов последовательности через предыдущие).

Например, если $X_1=1$, $X_{k+1}=3X_k+2$, то задана числовая последовательность 1, 5, 17, 53, ...

Последовательность $\overline{M}_1, \overline{M}_2, \ldots, \overline{M}_l, \ldots$ называют *подпоследовательностью* последовательности n-мерных точек $\{M_k\}$, если $\overline{M}_1=M_{k_1}$, $\overline{M}_2=M_{k_2}$, $\ldots$, $\overline{M}_l=M_{k_l}$, $\ldots$, где $k_1<k_2<\ldots<k_l<\ldots$. Таким образом, подпоследовательность всегда составлена из членов данной последовательности, а порядок следования членов подпоследовательности такой же, как у данной последовательности.

Например, числовая последовательность 4, 8, 12, 16, ..., $4k$, ... является подпоследовательностью последовательности 2, 4, 6, 8, 10, ..., $2k$, ...

3.8. Предел последовательности

Пределом последовательности $\{M_k\}$, $M_k \in \mathbf{R}^n$, называется n-мерная точка M_0, если каждая ε-окрестность точки M_0 содержит все члены данной последовательности начиная с некоторого номера, т. е. для любого $\varepsilon>0$ должен существовать номер K (зависящий от ε) такой, что $M_k \in S_\varepsilon(M_0)$ при всех $k>K$.

Если M_0 является пределом последовательности $\{M_k\}$, то пишут $M_0=\lim\limits_{k\to\infty}\{M_k\}$ или $M_k \to M_0$ при $k\to\infty$.

В частности, число a есть предел числовой последовательности $\{x_n\}$, если для любого числа $\varepsilon>0$ можно указать номер N (зависящий от ε) такой, что для всех номеров $n>N$ выполняется неравенство $|x_n-a|<\varepsilon$.

Например, последовательность $\left\{\frac{1}{n}\right\}$ имеет предел $a=0$. Действительно, для любого $\varepsilon>0$ всегда существует натуральное число $N=\left[\frac{1}{\varepsilon}\right]$ $\left(\text{целая часть числа } \frac{1}{\varepsilon}\right)$ такое, что для всех $n>N$ выполняется неравенство $\left|\frac{1}{n}-0\right|=\frac{1}{n}<\varepsilon$. При отыскании предела последовательности n-мерных точек ($n\geqslant 2$) важную роль играет предел числовой последовательности, так как имеют место следующие два утверждения:

1. Точка M_0 является пределом последовательности $\{M_k\}$, $M_k \in \mathbf{R}^n$ тогда и только тогда, когда предел числовой последовательности $\{\rho(M_k, M_0)\}$ равен нулю ($\rho(M_k, M_0)$ — расстояние между точками M_k и M_0).

2. Точка $M_0(x_1^0, x_2^0, \ldots, x_n^0)$ является пределом последовательности $\{M_k\}$, $M_k(x_1^k, x_2^k, \ldots, x_n^k)$ тогда и только тогда, когда $\lim\limits_{k\to\infty} x_1^k = x_1^0$, $\lim\limits_{k\to\infty} x_2^k = x_2^0$, ..., $\lim\limits_{k\to\infty} x_n^k = x_n^0$.

Пример. Точка $M_0(1; 1; \ldots; 1)$ является пределом последовательности $\{M_k\}$, $M_k\left(\frac{k}{k+1}, \frac{k}{k+1}, \ldots, \frac{k}{k+1}\right)$, так как

$$\lim_{k\to\infty} \frac{k}{k+1} = \lim_{k\to\infty} \frac{1}{1+\frac{1}{k}} = 1.$$

Последовательность n-мерных точек называют сходящейся, если она имеет предел.

Свойства сходящихся последовательностей

1°. Если последовательность сходится, то она имеет только один предел.

2°. Любая сходящаяся последовательность ограничена. Из любой ограниченной последовательности можно выделить сходящуюся подпоследовательность.

3°. Если последовательность n-мерных точек сходится к точке M_0, то и любая ее подпоследовательность сходится к M_0.

4°. Если M_0 — предельная точка некоторого множества V ($V \subseteq \mathbf{R}^n$), то существует последовательность точек из множества V, сходящаяся к точке M_0.

5°. Если последовательность точек замкнутого множества сходится к точке M_0, то $M_0 \in V$.

3.9. Бесконечно малые и бесконечно большие числовые последовательности

Числовая последовательность $\{\alpha_n\}$ называется *бесконечно малой*, если $\lim\limits_{n\to\infty} \{\alpha_n\} = 0$, т. е. $\{\alpha_n\}$ — бесконечно малая последовательность, если для любого числа $\varepsilon > 0$ можно указать номер N такой, что при всех $n > N$ выполняется неравенство $|\alpha_n| < \varepsilon$.

Например, если $\varphi(x)$, $g(x)$ — многочлены и $\deg\varphi(x) < \deg g(x)$, то последовательность $\left\{\frac{\varphi(n)}{g(n)}\right\}$ является бесконечно малой.

Бесконечно малые последовательности используют при вычислении пределов последовательностей, так как число a является пределом последовательности $\{x_n\}$ тогда и только тогда, когда $x_n = a + \alpha_n$, где $\{\alpha_n\}$ — бесконечно малая последовательность.

Например, $\lim\limits_{n\to\infty}\frac{2n^2+1}{n^2}=2$, так как $\frac{2n^2+1}{n^2}=2+\frac{1}{n^2}$ и $\left\{\frac{1}{n^2}\right\}$ — бесконечно малая последовательность.

Свойства бесконечно малых последовательностей

1°. Если $\{\alpha_n\}$, $\{\beta_n\}$ — бесконечно малые последовательности, то их сумма или разность $\{\alpha_n\pm\beta_n\}$ также последовательность бесконечно малая.

2°. Если $\{\alpha_n\}$ — бесконечно малая, а $\{x_n\}$ — ограниченная последовательность, то их произведение $\{\alpha_n\cdot x_n\}$ — последовательность бесконечно малая.

3°. Произведение конечного числа бесконечно малых последовательностей является бесконечно малой последовательностью.

Числовая последовательность $\{x_n\}$ называется *бесконечно большой*, если для любого числа $A>0$ существует натуральное число N такое, что при всех $n>N$ выполняется неравенство $|x_n|>A$.

Например, если $f(x)$, $\varphi(x)$ — многочлены и $\deg\varphi(x)>\deg f(x)$, то последовательность $\left\{\frac{\varphi(n)}{f(n)}\right\}$ является бесконечно большой.

Если $\{x_n\}$ — бесконечно большая последовательность, то пишут

$$\lim_{n\to\infty}\{x_n\}=\infty.$$

Если же $\{x_n\}$ — бесконечно большая последовательность и начиная с некоторого номера все x_n положительны (отрицательны), то пишут $\lim\limits_{n\to\infty}\{x_n\}=+\infty\left(\lim\limits_{n\to\infty}\{x_n\}=-\infty\right)$.

Свойства бесконечно больших последовательностей

1°. Бесконечно большая последовательность всегда неограничена. Однако не всякая неограниченная последовательность является бесконечно большой.

2°. Последовательность $\{x_n\}$, $x_n\neq0$ является бесконечно большой тогда и только тогда, когда последовательность $\left\{\frac{1}{x_n}\right\}$ бесконечно малая.

3.10. Арифметические свойства пределов числовых последовательностей

Если последовательность $\{x_n\}$ постоянна, т. е. при всех n $x_n=C=\text{const}$, то $\lim\limits_{n\to\infty}x_n=C$.

Предположим, что последовательности $\{x_n\}$, $\{y_n\}$ сходятся. Тогда:

1) сходятся последовательности $\{x_n+y_n\}$, $\{x_n\cdot y_n\}$, причем $\lim\limits_{n\to\infty}(x_n+y_n)=\lim\limits_{n\to\infty}x_n+\lim\limits_{n\to\infty}y_n$; $\lim\limits_{n\to\infty}(x_n\cdot y_n)=\lim\limits_{n\to\infty}x_n\cdot\lim\limits_{n\to\infty}y_n$;

2) если $\lim\limits_{n\to\infty}y_n\neq 0$, то сходится последовательность $\left\{\dfrac{x_n}{y_n}\right\}$, причем $\lim\limits_{n\to\infty}\dfrac{x_n}{y_n}=\dfrac{\lim\limits_{n\to\infty}x_n}{\lim\limits_{n\to\infty}y_n}$.

Отсюда следует, что если последовательность $\{x_n\}$ сходится, то:

а) $\lim\limits_{n\to\infty}(C\cdot x_n)=C\lim\limits_{n\to\infty}x_n$ $(C=\mathrm{const})$; б) $\lim\limits_{n\to\infty}(x_n)^k=(\lim\limits_{n\to\infty}x_n)^k$

(k — натуральное число).

Приведенные утверждения часто используют при вычислении пределов числовых последовательностей. В частности, можно показать, что если $\varphi(x)$, $g(x)$ — многочлены и $\deg\varphi(x)=\deg g(x)$, то $\lim\limits_{n\to\infty}\dfrac{\varphi(n)}{g(n)}$ равен отношению коэффициентов при старших степенях многочленов $\varphi(x)$ и $g(x)$.

Например,

$$\lim_{n\to\infty}\frac{3n^2-2n+8}{2-4n^2}=-\frac{3}{4}.$$

3.11. Переход к пределу в неравенствах (для числовых последовательностей)

Имеют место следующие утверждения:

1. Если последовательности $\{x_n\}$, $\{y_n\}$ сходятся и при всех номерах n выполняется неравенство $x_n\leqslant y_n$, то $\lim\limits_{n\to\infty}x_n\leqslant\lim\limits_{n\to\infty}y_n$.

2. Если при всех n выполняется неравенство $x_n\leqslant y_n\leqslant z_n$ и $\lim\limits_{n\to\infty}x_n=\lim\limits_{n\to\infty}z_n=a$, то и $\lim\limits_{n\to\infty}y_n=a$.

3.12. Монотонные последовательности. Число e

Числовую последовательность $\{x_n\}$ называют *возрастающей* (*неубывающей*), если каждый последующий член последовательности больше (не меньше) предыдущего, т. е. $x_{n+1}>x_n$ $(x_{n+1}\geqslant x_n)$.

Числовую последовательность $\{x_n\}$ называют *убывающей* (*невозрастающей*), если каждый последующий член последовательности меньше (не больше) предыдущего, т. е. $x_{n+1}<x_n$ $(x_{n+1}\leqslant x_n)$.

Определенные выше последовательности называют *монотонными*.

Например, $\left\{\frac{1}{n}\right\}$ — убывающая последовательность, так как $1>\frac{1}{2}>\frac{1}{3}>\ldots>\frac{1}{n}>\frac{1}{n+1}>\ldots$, а $\left\{\frac{n-1}{n}\right\}$ — возрастающая последовательность, так как $0<\frac{1}{2}<\frac{2}{3}<\ldots$ $\ldots<\frac{n-1}{n}<\frac{n}{n+1}<\ldots$.

Основные свойства монотонных последовательностей

Если числовая последовательность монотонна и ограничена, то она сходится.

Например, последовательность $\left\{\left(1+\frac{1}{n}\right)^n\right\}$ — возрастающая и ограниченная. Следовательно, она сходится. Предел этой последовательности называют числом e, т. е.

$$\mathrm{e}=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \qquad (\mathrm{e}\approx 2{,}718).$$

3.13. Выпуклые множества в n-мерном пространстве

Если $M(x_1;\ x_2;\ \ldots;\ x_n)$ и $N(y_1;\ y_2;\ \ldots;\ y_n)$ — две n-мерные точки, то отрезком $[MN]$ называют множество всех точек $P(z_1,\ \ldots,\ z_n)$, где

$z_1=\alpha x_1+(1-\alpha)y_1,\ z_2=\alpha x_2+(1-\alpha)y_2,\ \ldots,\ z_n=\alpha x_n+(1-$ $-\alpha)y_n$ при $0\leqslant\alpha\leqslant 1$.

Таким образом,

$$[MN]=\{P\in\mathbf{R}^n\mid\overrightarrow{OP}=\overrightarrow{OM}\cdot\alpha+\overrightarrow{ON}(1-\alpha)\ \text{при}\ 0\leqslant\alpha\leqslant 1\}.$$

○ **Пример.** Даны точки $M(1,\ -2;\ 3,\ 4)$ и $N(3;\ 4;\ 1;\ -8)$. Точка $P(2;\ 1;\ 2;\ -2)\in[MN]$, так как $2=\alpha\cdot 1;\ +(1-\alpha)\cdot 3$, $1=\alpha\cdot(-2)+(1-\alpha)\cdot 4$, $2=\alpha\cdot 3+(1-\alpha)\cdot 1$, $-2=\alpha\cdot 4+$ $+(1-\alpha)\cdot(-8)$ при $\alpha=1/2$. Точка $Q(4;\ 3;\ 2;\ -1)\notin[MN]$, так как соотношения $4=\alpha\cdot 1+(1-\alpha)\cdot 3$, $3=\alpha\cdot(-2)+$ $+(1-\alpha)\cdot 4$, $2=\alpha\cdot 3+(1-\alpha)\cdot 1$, $-1=\alpha\cdot 4+(1-\alpha)\cdot(-8)$ не выполняются ни при каком значении α. ●

Множество V в $\mathbf{R}^n$ называется *выпуклым*, если вместе с любыми двумя его точками ему принадлежит и отрезок, соединяющий эти две точки, т. е. если $M\in V$, $N\in V$, то $[MN]\subseteq V$.

Выпуклыми, например, являются следующие множества:

1) все n-мерное пространство $\mathbf{R}^n$;
2) $\{M(x,\ y)\in\mathbf{R}^2\mid x^2+y^2\leqslant r^2\}$;
3) $\{M(x_1,\ x_2,\ \ldots,\ x_n)\in\mathbf{R}^n\mid a_1x_1+a_2x_2+\ldots+a_nx_n=b\}$;

4) $\{M(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n \mid a_1x_1 + a_2x_2 + \ldots + a_nx_n \leqslant b\}$;

5) r-окрестность любой n-мерной точки.

Свойства выпуклых множеств

1°. Пересечение конечного числа выпуклых множеств является выпуклым множеством.

2°. Если точки $M_1, M_2, \ldots, M_k$ принадлежат выпуклому множеству V и $\overrightarrow{OP} = \lambda_1\overrightarrow{OM_1} + \lambda_2\overrightarrow{OM_2} + \ldots + \lambda_k\overrightarrow{OM_k}$ при $\lambda_1 \geqslant 0, \lambda_2 \geqslant 0, \ldots, \lambda_k \geqslant 0$, $\lambda_1 + \lambda_2 + \ldots + \lambda_k = 1$, то точка P принадлежит множеству V.

Выпуклой оболочкой точек $M_1, M_2, \ldots, M_k$ называется множество $\{P \in \mathbf{R}^n \mid \overrightarrow{OP} = \lambda_1\overrightarrow{OM_1} + \lambda_2\overrightarrow{OM_2} + \ldots + \lambda_k\overrightarrow{OM_k}$, $\lambda_1 \geqslant 0, \lambda_2 \geqslant 0, \ldots, \lambda_k \geqslant 0, \lambda_1 + \lambda_2 + \ldots + \lambda_k = 1\}$.

Выпуклая оболочка всегда является выпуклым множеством.

Если выпуклое множество содержит точки $M_1, M_2, \ldots, M_k$, то оно содержит и всю выпуклую оболочку этих точек.

3.14. Крайние точки выпуклых множеств

Точка P выпуклого множества V в n-мерном пространстве называется *крайней*, если она не может быть серединой отрезка, концы которого лежат в множестве V, т. е. если не существует точек $M_1, M_2 \in V$, $M_1 \neq M_2$ таких, что

$$\overrightarrow{OP} = \frac{1}{2}\overrightarrow{OM_1} + \frac{1}{2}\overrightarrow{OM_2}.$$

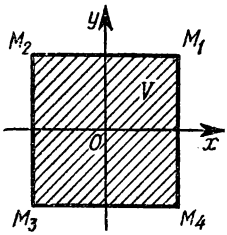


Рис. 3.4

Например, множество $V = \{M(x, y) \in \mathbf{R}^2 \mid |x| \leqslant a, |y| \leqslant a, a > 0\}$ имеет четыре крайние точки: $M_1(a; a)$, $M_2(-a; a)$, $M_3(-a; -a)$, $M_4(a; -a)$ (рис. 3.4).

Выпуклая оболочка n-мерных точек $M_1, M_2, \ldots, M_k$ имеет лишь конечное число крайних точек и совпадает с выпуклой оболочкой своих крайних точек.

Непустое выпуклое компактное множество в $\mathbf{R}^n$ имеет крайние точки.

3.15. Непрерывные отображения пространства и неподвижные точки

Говорят, что задано *отображение f пространства* $\mathbf{R}^n$ *в себя* ($f\colon \mathbf{R}^n \to \mathbf{R}^n$), если каждой точке $M \in \mathbf{R}^n$ поставлена в соответствие определенная точка $N = f(M) \in \mathbf{R}^n$.

Точка $M_0 \in \mathbf{R}^n$ называется *неподвижной точкой отображения* f: $\mathbf{R}^n \to \mathbf{R}^n$, если $f(M_0) = M_0$.

Рассмотрим, например, отображение пространства $\mathbf{R}^2$ в себя:

$$(x_1, x_2) \to (x_1^2, 2x_2 - 1).$$

Точки $M_1(0; 1)$, $M_2(1; 1)$ являются неподвижными точками этого отображения.

Отображение f: $\mathbf{R}^n \to \mathbf{R}^n$ называют *непрерывным в точке* $M_0 \in \mathbf{R}^n$, если для любой последовательности точек пространства $\mathbf{R}^n$: $M_1, M_2, \ldots, M_k, \ldots$, сходящейся к M_0, последовательность $f(M_1), f(M_2), \ldots, f(M_k), \ldots$ сходится к $f(M_0)$.

Отображение f: $\mathbf{R}^n \to \mathbf{R}^n$ непрерывно на множестве V $(V \subseteq \mathbf{R}^n)$, если оно непрерывно в каждой точке этого множества.

Теорема Брауэра. *Пусть V — непустое компактное выпуклое множество пространства $\mathbf{R}^n$, а f: $\mathbf{R}^n \to \mathbf{R}^n$ — отображение пространства $\mathbf{R}^n$. Если отображение f непрерывно на множестве V и $f(M) \in V$ для всех $M \in V$, то в множестве V существует неподвижная точка этого отображения.*

Отображение f: $\mathbf{R}^n \to \mathbf{R}^n$ называется *сжимающим*, если существует такое положительное число $\alpha < 1$, что для любых двух точек $M_1, M_2 \in \mathbf{R}^n$ выполняется неравенство

$$\rho(f(M_1), f(M_2)) \leqslant \alpha\rho(M_1, M_2).$$

Рассмотрим, например, отображение пространства $\mathbf{R}^n$:

$$(x_1, x_2, \ldots, x_n) \to (y_1, y_2, \ldots, y_n),$$

где $y_i = \sum_{k=1}^{n} a_{ik}x_k + b_i$, $i = 1, 2, \ldots, n$. Если $\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}^2 < 1$, то это отображение является сжимающим.

Свойства сжимающих отображений

1°. Сжимающее отображение непрерывно на всем пространстве $\mathbf{R}^n$.

2°. Сжимающее отображение пространства $\mathbf{R}^n$ имеет неподвижную точку, и притом единственную.

3.16. Точечно-множественные (многозначные) отображения пространства $\mathbf{R}^n$

Пусть V — некоторое непустое множество в $\mathbf{R}^n$, а $P(V)$ — множество всех подмножеств V.

Отображение F: $V \to P(V)$ называют *точечно-множе-*

ственным (*многозначным*) *отображением* множества V. Это отображение каждой точке $M \in V$ ставит в соответствие одно вполне определенное непустое подмножество $F(M)$ множества V.

Точка $M_0 \in V$ является неподвижной точкой точечно-множественного отображения $F: V \to P(V)$, если $M_0 \in F(M_0)$. Например, если $V = \{a, b, c\} \subseteq \mathbf{R}^1$, то соответствие $a \to \{c\}$, $b \to \{a, b\}$, $c \to \{b, c\}$ определяет точечно-множественное отображение. Точки b и c являются неподвижными точками этого отображения.

Точечно-множественное отображение $F: V \to P(V)$ называется *замкнутым в точке* $M_0 \in V$, если из сходимости последовательности точек множества V: $M_1, M_2, \ldots, M_k, \ldots$ к точке M_0 следует сходимость любой последовательности $N_1, N_2, \ldots, N_k, \ldots$, где $N_k \in F(M_k)$, к некоторой точке из множества $F(M_0)$.

Теорема Какутани. *Пусть V — компактное выпуклое множество в пространстве $\mathbf{R}^n$, а $F: V \to P(V)$ — точечно-множественное отображение, удовлетворяющее следующим условиям:*

а) *для любой точки $M \in V$ множество $F(M)$ является непустым выпуклым подмножеством V;*

б) *отображение F замкнуто в любой точке множества V. Тогда отображение F имеет неподвижную точку.*

3.17. Подпространства пространства $\mathbf{R}^n$

Множество P в пространстве $\mathbf{R}^n$ называется *подпространством этого пространства*, когда выполняются следующие условия:

1) если $M \in P$, $N \in P$ и $\overrightarrow{OL} = \overrightarrow{OM} + \overrightarrow{ON}$, то и $L \in P$;

2) если $M \in P$ и $\overrightarrow{OL} = k \cdot \overrightarrow{OM}$, где k — некоторое число, то $L \in P$.

Любое подпространство пространства $\mathbf{R}^n$ содержит точку $O(0, 0, \ldots, 0)$ и является выпуклым множеством. Следующие множества являются подпространствами $\mathbf{R}^n$:

а) множество, состоящее из одной точки $O(0; 0; \ldots; 0)$;

б) все пространство $\mathbf{R}^n$;

в) множество решений однородной системы линейных уравнений.

Пересечение подпространств пространства $\mathbf{R}^n$ само является подпространством этого пространства.

Линейной оболочкой точек $M_1, M_2, \ldots, M_k$ в пространстве $\mathbf{R}^n$ называется множество всех точек $M \in \mathbf{R}^n$

таких, что

$$\overrightarrow{OM}=\sum_{i=1}^{k}\lambda_i\overrightarrow{OM}_i,$$

где $\lambda_1, \lambda_2, \ldots, \lambda_k$ — некоторые числа.

Линейная оболочка всегда является подпространством. Если подпространство P содержит точки $M_1, M_2, \ldots, M_k$, то оно содержит и всю линейную оболочку этих точек.

В любом подпространстве пространства $\mathbf{R}^n$ существует конечное число точек, линейная оболочка которых совпадает с этим подпространством (наименьшее число точек с таким свойством называется размерностью подпространства).

3.18. Выпуклые конусы в пространстве $\mathbf{R}^n$

Выпуклое множество K в пространстве $\mathbf{R}^n$ называется *выпуклым конусом*, когда выполняется следующее условие: если $M\in K$ и $\overrightarrow{OL}=k\cdot\overrightarrow{OM}$, где $k\geqslant 0$, то $L\in K$.

Следующие множества являются выпуклыми конусами в $\mathbf{R}^n$:

а) множество всех точек пространства $\mathbf{R}^n$ с неотрицательными координатами;

б) любое подпространство пространства $\mathbf{R}^n$;

в) $K=\{M(x_1, x_2, x_3)\in\mathbf{R}^3\mid x_1^2+x_2^2-x_3^2\leqslant 0,\ x_3\geqslant 0\}$ (рис. 3.5).

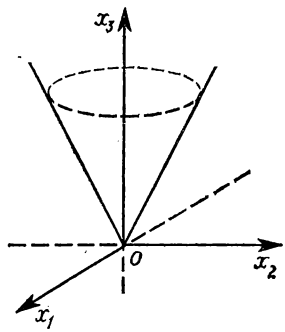


Рис. 3.5

Пересечение выпуклых конусов всегда является выпуклым конусом.

Выпуклый конус K называется *конечным* (*многогранным*), если существуют точки $M_1, M_2, \ldots, M_k$ такие, что

$$K=\left\{M\in\mathbf{R}^n\mid\overrightarrow{OM}=\sum_{i=1}^{k}\lambda_i\overrightarrow{OM}_i,\ \lambda_i\geqslant 0,\ i=1, 2, \ldots, k\right\}.$$

Например, множество решений однородной системы линейных неравенств

$$\sum_{j=1}^{n}a_{ij}x_j\leqslant 0,\quad i=1, 2, \ldots, m,$$

является конечным конусом в $\mathbf{R}^n$.

Конечный конус всегда замкнут. Пересечение двух конечных конусов снова является конечным конусом.

Если K — выпуклый конус в пространстве $\mathbf{R}^n$, то множество $K^* = \{L \in \mathbf{R}^n \mid \overrightarrow{OL} \cdot \overrightarrow{OM} \geqslant 0$ для всех $M \in K\}$ также является выпуклым конусом в $\mathbf{R}^n$. Конус K^* называется *сопряженным* (*двойственным*) конусу K.

В частности, если конус K задается однородной системой линейных неравенств

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant 0, \quad i = 1, 2, \ldots, m,$$

то

$$K^* = \left\{ L(x_1, x_2, \ldots, x_n) \in \mathbf{R}^n \mid x_j = \sum_{i=1}^{m} a_{ij}y_i; \ y_i \geqslant 0 \right\}.$$

Свойства сопряженных конусов

1°. Сопряженный конус K^* всегда замкнут.

2°. Конус, сопряженный к конечному конусу, сам будет конечным.

3°. Если K — замкнутый выпуклый конус, то

$$K^{**} = K.$$

3.19. Суммы выпуклых множеств в пространств $\mathbf{R}^n$

Пусть V и W — множества в пространстве $\mathbf{R}^n$. *Суммой* $V + W$ называется множество всех точек $M \in \mathbf{R}^n$ таких, что

$$\overrightarrow{OM} = \overrightarrow{OM}_1 + \overrightarrow{OM}_2,$$

где

$$M_1 \in V, \quad M_2 \in W.$$

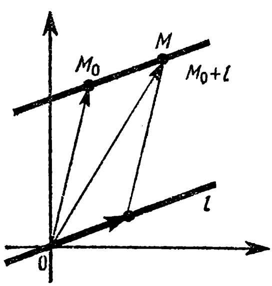


Рис. 3.6

Например, суммой множества, состоящего из одной точки $M_0 \in \mathbf{R}^2$ и прямой $l \subseteq \mathbf{R}^2$, проходящей через точку $O(0; 0)$, является прямая, проходящая через точку M_0 параллельно прямой l (рис. 3.6).

Свойства суммы выпуклых множеств в пространстве $\mathbf{R}^n$.

1°. Сумма выпуклых множеств всегда является выпуклым множеством.

2°. Сумма подпространств пространства $\mathbf{R}^n$ будет подпространством этого пространства.

3°. Сумма выпуклых конусов в $\mathbf{R}^n$ является выпуклым конусом, а сумма конечных конусов — конечным конусом.

Имеют место следующие два утверждения:

1) Множество всех решений системы линейных уравнений

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \ldots, m$$

(если оно не пусто), является суммой множества, состоящего из одной точки и подпространства пространства $\mathbf{R}^n$.

2) Множество всех решений системы линейных неравенств

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \quad i = 1, 2, \ldots, m,$$

является суммой выпуклой оболочки конечного числа точек в пространстве $\mathbf{R}^n$ и конечного конуса.

РАЗДЕЛ IV. АНАЛИЗ ФУНКЦИЙ ОДНОЙ И МНОГИХ ПЕРЕМЕННЫХ

4.1. Понятие функции

Пусть V — некоторое множество точек n-мерного пространства, т. е. $V=\{M(x_1;\ x_2;\ \ldots;\ x_n)\}\subseteq \mathbf{R}^n$. Говорят, что на множестве V задана функция $y=f(M)=f(x_1,\ \ldots,\ x_n)$ переменных $x_1,\ x_2,\ \ldots,\ x_n$, если каждой точке $M\in V$ поставлено в соответствие определенное действительное число $f(M)$. Число $f(M)$ называют при этом *значением функции y в точке M*.

В частности, если $V\subseteq \mathbf{R}^1$, т. е. V является подмножеством множества действительных чисел $\mathbf{R}^1=\{x\}$, говорят, что на множестве V задана функция одной переменной $y=f(x)$.

○ **Примеры.**

1. $f(x)=\lg x$ — функция одной переменной x, заданная на множестве $V=\{x\in \mathbf{R}^1 \mid x>0\}$. В частности, $f(10)=\lg 10=1$.

2. $f(M)=\dfrac{1-x_1x_2}{x_1^2+x_2^2}$ — функция двух переменных $x_1,\ x_2$, заданная на множестве $V=\mathbf{R}^2\setminus\{O(0,\ 0)\}$. В частности, в точке $M(1;\ -1)$ имеем $f(M)=\dfrac{1-1(-1)}{1^2+(-1)^2}=1$.

3. $f(M)=\sqrt{4-x_1^2-x_2^2-x_3^2}$ — функция трех переменных $x_1,\ x_2,\ x_3$, заданная на множестве $V=\{M(x_1;\ x_2;\ x_3)\in \mathbf{R}^3 \mid x_1^2+x_2^2+x_3^2\leqslant 4\}$. В частности, в точке $M(1,\ 1,\ 1)$ имеем $f(M)=\sqrt{4-1^2-1^2-1^2}=1$. ●

4.2. Область определения и множество значений функции

Множество, на котором задана функция $y=f(M)=f(x_1,\ x_2,\ \ldots,\ x_n)$, называется *областью определения функции* и обозначается $D(y)$; $D(y)\subseteq \mathbf{R}^n$.

Множество всех значений, которые принимает функция $y=f(M)$ во всех точках своей области определения $D(y)$, называется *множеством значений функции* и обозначается $E(y)$; $E(y) \subseteq \mathbf{R}^1$.

Примеры.

1. $y=\sqrt{x-1}$ — функция одной переменной; $D(y) = [1, +\infty[\subseteq \mathbf{R}^1$; $E(y)=[0; +\infty[\subseteq \mathbf{R}^1$.

2. $y=\frac{1}{x_1^2+x_2^2}$ — функция двух переменных; $D(y) = \mathbf{R}^2 \setminus \{(0, 0)\} \subseteq \mathbf{R}^2$; $E(y)=]0, +\infty[\subseteq \mathbf{R}^1$.

3. $y=\sqrt{1-x_1^2-x_2^2-x_3^2}$ — функция трех переменных; $D(y)=\{M(x_1, x_2, x_3) \mid x_1^2+x_2^2+x_3^2 \leqslant 1\} \subseteq \mathbf{R}^3$ (представляет шар единичного радиуса с центром в начале координат); $E(y)=[0, 1] \subseteq \mathbf{R}^1$.

4.3. Ограниченные функции

Функция $y=f(M)$, определенная на множестве V, называется *ограниченной сверху* (*снизу*), если множество принимаемых ею на V значений ограничено сверху (снизу).

Ограниченность сверху (снизу) функции $y=f(M)$ на множестве V означает существование такого числа k, что для всех точек $M \in V$ выполняется неравенство $f(M) \leqslant k$ $(f(M) \geqslant k)$.

Функция $y=f(M)$ называется *ограниченной на множестве* V, если она ограничена на этом множестве и сверху, и снизу.

В частности, если V является окрестностью некоторой точки M_0, т. е. $V=S_r(M_0)=\{M \in \mathbf{R}^n \mid \rho(M, M_0) < r\}$, то говорят об ограниченности функции $y=f(M)$ в данной окрестности точки M_0.

Если V — область определения $D(f)$ функции $y=f(M)$, то говорят об ограниченности функции в области определения, при этом множество значений $E(f)$ является ограниченным множеством.

Если функция $y=f(M)$ не ограничена сверху (снизу) на множестве V, то существует последовательность $\{M_k\}$ точек, принадлежащих V $(k=1, 2, 3, \ldots)$, такая, что

$$\lim_{k\to\infty}\{f(M_k)\}=+\infty \quad \left(\lim_{k\to\infty}\{f(M_k)\}=-\infty\right).$$

○ **Примеры.**

1. Функция $f(x)=\sin x$ ограничена во всей области определения $D(f)=]-\infty, +\infty[$, так как множество ее значений $E(f)=[-1, 1]$ — множество ограниченное $(-1 \leqslant \sin x \leqslant 1)$ (рис. 1.13).

2. Функция $f(M)=\frac{1}{x_1^2+x_2^2}$ ограничена лишь снизу во всей области определения $D(f)=R^2\setminus\{(0;\ 0)\}$, так как множество ее значений $E(f)$ ограничено только снизу так, что $f(M)>0$. Функция не ограничена сверху в любой окрестности точки $(0,\ 0)$: существует последовательность $\left\{M_k\left(\frac{1}{k};\ \frac{1}{k}\right)\right\}$, $k=1,\ 2,\ \ldots$, сходящаяся к точке $O(0,\ 0)$ и такая, что последовательность значений функции $f(M_k)=$ $=\frac{1}{\left(\frac{1}{k}\right)^2+\left(\frac{1}{k}\right)^2}=\frac{k^2}{2}$ стремится к $+\infty$. ●

4.4. Сложные функции (суперпозиции)

Пусть функция $y=f(u_1;\ u_2;\ \ldots;\ u_m)$ определена на некотором множестве $W\subseteq R^m$ переменных $u_1,\ u_2,\ \ldots,\ u_m$, а каждая из функций

$$u_1=g_1(x_1;\ x_2;\ \ldots;\ x_n),\ u_2=g_2(x_1;\ x_2;\ \ldots;\ x_n),\ \ldots$$
$$\ldots,\ u_m=g_m(x_1;\ x_2;\ \ldots;\ x_n)$$

определена на некотором множестве $V\subseteq R^n$ переменных $x_1,\ x_2,\ \ldots,\ x_n$.

Если при этом каждой точке $M(x_1;\ \ldots;\ x_n)\in V$ можно поставить в соответствие точку $N(u_1;\ u_2;\ \ldots;\ u_m)\in W$, где

$$u_1=g_1(x_1;\ \ldots;\ x_n),\ u_2=g_2(x_1;\ \ldots;\ x_n),\ \ldots$$
$$\ldots,\ u_m=g_m(x_1;\ \ldots;\ x_n)$$

(см. п. 1.22), то на множестве V определяется функция $y=f[g_1(x_1,\ x_2,\ \ldots,\ x_n),\ g_2(x_1,\ x_2,\ \ldots,\ x_n),\ \ldots,\ g_m(x_1,\ x_2,\ \ldots$ $\ldots,\ x_n)]$, называемая *сложной функцией* переменных x_1, $x_2,\ \ldots,\ x_n$ или *суперпозицией* функций $f,\ g_1,\ g_2,\ \ldots,\ g_m$.

В частности, если даны две функции одной переменной $y=f(u)$, $u=g(x)$ и при этом $E(g)\subseteq D(f)$ (множество значений функции g является подмножеством области определения функции f), то говорят о сложной функции $y=f[g(x)]$ одной переменной x. Например, пара функций $y=2^u$, $u=\sin x$ задает сложную функцию $y=2^{\sin x}$, определенную на множестве $\mathbf{R}^1$ и имеющую множеством значений отрезок $[1/2,\ 2]$. Аналогично, функция $y=$ $=\ln\cos\frac{1}{\sqrt{x}}$ является суперпозицией следующих функций

$$y=\ln u,\ u=\cos v,\ v=\frac{1}{z},\ z=\sqrt{x}.$$

4.5. Неявные функции

Говорят, что функция $y=f(M)=f(x_1, x_2, \ldots, x_n)$ неявно задана уравнением $F(x_1, x_2, \ldots, x_n, y)=0$, если существует множество $V \subseteq \mathbb{R}^n$ такое, что для всех точек $M(x_1, x_2, \ldots, x_n) \in V$ справедливо тождество

$$F[x_1, x_2, \ldots, x_n, f(x_1, x_2, \ldots, x_n)] \equiv 0.$$

Одно и то же уравнение может задавать неявно не одну, а несколько функций. Например, уравнение $x_1^2+x_2^2+\ldots$ $\ldots+x_n^2+y^2=1$ задает неявно две функции

$$y_1=f_1(M)=+\sqrt{1-x_1^2-x_2^2-\ldots-x_n^2}$$

и

$$y_2=f_2(M)=-\sqrt{1-x_1^2-x_2^2-\ldots-x_n^2},$$

определенные на множестве

$$V=\{M \in \mathbb{R}^n \mid x_1^2+x_2^2+\ldots+x_n^2 \leqslant 1\}.$$

В частности, уравнение $F(x, y)=0$ при указанных предположениях задает неявно функцию $y=f(x)$ одной переменной x; уравнение $F(x_1, x_2, y)=0$ задает неявно функцию $y=f(x_1, x_2)$ двух переменных и т. д.

Название «неявная функция» отражает способ задания функциональной зависимости.

4.6. Параметрическое задание функций

Часто бывает полезно (например, при изучении неявных функций) функциональную зависимость между несколькими переменными выражать через вспомогательные переменные — параметры. Так, для функции, неявно заданной уравнением $F(x, y)=0$, необходимо каждую из переменных x и y выразить через один параметр; для функции, неявно заданной уравнением $F(x_1, x_2, y)=0$, необходимо каждую из переменных x_1, x_2, y выразить через два параметра. Выражение переменных через параметры называют *параметрическим заданием* функциональной зависимости.

○ **Примеры.**

1. Эллипс $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ задается параметрически в виде $x=a \cdot \cos t$, $y=b \cdot \sin t$, где $0 \leqslant t \leqslant 2\pi$.

2. Прямая линия в пространстве имеет параметрическое задание $x=x_0+mt$, $y=y_0+nt$, $z=z_0+pt$, где $(x_0; y_0; z_0)$ — точка, через которую проходит прямая; $(m; n; p)$ — вектор, параллельный прямой; $-\infty < t < +\infty$.

3. Зависимость $z=x^2+y^2$ (параболоид вращения) может быть задана параметрически в виде $x=r\cdot\cos t$, $y=r\cdot\sin t$, $z=r^2$, где параметры r и t изменяются в следующих пределах: $0\leqslant r<+\infty$; $0\leqslant t\leqslant 2\pi$. ●

4.7. Выпуклые и вогнутые функции

Пусть функция $y=f(M)$ определена на выпуклом множестве $V\subseteq \mathbf{R}^n$.

Функция $y=f(M)$ называется *выпуклой* (*вогнутой*) на множестве V, если для любых двух точек $M_1(x_1, x_2, \ldots, x_n)$ и $M_2(y_1, y_2, \ldots, y_n)$, принадлежащих V, и для любого действительного числа $0\leqslant\alpha\leqslant 1$ выполняется неравенство

$$f(N)\leqslant \alpha f(M_1)+(1-\alpha)f(M_2)\ (f(N)\geqslant \alpha f(M_1)+(1-\alpha)f(M_2)),$$

где

$$N(\alpha x_1+(1-\alpha)y_1;\ \alpha x_2+(1-\alpha)y_2;\ \ldots;\ \alpha x_n+(1-\alpha)y_n).$$

○ **Примеры.**

1. Функция $f(x)=x^2$ — выпуклая на $\mathbf{R}^1$. Действительно, для произвольных x_1, $x_2\in\mathbf{R}^1$ и любого $\alpha\in[0, 1]$ получим

$$\alpha f(x_1)+(1-\alpha)f(x_2)-f[\alpha x_1+(1-\alpha)x_2]=\alpha x_1^2+(1-\alpha)x_2^2-$$
$$-[\alpha x_1+(1-\alpha)x_2]^2=\alpha(1-\alpha)x_1^2-2\alpha(1-\alpha)x_1x_2+$$
$$+\alpha(1-\alpha)x_2^2=\alpha(1-\alpha)(x_1-x_2)^2\geqslant 0.$$

2. Линейная функция $f(M)=a_1x_1+a_2x_2+\ldots+a_nx_n$ является одновременно и выпуклой, и вогнутой на всем пространстве $\mathbf{R}^n$.

3. Квадратичная функция

$$f(M)=a_{11}x_1^2+a_{22}x_2^2+\ldots+a_{nn}x_n^2+2a_{12}x_1x_2+\ldots$$
$$\ldots+2a_{1n}x_1x_n+\ldots+2a_{n-1n}x_{n-1}x_n$$

является выпуклой (вогнутой) на $\mathbf{R}^n$ тогда и только тогда, когда она положительно (отрицательно) определена, т. е. принимает неотрицательные (неположительные) значения. Например, функция $f(M)=2x_1^2+11x_2^2+52x_3^2+8x_1x_2+$ $+4x_1x_3-16x_2x_3$ является выпуклой на пространстве $\mathbf{R}^3$.

Действительно,

$$f(M)=2(x_1^2+4x_1x_2+2x_1x_3)+11x_2^2+52x_3^2-16x_2x_3=$$
$$=2(x_1^2+4x_1x_2+2x_1x_3+4x_2^2+x_3^2+4x_2x_3)+3x_2^2+50x_3^2-24x_2x_3=$$
$$=2(x_1+2x_2+x_3)^2+3(x_2^2-8x_2x_3+16x_3^2)+2x_3^2=$$
$$=2(x_1+2x_2+x_3)^2+3(x_2-4x_3)^2+2x_3^2\geqslant 0$$

во всех точках пространства $\mathbf{R}^3$, т. е. функция $f(M)$ положительно определенная. ●

Свойства выпуклых функций

1°. Функция $f(M)$ выпукла на множестве V тогда и только тогда, когда функция $-f(M)$ вогнута на V.

2°. Если функции $f_1(M)$ и $f_2(M)$ выпуклы на множестве V, то функция $k_1f_1(M)+k_2f_2(M)$, где k_1, k_2 — произвольные неотрицательные числа, также является выпуклой на V.

3°. Если функция $f(M)$ выпукла на множестве V, то множество $\{M \in V \mid f(M) \leqslant b\}$, где b — любое число, если только оно не пусто, само является выпуклым множеством.

4°. Если выпуклая функция $f(M)$ определена на открытом множестве V, то на этом множестве она непрерывна.

Аналогичные свойства имеют место и для вогнутых функций.

4.8. Специфические свойства функций одной переменной

Функция $y=f(x)$, определенная на множестве $V \subseteq \mathbf{R}^1$, называется *четной* на этом множестве, если множество V симметрично относительно точки $x=0$ и имеет место равенство $f(-x)=f(x)$ для любого $x \in V$.

График четной функции симметричен относительно оси ординат Oy.

Функция $y=f(x)$, определенная на множестве $V \subseteq \mathbf{R}^1$, называется *нечетной* на этом множестве, если множество V симметрично относительно точки $x=0$ и имеет место равенство $f(-x)=-f(x)$ для любого $x \in V$.

График нечетной функции симметричен относительно начала координат.

○ **Примеры.**

1. Функция $y=\cos x$, для которой $D(y)=]-\infty, +\infty[$, является четной функцией, так как $\cos(-x)=\cos x$ для всех $x \in D(y)$.

2. Функция $y=\arcsin x$, для которой $D(y)=[-1, 1]$, является нечетной функцией, так как $\arcsin(-x)=$ $=-\arcsin x$ для всех $x \in D(y)$. ●

Функция $y=f(x)$ называется *периодической*, если существует такое положительное действительное число t, что для всех точек x и $x+t$ из области определения функции имеет место равенство $f(x+t)=f(x)$. При этом число t называют *периодом* функции.

Практически всегда ставится вопрос о наименьшем из всех возможных периодов, т. е. о числе $T=\min\limits_i t_i$.

Если функция $y=f(x)$ непрерывна, отлична от постоянной и периодическая на $\mathbf{R}^1$, то существует наименьший период T этой функции. Все остальные периоды кратны T, т. е. $t_i=nT$, где $n=1, 2, 3, \ldots$.

○ **Примеры.**

1. $y=\sin x$ и $y=\cos x$ имеют период $T=2\pi$.
2. $y=\operatorname{tg} x$ и $y=\operatorname{ctg} x$ имеют период $T=\pi$.
3. Функция Дирихле

$$y=\begin{cases} 1, & \text{если } x\text{—рациональное,} \\ 0, & \text{если } x\text{—иррациональное,} \end{cases}$$

имеет периодом любое положительное рациональное число, однако не имеет наименьшего периода. ●

Функция $y=f(x)$ называется *возрастающей* (*убывающей*) на некотором множестве $V \subseteq \mathbf{R}^1$, если она определена на этом множестве и если для любых значений x_1, $x_2 \in V$ из условия $x_1 < x_2$ следует неравенство

$$f(x_1) < f(x_2) \quad (f(x_1) > f(x_2)).$$

Функция $y=f(x)$ называется *неубывающей* (*невозрастающей*) на некотором множестве $V \subseteq \mathbf{R}^1$, если она определена на этом множестве и если для любых значений x_1, $x_2 \in V$ из условия $x_1 < x_2$ следует неравенство

$$f(x_1) \leqslant f(x_2) \quad (f(x_1) \geqslant f(x_2)).$$

Возрастающие, убывающие, неубывающие и невозрастающие функции называют *монотонными* функциями. Возрастающие и убывающие функции называют *строго монотонными*.

○ **Примеры.**

1. $y=\lg x$—строго монотонно возрастающая функция во всей области определения (см. рис. 1.11).
2. $y=\left(\frac{1}{2}\right)^x$—строго монотонно убывающая функция в области определения (см. рис. 1.10).
3. $y=x^2$—возрастающая в промежутке $[0, +\infty[$ и убывающая в промежутке $]-\infty, 0]$ (см. рис. 1.5).
4. $y=[x]$—целая часть числа x (см. рис. 1.18)—неубывающая функция. ●

4.9. Обратная функция

Пусть функция $y=f(x)$ определена в области $D(y) \subseteq \mathbf{R}^1$ и имеет множество значений $E(y)$. Если эта функция такова, что для любых x_1, $x_2 \in D(y)$ из условия $x_1 \neq x_2$

следует условие $f(x_1) \neq f(x_2)$, то каждому $y \in E(y)$ можно поставить в соответствие определенное $x \in D(y)$ такое, что $f(x)=y$, т. е. на множестве $E(y)$ можно определить функцию $x=g(y)$, называемую *обратной* к заданной функции $y=f(x)$.

Областью определения обратной функции является множество значений $E(y)$ функции $y=f(x)$. Множеством значений обратной функции является область определения $D(y)$ функции $y=f(x)$.

Например, функция $y=x^2$, заданная в промежутке $[0, +\infty[$, имеет обратную функцию $x=+\sqrt{y}$, определенную на множестве $E(y)=[0, +\infty[$. Эта же функция, заданная в промежутке $]-\infty, 0]$, имеет обратную функцию $x=-\sqrt{y}$, определенную на множестве $E(y)=[0, +\infty[$. Однако функция $y=f(x)=x^2$, заданная, например, на отрезке $[-2, 2]$, не имеет обратной функции, так как $f(-1)=f(1)=1$ (двум различным значениям аргумента $x=-1$ и $x=1$ соответствует одно и то же значение y).

Если функция $y=f(x)$ определена, непрерывна и строго монотонна на отрезке $[a, b]$, то она имеет обратную функцию $x=g(y)$, определенную, непрерывную и строго монотонную на отрезке с концами в точках $f(a)$ и $f(b)$.

4.10. Понятие предела функции

Пусть функция $y=f(M)=f(x_1, x_2, \ldots, x_n)$ определена на множестве $V \subseteq \mathbf{R}^n$ и $M_0(x_1^0; x_2^0; \ldots; x_n^0)$ — предельная точка множества V.

Имеют место два эквивалентных между собой определения предела функции.

Число b называется *пределом функции* $f(M)$ при M, стремящемся к M_0 $(M \to M_0)$, если для любой последовательности точек $M_1, M_2, \ldots, M_k, \ldots$, где $M_k \in V$ $(k=1, 2, 3, \ldots)$, $M_k \neq M_0$, сходящейся к M_0, последовательность значений функции $f(M_1), f(M_2), \ldots, f(M_k), \ldots$ сходится к числу b. При этом пишут

$$b = \lim_{M \to M_0} f(M) \text{ или } b = \lim_{\substack{x_1 \to x_1^0 \\ x_2 \to x_2^0 \\ \cdots \\ x_n \to x_n^0}} f(x_1; x_2; \ldots, x_n).$$

В частности, для функции одной переменной $y=f(x)$ число b называется пределом при $x \to x_0$, если для любой последовательности значений аргумента $x_1, x_2, \ldots, x_k, \ldots$, где $x_k \in V$, $x_k \neq x_0$ $(k=1, 2, 3, \ldots)$, сходящейся к x_0, последовательность значений функции $f(x_1), f(x_2), \ldots, f(x_k), \ldots$ сходится к числу b:

$$b = \lim_{x \to x_0} f(x) \text{ или } f(x) \to b \text{ при } x \to x_0.$$

Число b называется пределом функции $f(M)$ при $M \to M_0$, если для любого числа $\varepsilon > 0$ можно указать такую окрестность $S_r(M_0)$ точки M_0, что для всех точек $M \in S_r(M_0) \cap V$, $M \neq M_0$ выполняется неравенство

$$|f(M) - b| < \varepsilon.$$

В частности, для функции одной переменной $y = f(x)$ число b называется пределом при $x \to x_0$, если для любого числа $\varepsilon > 0$ можно указать такое положительное число δ, что для всех $x \in V$, $x \neq x_0$ и удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство $|f(x) - b| < \varepsilon$.

○ **Примеры.**

1. $\lim\limits_{x \to 0} \cos x = 1$.

Действительно, возьмем произвольное $\varepsilon > 0$. Так как $|\cos x - 1| = 2\sin^2 \frac{x}{2} < \frac{x^2}{2}$, например, в промежутке $|x| < \pi/2$, то, положив $\delta = \min(\pi/2, \sqrt{2\varepsilon})$, получим, что для всех $x \neq 0$, удовлетворяющих условию $|x| < \delta$, выполняется неравенство $|\cos x - 1| < \varepsilon$.

2. $\lim\limits_{x \to 0} \sin \frac{1}{x}$ не существует. Действительно, рассмотрим две последовательности $\{x_k\}$ и $\{x'_k\}$, где $x_k = \frac{1}{\pi k}$, $x'_k = \frac{1}{\pi/2 + 2\pi k}$, $k = 1, 2, \ldots$, которые сходятся к нулю. Последовательность значений функции $\{f(x_k)\}$ сходится к нулю, так как $f(x_k) = \sin \pi k = 0$ при всех k. Последовательность же $\{f(x'_k)\}$ сходится к единице, так как $f(x'_k) = \sin(\pi/2 + 2\pi k) = 1$.

3. $\lim\limits_{\substack{x_1 \to 0 \\ x_2 \to 0}} \frac{x_1 \cdot x_2}{x_1^2 + x_2^2}$ не существует. Действительно, рассмотрим две последовательности точек $\{M_k(1/k, 1/k)\}$ и $\{M'_k(2/k, 1/k)\}$, сходящиеся к точке $O(0; 0)$. Последовательность значений функции $f(M_1) = f(1, 1) = 1/2$, $f(M_2) = f(1/2, 1/2) = 1/2, \ldots, f(M_k) = f(1/k, 1/k) = 1/2, \ldots$ сходится к $1/2$, а последовательность $f(M'_1) = f(2, 1) = 2/5$, $f(M'_2) = f(1, 1/2) = 2/5, \ldots, f(M'_k) = f(2/k, 1/k) = 2/5, \ldots$ сходится к $2/5$. ●

4.11. Некоторые замечательные пределы

1. $\lim\limits_{x \to 0} \frac{\sin x}{x} = 1.$ 2. $\lim\limits_{x \to 0} \frac{\operatorname{tg} x}{x} = 1.$ 3. $\lim\limits_{x \to 0} \frac{\arcsin x}{x} = 1.$

4. $\lim\limits_{x \to 0} \frac{\operatorname{arctg} x}{x} = 1.$ 5. $\lim\limits_{x \to 0} (1+x)^{1/x} = e$ $(e = 2{,}718\ldots)$.

6. $\lim\limits_{x \to 0} \frac{\log_a(1+x)}{x} = \log_a e.$ 7. $\lim\limits_{x \to 0} \frac{\ln(1+x)}{x} = 1.$

8. $\lim\limits_{x \to 0} \frac{a^x - 1}{x} = \ln a.$ 9. $\lim\limits_{x \to 0} \frac{e^x - 1}{x} = 1.$

4.12. Свойства функций, имеющих предел

Пусть функция $y = f(M)$ определена на множестве V.

1°. Если функция $y = f(M)$ имеет при $M \to M_0$ предел, то этот предел единственный.

2°. Если функция $y = f(M)$ имеет при $M \to M_0$ конечный предел, то существует окрестность $S_r(M_0)$ точки M_0 такая, что функция $f(M)$ ограничена в $S_r(M_0) \cap V$.

4.13. Предел функции при $x \to \infty$

Пусть дана функция $y = f(x)$ одной переменной x. Число b называют *пределом функции* $f(x)$ при $x \to \infty$,

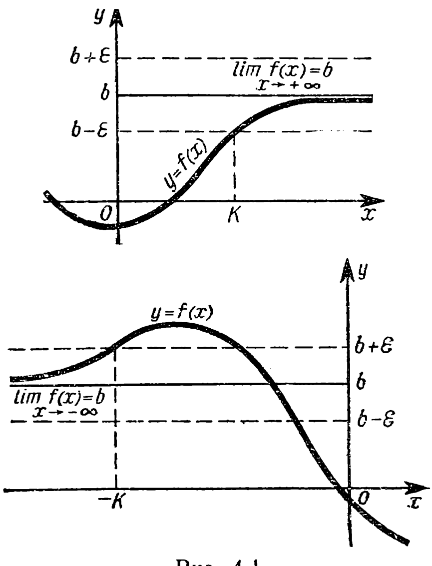


Рис. 4.1

если для любого числа $\varepsilon > 0$ можно указать такое число $K > 0$, что для всех $|x| > K$ выполняется неравенство $|f(x) - b| < \varepsilon$ (рис. 4.1).

Примеры.

1. $\lim\limits_{x \to +\infty} \frac{4}{1+(1/2)^x} = 4$, так как $\left(\frac{1}{2}\right)^x \to 0$ при $x \to +\infty$;

$\lim\limits_{x \to -\infty} \frac{4}{1+(1/2)^x} = 0$, так как $\left(\frac{1}{2}\right)^x \to +\infty$ при $x \to -\infty$.

2. $\lim\limits_{x \to \infty} \left(1 + \frac{1}{x}\right)^x = e$ (см. п. 3.12).

4.14. Односторонние пределы

Пусть функция $y = f(x)$ определена при $x < x_0$, где x_0 — предельная точка области определения функции.

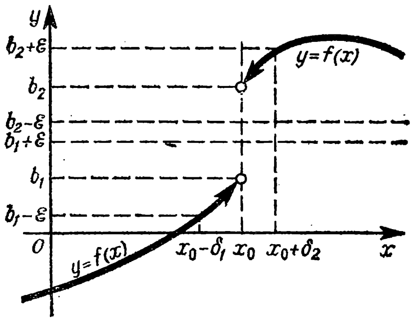


Рис. 4.2

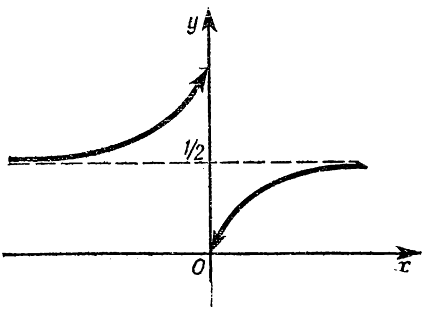


Рис. 4.3

Число $b_1 = f(x_0 - 0)$ называют *пределом слева* функции $f(x)$ при $x \to x_0$, если для любого числа $\varepsilon > 0$ можно указать такое положительное число δ_1, что для всех $x \neq x_0$, удовлетворяющих условию $x_0 - x < \delta_1$, выполняется неравенство $|f(x) - b_1| < \varepsilon$ (рис. 4.2): $f(x_0 - 0) = \lim\limits_{x \to x_0 - 0} f(x)$.

Пусть функция $y = f(x)$ определена при $x > x_0$, где x_0 — предельная точка области определения функции.

Число $b_2 = f(x_0 + 0)$ называют *пределом справа* функции $f(x)$ при $x \to x_0$, если для любого числа $\varepsilon > 0$ можно указать такое положительное число δ_2, что для всех $x \neq x_0$, удовлетворяющих условию $x - x_0 < \delta_2$, выполняется неравенство $|f(x) - b_2| < \varepsilon$ (рис. 4.2):

$$f(x_0 + 0) = \lim_{x \to x_0 + 0} f(x).$$

Функция $y=f(x)$ имеет предел в точке x_0 тогда и только тогда, когда в этой точке существуют пределы как слева, так и справа и они равны, т. е. $f(x_0-0)=f(x_0+0)$.

Например, для функции $f(x)=\dfrac{1}{1+2^{1/x}}$ предел слева при $x\to 0$ $f(-0)=1$, а предел справа при $x\to 0$ $f(+0)=0$ и, следовательно, $\lim\limits_{x\to 0} f(x)$ не существует (рис. 4.3).

4.15. Основные теоремы о пределах

1. *Если функция* $y=f(M)=C$ (C — постоянная), *то* $\lim\limits_{M\to M_0} f(M)=C$.

2. Если $\lim\limits_{M\to M_0} f(M)$ *существует, то для любого числа* k

$$\lim_{M\to M_0} k\cdot f(M)=k\cdot \lim_{M\to M_0} f(M).$$

3. *Если существуют* $\lim\limits_{M\to M_0} f(M)$ и $\lim\limits_{M\to M_0} g(M)$, то:

а) *существует* $\lim\limits_{M\to M_0} [f(M)+g(M)]$, *причем*

$$\lim_{M\to M_0} [f(M)+g(M)]=\lim_{M\to M_0} f(M)+\lim_{M\to M_0} g(M);$$

б) *существует* $\lim\limits_{M\to M_0} [f(M)\cdot g(M)]$, *причем*

$$\lim_{M\to M_0} [f(M)\cdot g(M)]=\lim_{M\to M_0} f(M)\cdot \lim_{M\to M_0} g(M);$$

в) *если* $\lim\limits_{M\to M_0} g(M)\neq 0$, *существует* $\lim\limits_{M\to M_0} \dfrac{f(M)}{g(M)}$, *причем*

$$\lim_{M\to M_0} \frac{f(M)}{g(M)}=\frac{\lim\limits_{M\to M_0} f(M)}{\lim\limits_{M\to M_0} g(M)}.$$

4. *Пусть* $f(M)\leqslant g(M)$ *в некоторой окрестности точки* M_0. *Тогда* $\lim\limits_{M\to M_0} f(M)\leqslant \lim\limits_{M\to M_0} g(M)$, *если эти пределы существуют.*

В частности, если $f(M)\leqslant 0$ ($f(M)\geqslant 0$), то и $\lim\limits_{M\to M_0} f(M)\leqslant 0$ $\left(\lim\limits_{M\to M_0} f(M)\geqslant 0\right)$.

5. *Если* $\varphi(M)\leqslant f(M)\leqslant g(M)$ *в некоторой окрестности точки* M_0 и $\lim\limits_{M\to M_0} \varphi(M)=\lim\limits_{M\to M_0} g(M)=b$, то $\lim\limits_{M\to M_0} f(M)=b$.

○ **Пример.** Имеем

$$\lim_{\substack{x_1\to 2\\ x_2\to -1}} \frac{x_1\cdot x_2}{x_1^2+x_2^2} = \frac{\lim\limits_{\substack{x_1\to 2\\ x_2\to -1}} x_1 \cdot \lim\limits_{\substack{x_1\to 2\\ x_2\to -1}} x_2}{\lim\limits_{\substack{x_1\to 2\\ x_2\to -1}} x_1^2 + \lim\limits_{\substack{x_1\to 2\\ x_2\to -1}} x_2^2} = \frac{2\cdot(-1)}{2^2+(-1)^2} = -\frac{2}{5}. \text{ ●}$$

4.16. Бесконечно малые и бесконечно большие функции

Функция $y=\alpha(M)$ называется *бесконечно малой* при $M\to M_0$, если

$$\lim_{M\to M_0} \alpha(M)=0.$$

В частности, функция $y=\alpha(x)$ называется бесконечно малой при $x\to x_0$, если

$$\lim_{x\to x_0} \alpha(x)=0.$$

○ **Примеры.**

1. Функция $\alpha(x)=\frac{x-2}{x^2}$ бесконечно малая при $x\to 2$, так как $\lim\limits_{x\to 2}\frac{x-2}{x^2}=0$. Данная функция является бесконечно малой также при $x\to\infty$, так как $\lim\limits_{x\to\infty}\frac{x-2}{x^2}=0$.

2. Функция $\alpha(x_1, x_2)=\frac{x_2-x_1}{x_2+x_1}$ бесконечно малая при стремлении точки $M(x_1, x_2)$ к любой точке прямой $x_2=x_1$, за исключением начала координат $O(0, 0)$. Функция $\alpha(x_1, x_2)$ не имеет предела при $M(x_1, x_2)\to O(0, 0)$. ●

Свойства бесконечно малых функций

1°. Предел $\lim\limits_{M\to M_0} f(M)$ существует и равен числу b тогда и только тогда, когда $f(M)=b+\alpha(M)$, где $\alpha(M)$ — бесконечно малая при $M\to M_0$.

2°. Сумма и произведение конечного числа бесконечно малых при $M\to M_0$ функций являются бесконечно малыми функциями.

3°. Произведение бесконечно малой при $M\to M_0$ функции на ограниченную в некоторой окрестности точки M_0 функцию является бесконечно малой функцией.

Функция $y=f(M)$ называется бесконечно большой при $M\to M_0$, если для любого числа $K>0$ можно указать окрестность $S_r(M_0)$ точки M_0 такую, что для всех точек $M\in S_r(M_0)$, $M\neq M_0$ выполняется неравенство $|f(M)|>K$.

В этом случае пишут $\lim\limits_{M \to M_0} f(M) = \infty$ или $f(M) \to \infty$ при $M \to M_0$.

В частности, функция одной переменной $y = f(x)$ является бесконечно большой при $x \to x_0$, если для любого числа $K > 0$ можно указать такое зависящее от K положительное число δ, что для всех $x \neq x_0$, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство $|f(x)| > K$.

○ **Примеры.**

1. Функция $\alpha(x) = \frac{x-2}{x^2}$ бесконечно большая при $x \to 0$, так как $\lim\limits_{x \to 0} \alpha(x) = -\infty$.

2. Функция $f(x) = \frac{1}{x-1}$ бесконечно большая при $x \to 1$, так как для любого $K > 0$ найдется $\delta = 1/K$ такое, что для всех $x \neq 1$, удовлетворяющих условию $|x-1| < 1/K$, выполняется неравенство $|f(x)| > K$. При этом

$$\lim_{x \to 1-0} f(x) = -\infty, \quad \lim_{x \to 1+0} f(x) = +\infty.$$

3. Функция $f(M) = \frac{1}{x_1^2 + x_2^2}$ бесконечно большая при $M \to O(0, 0)$, так как

$$\lim_{M \to (0, 0)} f(M) = +\infty. \quad ●$$

Функция $f(M)$, $f(M) \neq 0$ при $M \neq M_0$, является бесконечно большой при $M \to M_0$ тогда и только тогда, когда функция $\alpha(M) = \frac{1}{f(M)}$ является бесконечно малой при $M \to M_0$.

Функция $f(x)$ одной переменной является бесконечно большой при $x \to \infty$, если для любого числа $L > 0$ можно указать такое зависящее от L положительное число K, что для всех $|x| > K$ выполняется неравенство $|f(x)| > L$.

4.17. Сравнение функций. Эквивалентные бесконечно малые

Пусть функции $f(x)$ и $g(x)$ таковы, что $f(x) \neq 0$, $g(x) \neq 0$ при $x \neq x_0$ и существует предел $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)} = l$. Тогда:

а) если $l \neq 0$ и $l \neq \infty$, то говорят, что функции $f(x)$ и $g(x)$ одного порядка при $x \to x_0$, и пишут: $f(x) = O^*(g(x))$;

б) если $l=1$, то функции $f(x)$ и $g(x)$ называют эквивалентными при $x \to x_0$ и пишут $f(x) \backsim g(x)$;

в) если $l=0$, то функцию $f(x)$ называют функцией более высокого порядка малости по сравнению с функцией $g(x)$ при $x \to x_0$ и пишут $f(x)=o(g(x))$ (читается: «$f(x)$ есть o малое от $g(x)$ при $x \to x_0$»);

г) если $l=\infty$, то $g(x)=o(f(x))$.

○ **Примеры.**

1. $\sin^2 x = O^*(3x^2)$ при $x \to 0$, так как $\lim\limits_{x \to 0} \frac{\sin^2 x}{3x^2} = \frac{1}{3} \neq 0$ (см. п. 4.11).

2. $x = o(x^2)$ при $x \to \infty$, так как $\lim\limits_{x \to \infty} \frac{x}{x^2} = 0$. В то же время $x^2 = o(x)$ при $x \to 0$, так как $\lim\limits_{x \to 0} \frac{x^2}{x} = 0$. ●

Если $f(x) = O^*(g(x))$ при $x \to x_0$, т. е. $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)} = l \neq 0$, то $f(x) \backsim lg(x)$ и $f(x) = lg(x) + o(g(x))$ при $x \to x_0$. Если $f(x) \backsim u(x)$, а $g(x) \backsim v(x)$ при $x \to x_0$, то при условии существования $\lim\limits_{x \to x_0} \frac{u(x)}{v(x)}$ существует и $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)}$, причем $\lim\limits_{x \to x_0} \frac{f(x)}{g(x)} = \lim\limits_{x \to x_0} \frac{u(x)}{v(x)}$.

○ **Примеры.**

1. При $a > 1$ и $p > 0$ $\log_a x = o(x^p)$ и $x^p = o(a^x)$ при $x \to \infty$, т. е. логарифмическая функция растет медленнее степенной функции, которая, в свою очередь, растет медленнее показательной функции.

2. Если $\alpha(x)$ — бесконечно малая при $x \to x_0$, то имеют место следующие эквивалентности:

$$\sin \alpha(x) \backsim \alpha(x);$$
$$\operatorname{tg} \alpha(x) \backsim \alpha(x);$$
$$\arcsin \alpha(x) \backsim \alpha(x);$$
$$\operatorname{arctg} \alpha(x) \backsim \alpha(x);$$
$$\log_a [1+\alpha(x)] \backsim \alpha(x) \log_a e;$$
$$1 - \cos \alpha(x) \backsim \frac{\alpha^2(x)}{2};$$
$$a^{\alpha(x)} - 1 \backsim \alpha(x) \ln a;$$
$$[1+\alpha(x)]^n - 1 \backsim n \cdot \alpha(x);$$
$$\sqrt[n]{1+\alpha(x)} - 1 \backsim \frac{1}{n} \alpha(x).$$

Указанные эквивалентности полезно использовать при вычислении пределов функций. Например,

$$\lim_{x\to 0}\frac{e^{x^2}-1}{\sin x\left(\sqrt[3]{1-x}-1\right)}=\lim_{x\to 0}\frac{x^2}{x\cdot\frac{1}{3}(-x)}=-3.$$

$$\lim_{x\to\infty}\frac{3x^3-7x+2}{4-6x+9x^2-x^3}=\lim_{x\to\infty}\frac{3x^3+o(3x^3)}{-x^3+o(-x^3)}=\lim_{x\to\infty}\frac{3x^3}{-x^3}=-3.$$ ●

4.18. Асимптоты графика функции одной переменной

Пусть функция $y=f(x)$ определена при всех $x>x_0$ ($x<x_0$). Если существуют числа k и b такие, что функция $f(x)-kx-b=o(1)$ при $x\to+\infty$ ($x\to-\infty$), то прямую линию $y=kx+b$ называют *асимптотой графика* функции $y=f(x)$ при $x\to+\infty$ ($x\to-\infty$).

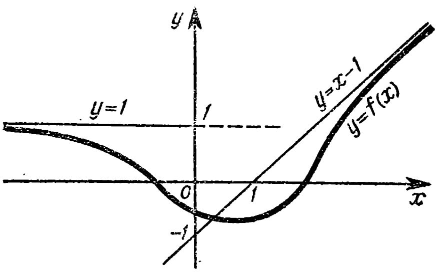


Рис. 4.4

При этом если $k\neq 0$, то асимптоту называют *наклонной*, если же $k=0$ (тогда $y=b$), то *горизонтальной*.

Условие $f(x)-kx-b=o(1)$ означает, что $\lim\limits_{\substack{x\to+\infty\\(x\to-\infty)}}[f(x)-kx-b]=0$ и, следовательно, функция $f(x)$ при $x\to+\infty$ ($x\to-\infty$) неограниченно приближается к прямой $y=kx+b$ («ведет себя почти как линейная функция»). Например, на рис. 4.4 изображен график функции, имеющий наклонную асимптоту $y=x-1$ при $x\to+\infty$ (правая наклонная асимптота) и горизонтальную асимптоту $y=1$ при $x\to-\infty$ (левая горизонтальная асимптота).

Если существуют пределы

$$\lim_{x\to+\infty}\frac{f(x)}{x}=k_1 \text{ и } \lim_{x\to+\infty}[f(x)-k_1x]=b_1,$$

то прямая $y=k_1x+b_1$ является правой наклонной (при $k_1=0$ — горизонтальной) асимптотой графика функции $y=f(x)$.

Если существуют пределы $\lim\limits_{x\to-\infty}\frac{f(x)}{x}=k_2$ и $\lim\limits_{x\to-\infty}[f(x)-$ $-kx]=b_2$, то прямая $y=k_2x+b_2$ является левой наклонной (при $k_2=0$ — горизонтальной) асимптотой графика функции $y=f(x)$.

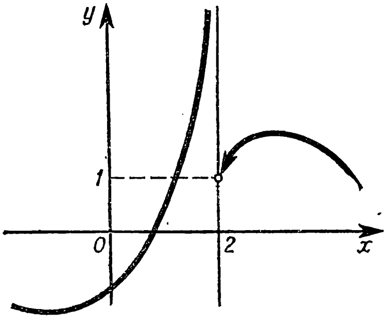


Рис. 4.5

Пусть функция $y=f(x)$ определена в некоторой окрестности точки $x=x_0$. Если $\lim\limits_{x\to x_0-0}f(x)=\infty$ или $\lim\limits_{x\to x_0+0}f(x)=\infty$, то прямая $x=x_0$ является вертикальной асимптотой графика функции $y=f(x)$.

Например, на рис. 4.5 изображен график функции $y=f(x)$, имеющий вертикальную асимптоту $x=2$, так как $\lim\limits_{x\to 2-0}f(x)=+\infty$. Предел справа $\lim\limits_{x\to 2+0}f(x)=1$.

4.19. Понятие непрерывности функции в точке

Пусть функция $f(M)=f(x_1, x_2, \ldots, x_n)$ определена на множестве $V\subseteq R^n$ и пусть точка $M_0(x_1^0; x_2^0; \ldots, x_n^0)\in V$ является его предельной точкой.

Функция $f(M)$ называется *непрерывной в точке* M_0, если для любого числа $\varepsilon>0$ можно указать окрестность $S_r(M_0)$ точки M_0 так, что для всех точек $M\in S_r(M_0)\cap V$ выполняется неравенство $|f(M)-f(M_0)|<\varepsilon$.

Непрерывность функции $f(M)$ в точке M_0 означает существование предела $\lim\limits_{M\to M_0}f(M)$ и равенство этого предела значению функции в точке M_0, т. е. $\lim\limits_{M\to M_0}f(M)=$ $=f(M_0)$.

В этом же случае для любой последовательности точек $\{M_k\}$, $M_k\in V$, сходящейся к точке M_0, последовательность значений функции $f(M_1), f(M_2), \ldots, f(M_k), \ldots$ сходится к $f(M_0)$ (см. п. 4.10).

Условие $\lim\limits_{M\to M_0}f(M)=f(M_0)$ равносильно условию $\lim\limits_{M\to M_0}[f(M)-f(M_0)]=0$. Если при этом точка M имеет координаты $(x_1; x_2; \ldots; x_n)$, то разности $x_1-x_1^0, x_2-x_2^0, \ldots$ $\ldots, x_n-x_n^0$ обозначают соответственно через $\Delta x_1, \Delta x_2, \ldots, \Delta x_n$ и называют *приращениями аргументов*, а разность $f(M)-$

$-f(M_0)$ — через $\Delta f(M_0)$ и называют *приращением функции* в точке M_0, соответствующим данным приращениям аргументов $\Delta x_1, \Delta x_2, \ldots, \Delta x_n$. Тогда условие $\lim\limits_{M \to M_0} [f(M) - f(M_0)] = 0$ может быть записано в виде $\Delta f(M_0) \to 0$, при $\Delta x_1 \to 0, \ldots, \Delta x_n \to 0$ и, следовательно, непрерывность функции $f(M)$ в точке M_0 означает, что ее приращение стремится к нулю, когда приращения всех ее аргументов также стремятся к нулю.

В частности, функция $y = f(x)$ одной переменной, определенная в некоторой окрестности точки x_0, является непрерывной в этой точке, если для любого числа $\varepsilon > 0$ можно указать зависящее от ε число $\delta > 0$ такое, что для всех x, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство $|f(x) - f(x_0)| < \varepsilon$, т. е. $\lim\limits_{x \to x_0} f(x) = f(x_0)$ или $\lim\limits_{\Delta x \to 0} \Delta f(x_0) = 0$.

○ **Примеры.**

1. Линейная функция $y = a_1x_1 + \ldots + a_nx_n$ непрерывна в любой точке $M(x_1; x_2; \ldots; x_n) \in \mathbb{R}^n$.

2. Квадратичная функция $y = a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + 2a_{23}x_2x_3 + \ldots + 2a_{n-1\,n}x_{n-1}x_n$ непрерывна в любой точке из $\mathbb{R}^n$.

3. Функция $f(x) = \sin x$ непрерывна при любом $x \in \mathbb{R}^1$. Действительно, взяв произвольно точку $x_0 \in \mathbb{R}^1$ и приращение Δx, найдем, что $\Delta f(x_0) = \sin(x_0 + \Delta x) - \sin x_0 = 2\sin\frac{\Delta x}{2}\cos\left(x_0 + \frac{\Delta x}{2}\right)$, откуда

$$\lim_{\Delta x \to 0} \Delta f(x_0) = \lim_{\Delta x \to 0} 2\sin\frac{\Delta x}{2}\cos\left(x_0 + \frac{\Delta x}{2}\right) = 0. \quad ●$$

4.20. Свойства функций, непрерывных в точке

1°. Если функция $f(M)$ и функция $g(M)$ непрерывны в точке M_0, то:

а) функция $f(M) + g(M)$ непрерывна в точке M_0;

б) функция $k \cdot f(M)$, где k — постоянная, непрерывна в точке M_0;

в) функция $f(M) \cdot g(M)$ непрерывна в точке M_0;

г) функция $\frac{f(M)}{g(M)}$ непрерывна в точке M_0, если $g(M_0) \neq 0$.

2°. Если функция $f(M)$ определена на множестве V, непрерывна в точке $M_0 \in V$ и $f(M_0) > 0$ $(f(M_0) < 0)$, то существует окрестность $S_r(M_0)$ точки M_0 такая, что $f(M) > 0$ $(f(M) < 0)$ для всех точек $M \in S_r(M_0) \cap V$.

4.21. Свойства функций, непрерывных на множестве

Функция $f(M)$ называется *непрерывной на множестве* V, если она непрерывна в каждой точке этого множества.

1°. Если функция $f(M)$ непрерывна на ограниченном замкнутом множестве V, то она ограничена на этом множестве.

2°. Если функция $f(M)$ непрерывна на ограниченном замкнутом множестве V, то она достигает на этом множестве как наименьшего, так и наибольшего своих значений, т. е. на множестве V найдутся точки M_1 и M_2 такие, что $f(M_1)=\inf\limits_V\{f(M)\}$ и $f(M_2)=\sup\limits_V\{f(M)\}$.

В частности, для функций одной переменной справедливы следующие утверждения:

а) если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она ограничена на этом отрезке, т. е. существует число d такое, что $|f(x)|\leqslant d$ для всех $x\in[a, b]$;

б) если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она достигает на этом отрезке как наименьшего l, так и наибольшего L своих значений, т. е. найдутся точки $x_1\in[a, b]$ и $x_2\in[a, b]$ такие, что $f(x_1)=l=\inf\limits_{[a, b]}\{f(x)\}$, $f(x_2)=L=\sup\limits_{[a, b]}\{f(x)\}$.

Кроме указанных свойств для функций одной переменной имеют место следующие свойства:

3°. Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и на концах этого отрезка принимает значения разных знаков $(f(a)\cdot f(b)<0)$, то найдется хотя бы одна точка $c\,(a<c<b)$ такая, что $f(c)=0$.

4°. Если функция $f(x)$ непрерывна на отрезке $[a, b]$, то она принимает хотя бы по одному разу все промежуточные значения от наименьшего l до наибольшего L, т. е. для любого числа μ, заключенного между l и L $(l\leqslant\mu\leqslant L)$, найдется точка $c\in[a, b]$ такая, что $f(c)=\mu$.

4.22. Непрерывность сложной функции

Пусть задана функция $y=f(u_1, u_2, \ldots, u_m)$, определенная на множестве $W\subseteq\mathbf{R}^m$. Пусть, кроме того, каждая из функций $u_1=g_1(x_1, x_2, \ldots, x_n)$, $u_2=g_2(x_1, x_2, \ldots$ $\ldots, x_n)$, $\ldots$, $u_m=g_m(x_1, x_2, \ldots, x_n)$ определена на множестве $V\subseteq\mathbf{R}^n$ (см. п. 4.4).

Если функции $g_1, g_2, \ldots, g_m$ непрерывны в точке $M_0(x_1^0; x_2^0; \ldots, x_n^0)\in V$, а функция $y=f(u_1, u_2, \ldots, u_m)$

непрерывна в соответствующей точке $P_0(u_1^0;\ u_2^0;\ \ldots,\ u_m^0)$, где $u_1^0=g_1(M_0)$, $u_2^0=g_2(M_0)$, $\ldots$, $u_m^0=g_m(M_0)$, то сложная функция

$$y=f[g_1(x_1,\ x_2,\ \ldots,\ x_n),\ g_2(x_1,\ x_2,\ \ldots,\ x_n),\ \ldots$$
$$\ldots,\ g_m(x_1,\ x_2,\ \ldots,\ x_n)]$$

непрерывна в точке M_0.

В частности, если функция $u=g(x)$ одной переменной x непрерывна в точке x_0, а функция $y=f(u)$ одной переменной u непрерывна в точке $u_0=g(x_0)$, то сложная функция $y=f[g(x)]$ непрерывна в точке x_0, т. е. $\lim\limits_{u\to u_0} f(u)=$ $=\lim\limits_{x\to x_0} f[g(x)]=f[\lim\limits_{x\to x_0} g(x)]$.

Последняя формула показывает, что, с одной стороны, операция предельного перехода перестановочна с операцией взятия непрерывной функции (правое равенство), а с другой стороны, дает правило замены переменной при вычислении пределов непрерывных функций (левое равенство).

○ **Примеры.**

1. $\lim\limits_{x\to+\infty} x[\ln(x+1)-\ln x]=\lim\limits_{x\to+\infty} x\cdot\ln\dfrac{x+1}{x}=$
$=\lim\limits_{x\to+\infty}\ln\left(1+\dfrac{1}{x}\right)^x=\ln\left[\lim\limits_{x\to+\infty}\left(1+\dfrac{1}{x}\right)^x\right]=\ln e=1$
(см. п. 4.13).

2. $\lim\limits_{x\to 1}\dfrac{\sin\pi x}{x^2-1}=\left|\begin{matrix} x=t+1 \\ x\to 1\Rightarrow t\to 0\end{matrix}\right|=\lim\limits_{t\to 0}\dfrac{\sin(\pi t+\pi)}{t^2+2t+1-1}=$
$=\lim\limits_{t\to 0}\dfrac{-\sin\pi t}{t(t+2)}=-\lim\limits_{t\to 0}\left(\dfrac{\sin\pi t}{\pi t}\cdot\dfrac{\pi}{t+2}\right)=(-1)\cdot\dfrac{\pi}{2}=-\dfrac{\pi}{2}$. ●

4.23. Односторонняя непрерывность

Пусть функция $y=f(x)$ одной переменной x определена при $x\leqslant x_0$ $(x\geqslant x_0)$. Функция $f(x)$ называется *непрерывной слева* (*справа*) в точке x_0, если

$$\lim_{x\to x_0-0} f(x)=f(x_0)\quad \left(\lim_{x\to x_0+0} f(x)=f(x_0)\right).$$

○ **Примеры.**

1. Функция $f(x)=\begin{cases} e^{1/x}, & x<0, \\ 0, & x=0,\end{cases}$ непрерывна слева в точке x_0, так как

$$\lim_{x\to-0} f(x)=\lim_{x\to-0} e^{1/x}=0=f(0).$$

2. Функция $f(x)$ является непрерывной на отрезке $[a, b]$, если она непрерывна в каждой внутренней точке этого отрезка, непрерывна справа в точке $x=a$ и непрерывна слева в точке $x=b$ (см. п. 4.21). ●

4.24. Непрерывность обратной функции

Если функция одной переменной $y=f(x)$ строго монотонна и непрерывна на отрезке $[a, b]$, то обратная функция $x=g(y)$ определена, строго монотонна и непрерывна на отрезке с концами в точках $f(a)$ и $f(b)$.

Если функция одной переменной $y=f(x)$ строго монотонна и непрерывна на интервале $]a, b[$ (конечном или бесконечном) и если существуют (конечные или бесконечные) односторонние пределы $c=\lim\limits_{x\to a+0} f(x)$ и $d=\lim\limits_{x\to b-0} f(x)$, то обратная функция $x=g(y)$ определена, строго монотонна и непрерывна на интервале $]c, d[$.

4.25. Точки разрыва функции

Пусть функция одной переменной $y=f(x)$ определена в некоторой окрестности точки x_0, за исключением, быть может, самой точки x_0. Если функция $f(x)$ не является непрерывной в точке x_0, то говорят, что в точке x_0 функция *терпит разрыв*, и точку x_0 называют *точкой разрыва функции*.

Точку x_0 называют *точкой разрыва первого рода*, если существуют конечные односторонние пределы $f(x_0-0)=$ $=\lim\limits_{x\to x_0-0} f(x)$ и $f(x_0+0)=\lim\limits_{x\to x_0+0} f(x)$, но $f(x_0-0)\neq f(x_0+0)$. В этом случае наибольшую из разностей между числами $f(x_0)$, $f(x_0-0)$, $f(x_0+0)$ называют *скачком функции* $f(x)$ в точке x_0. Например, для функции $f(x)=\frac{1}{1+2^{1/x}}$ точка $x_0=0$ является точкой разрыва, так как в этой точке функция не определена ($f(0)$ не существует). При этом $f(-0)=\lim\limits_{x\to -0} f(x)=1$, $f(+0)=\lim\limits_{x\to +0} f(x)=0$. Следовательно, точка $x_0=0$ является точкой разрыва первого рода, а разность $f(-0)-f(+0)=1$ — скачком данной функции (см. рис. 4.3).

Точку x_0 называют *точкой устранимого разрыва*, если конечные односторонние пределы $f(x_0-0)$ и $f(x_0+0)$ равны между собой, но не совпадают со значением $f(x_0)$, если только оно существует.

Например, для функции $f(x)=\begin{cases} x^2 & \text{при } x \neq 2, \\ 1 & \text{при } x=2 \end{cases}$ имеем $f(2-0)=f(2+0)=\lim\limits_{x\to 2} x^2=4$, однако $4=f(2-0)=$ $=f(2+0)\neq f(2)=1$. Следовательно, точка $x=2$ является точкой устранимого разрыва функции $f(x)$ (рис. 4.6).

Термин «устранимый разрыв» оправдан тем, что достаточно доопределить или переопределить функцию в точке x_0 для того, чтобы она стала непрерывной в этой точке.

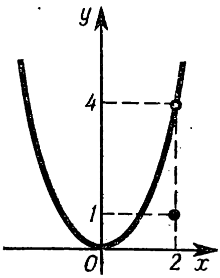


Рис. 4.6

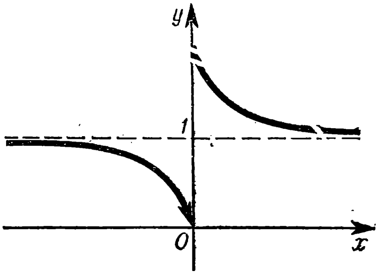


Рис. 4.7

В рассмотренном примере надо положить $g(2)=4=\lim\limits_{x\to 2} f(x)$; тогда функция $g(x)=\begin{cases} x^2 & \text{при } x \neq 2, \\ 4 & \text{при } x=2 \end{cases}$ является непрерывной в точке $x_0=2$.

Точку x_0 называют *точкой разрыва второго рода*, если хотя бы один из односторонних пределов $f(x_0-0)$ и $f(x_0+0)$ не существует (в частности, бесконечен). Например, для функции $f(x)=e^{1/x}$ точка $x_0=0$ является точкой разрыва второго рода, так как $f(-0)=0$, $f(+0)=+\infty$ (рис. 4.7).

Замечание. Функция n переменных $y=f(x_1, x_2, \ldots$ $\ldots, x_n)$ может иметь не только изолированные точки разрыва, а целые множества точек разрывов (линии, поверхности разрывов).

Например, функция $f(x_1, x_2)=\dfrac{x_1^2+x_2^2}{(x_1^2-x_2)\cdot(x_1+3x_2)}$ имеет разрыв во всех точках параболы $x_2=x_1^2$ и во всех точках прямой $x_2=-\dfrac{1}{3}x_1$.

РАЗДЕЛ V. ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ

5.1. Производная

Пусть функция $y=f(x)$ определена в некоторой окрестности точки x.

Первой производной (*производной первого порядка*) функции $f(x)$ в точке x называют конечный предел отношения приращения функции $\Delta y=\Delta f(x)$ к приращению аргумента Δx при условии, что Δx стремится к нулю. Обозначения производной:

$$f'(x),\ y'_x,\ \dot{y}_x,\ \frac{df(x)}{dx},\ \frac{dy}{dx}.$$

Таким образом, $f'(x)=\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim\limits_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$.

Если в некоторой точке x $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\infty\ (+\infty,\ -\infty)$ и функция $f(x)$ непрерывна в точке x, то говорят о наличии у этой функции в точке x «бесконечной производной» $f'(x)=\infty\ (+\infty,\ -\infty)$.

Конечные или бесконечные пределы

$$f'_-(x)=\lim_{\Delta x\to -0}\frac{f(x+\Delta x)-f(x)}{\Delta x} \text{ и } f'_+(x)=\lim_{\Delta x\to +0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

называют соответственно *левой* и *правой производными* функции $f(x)$ в точке x.

Функция $f(x)$ имеет в точке x производную $f'(x)$ тогда и только тогда, когда односторонние производные $f'_-(x)$ и $f'_+(x)$ существуют и совпадают, т. е. $f'(x)=f'_-(x)=f'_+(x)$.

Операцию нахождения производной $f'(x)$ называют *операцией дифференцирования* функции $f(x)$.

Примеры.

1. Функция $f(x)=x^2$ имеет конечную производную при любом действительном x. Действительно, при любом x

имеем

$$f'(x)=\lim_{\Delta x\to 0}\frac{(x+\Delta x)^2-x^2}{\Delta x}=\lim_{\Delta x\to 0}\frac{2x\,\Delta x+(\Delta x)^2}{\Delta x}=$$
$$=\lim_{\Delta x\to 0}(2x+\Delta x)=2x.$$

2. Функция $f(x)=\sqrt[3]{x}$ имеет в точке $x=0$ бесконечную производную. В самом деле,

$$f'(0)=\lim_{\Delta x\to 0}\frac{\sqrt[3]{0+\Delta x}-0}{\Delta x}=\lim_{\Delta x\to 0}\frac{1}{\sqrt[3]{(\Delta x)^2}}=+\infty.$$

3. Функция $f(x)=e^{|x|}$ не имеет в точке $x=0$ производной, хотя в этой точке существуют конечные односторонние производные. В самом деле,

$$f'_+(0)=\lim_{\Delta x\to +0}\frac{e^{|0+\Delta x|}-1}{\Delta x}=\lim_{\Delta x\to +0}\frac{e^{\Delta x}-1}{\Delta x}=1;$$
$$f'_-(0)=\lim_{\Delta x\to -0}\frac{e^{|0+\Delta x|}-1}{\Delta x}=\lim_{\Delta x\to -0}\frac{e^{-\Delta x}-1}{\Delta x}=-1$$

(см. п. 4.11), но $f'_-(0)\neq f'_+(0)$. ●

5.2. Дифференцируемость и дифференциал функции

Дифференциалом dx независимой переменной x называют ее приращение $\Delta x\,(dx=\Delta x)$.

Функция $y=f(x)$ называется дифференцируемой в точке x, если в этой точке ее приращение $\Delta y=f(x+dx)-f(x)$ представимо в виде $\Delta y=A(x)\,dx+\alpha(dx)$, где $A(x)$ — постоянная, а $\alpha(dx)=o(dx)$ при $dx\to 0$, т. е. является бесконечно малой высшего порядка малости по сравнению с dx.

Главную линейную относительно dx часть приращения $A(x)\,dx$ называют *первым дифференциалом* (*дифференциалом первого порядка*) функции $f(x)$ в точке x и обозначают $df(x)$ или dy. Таким образом, $dy=A(x)\,dx$, так что $\Delta y=dy+o(dx)$.

Если функция $f(x)$ дифференцируема в точке x, то она и непрерывна в этой точке.

Функция $f(x)$ дифференцируема в точке x тогда и только тогда, когда в этой точке существует конечная производная $f'(x)$. При этом $A(x)=f'(x)$, так что $dy=f'(x)\,dx$.

Из равенства $\Delta y=dy+o(dx)$ при условии $f'(x)\neq 0$ следует, что при достаточно малых dx справедливо приближенное равенство $\Delta y \approx dy$ или $f(x+dx) \approx f(x) + f'(x)\,dx$, используемое в приближенных вычислениях.

Если функция $f(x)$ дифференцируема в каждой точке некоторого промежутка, то говорят о дифференцируемости функции на этом промежутке. Если, кроме того, производная $f'(x)$ непрерывна на данном промежутке, то говорят, что функция $f(x)$ *непрерывно дифференцируема на этом промежутке.*

○ **Примеры.**

1. Функция $y = x^2$ дифференцируема при любом x, так как $\Delta y = (x + dx)^2 - x^2 = 2x dx + (dx)^2 = 2x \cdot dx + o(dx)$. При этом $dy = 2x dx$.

2. Вычислить приближенно $\sqrt{40}$.

Рассмотрим функцию $f(x) = \sqrt{x}$. Ее производная $f'(x) = \frac{1}{2\sqrt{x}}$. Пусть $x = 36$, $x + dx = 40$. Тогда $dx = (x + dx) - x = 4$; $f(x) = \sqrt{36}$; $f'(x) = \frac{1}{2\sqrt{36}} = \frac{1}{12}$. Отсюда $f(x + dx) = \sqrt{40} \approx 6 + \frac{1}{12} \cdot 4 = 6\frac{1}{3} = 6{,}33\ldots$ ●

5.3. Геометрический смысл производной и дифференциала

Касательной к графику функции $y = f(x)$ в точке $M(x_0; f(x_0))$ называют предельное положение секущей MN при произвольном стремлении точки N к точке M по графику функции (или, что то же самое, при $dx \to 0$) (рис. 5.1).

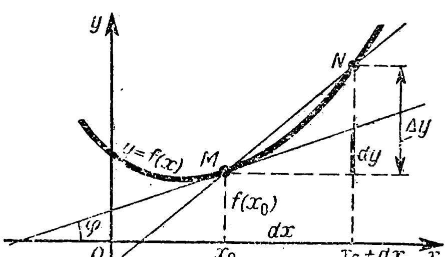


Рис. 5.1

Значение производной $f'(x_0)$ в точке x_0 определяет угловой коэффициент касательной, проведенной к графику функции $f(x)$ в точке $M(x_0; f(x_0))$, т. е. $f'(x_0) = \operatorname{tg} \varphi$, где φ — угол между положительным направлением оси Ox

и касательной, отсчитываемый против часовой стрелки (рис. 5.1).

Уравнение касательной к графику функции $y=f(x)$ в точке $M(x_0;\ f(x_0))$ имеет вид $y-f(x_0)=f'(x_0)(x-x_0)$.

Если $f'(x_0)=\infty\ (-\infty,\ +\infty)$, то касательная к графику непрерывной функции $f(x)$ в точке $M(x_0;\ f(x_0))$ перпендикулярна оси Ox (вертикальная касательная).

Уравнение такой касательной имеет вид $x=x_0$.

Величина дифференциала dy в точке x_0 равна приращению ординаты касательной к графику $f(x)$ в точке

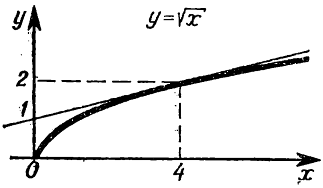


Рис. 5.2

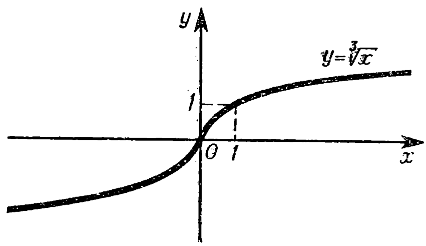


Рис. 5.3

$M(x_0;\ f(x_0))$ при переходе от точки x_0 к точке (x_0+dx) (рис. 5.1).

○ **Примеры.**

1. Написать уравнение касательной к графику функции $f(x)=\sqrt{x}$ в точке с абсциссой $x_0=4$.

Имеем $f(x_0)=\sqrt{4}=2$, $f'(x)=1/2\sqrt{x}$, $f'(x_0)=1/2\sqrt{4}=$ $=1/4$. Следовательно, уравнение касательной имеет вид $y-2=\frac{1}{4}(x-4)$ или $y=\frac{1}{4}x+1$ (рис. 5.2).

2. Касательная к графику функции $y=\sqrt[3]{x}$ в точке $O(0;\ 0)$ будет вертикальной, так как данная функция непрерывна при $x=0$, а $f'(0)=+\infty$ (см. п. 5.1) (рис. 5.3). ●

5.4. Физический смысл производной и дифференциала

В каждой точке, где функция $y=f(x)$ имеет конечную производную $f'(x)$, последняя может быть интерпретирована как мера скорости изменения y относительно x. Замена приращения функции ее дифференциалом позволяет считать процесс изменения зависимой переменной «в малом» линейным относительно изменения аргумента.

○ **Примеры.**

1. Если $s=s(t)$ — закон движения материальной точки, определяющий зависимость пути s от времени t, то производная $v=\frac{ds}{dt}$ определяет мгновенную скорость материальной точки в момент времени t. Дифференциал $ds=v\cdot dt$ определяет путь, который прошла бы материальная точка, двигаясь равномерно с мгновенной скоростью v в момент времени t, за промежуток времени от момента t до $(t+dt)$.

2. Если $q=q(t)$ — закон, определяющий зависимость количества электричества, протекающего через поперечное сечение проводника, от времени t, то производная $I=\frac{dq}{dt}$ определяет силу тока в момент времени t. Дифференциал $dq=I\cdot dt$ определяет количество электричества, которое могло бы пройти через поперечное сечение проводника при постоянной силе тока I в момент времени t за промежуток времени dt. ●

5.5. Приложения производной к экономике

В практике экономических исследований широкое применение получили *производственные функции*, используемые для установления зависимостей выпуска продукции от затрат ресурсов, при прогнозировании развития отраслей, при решении оптимизационных задач. Например, производственная функция Кобба—Дугласа связывает выпуск y с величиной производственных фондов K и затратами живого труда L:

$$y=qK^{\alpha}L^{1-\alpha},$$

где q и α — постоянные, т. е. является функцией двух переменных K и L (см. п. 4.1).

В предположении дифференцируемости производственных функций важное значение приобретают их дифференциальные характеристики, связанные с понятием производной. Так, например, если производственная функция $y=f(x)$ устанавливает зависимость выпуска продукции y от затрат ресурса x, то $f'(x)$ называют *предельным продуктом*, если же $y=f(x)$ устанавливает зависимость издержек производства y от объема продукции x, то $f'(x)$ называют *предельными издержками*.

Характеристикой относительного изменения прироста функции $y=f(x)$ при малых относительных изменениях прироста аргумента x является *эластичность функции*. Коэффициент эластичности ε определяется по формуле

$$\varepsilon=\frac{dy}{y}:\frac{dx}{x}, \quad \text{или} \quad \varepsilon=y'\cdot\frac{x}{y}.$$

Коэффициент эластичности широко используют в исследованиях потребительского спроса на товары в зависимости от цен этих товаров или доходов потребителей. Высокий коэффициент эластичности означает слабую степень удовлетворения потребности; низкий указывает на большую настоятельность данной потребности.

Если производственная функция устанавливает зависимость выпуска y от n производственных факторов x_1, x_2, ..., x_n в виде $y=f(x_1;\ x_2;\ \ldots;\ x_n)$ (см. п. 4.1), то наиболее важными дифференциальными характеристиками такой функции являются:

а) $\frac{\partial y}{\partial x_i}$ — предельная эффективность фактора x_i;

б) $\frac{\partial y}{\partial x_i}\cdot\frac{x_i}{y}$ — эластичность выпуска y относительно фактора x_i;

в) $\frac{\partial y}{\partial x_i}:\frac{\partial y}{\partial x_j}$ — предельная норма замены факторов x_j и x_i;

г) $\dfrac{d\left(\frac{x_i}{x_j}\right)}{d\left(\frac{\partial y}{\partial x_i}:\frac{\partial y}{\partial x_j}\right)}\cdot\dfrac{\frac{\partial y}{\partial x_i}:\frac{\partial y}{\partial x_j}}{\frac{x_i}{x_j}}$ — эластичность замены факторов x_j и x_i (см. п. 6.1).

Теоретический и практический интерес представляют производственные функции с постоянной (отличной от единицы) эластичностью замещения труда производственными фондами и с постоянной (переменной) отдачей на единицу масштаба производства.

Примером такого рода функций является функция CES (Canstanta Elastisity of Substantion) $y=C_0[CL^{-\rho}+(1-C)\cdot K^{\rho}]^{-1/\rho}$, для которой эластичность замещения равна $\frac{1}{1-\rho}\neq 1$; ρ, C_0 и C — постоянные.

5.6. Правила вычисления производных и дифференциалов

Пусть функции $f(x)$ и $g(x)$ дифференцируемы в точке x и k — постоянная. Тогда:

1°. $[f(x)\pm g(x)]'=f'(x)\pm g'(x)$; $d[f(x)\pm g(x)]=df(x)\pm dg(x)$.

2^0. $[k\cdot f(x)]'=k\cdot f'(x)$; $\mathrm{d}\,[k\cdot f(x)]=k\cdot \mathrm{d}f(x)$.

3^0. $[f(x)\cdot g(x)]'=f'(x)\,g(x)+f(x)\,g'(x)$;

$\mathrm{d}\,(f(x)\,g(x))=\mathrm{d}f(x)\,g(x)+f(x)\,\mathrm{d}g(x)$.

4^0. $\left[\dfrac{f(x)}{g(x)}\right]'=\dfrac{f'(x)\,g(x)-f(x)\,g'(x)}{g^2(x)}$;

$\mathrm{d}\left[\dfrac{f(x)}{g(x)}\right]=\dfrac{\mathrm{d}f(x)\,g(x)-f(x)\,\mathrm{d}g(x)}{g^2(x)}\quad (g(x)\neq 0)$.

5.7. Таблица производных

$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
C (постоянная)	0	$\sin x$	$\cos x$
x	1	$\cos x$	$-\sin x$
$\sqrt{x}$	$\dfrac{1}{2\sqrt{x}}$	$\operatorname{tg} x$	$\dfrac{1}{\cos^2 x}$
x^α	$\alpha\cdot x^{\alpha-1}$	$\operatorname{ctg} x$	$-\dfrac{1}{\sin^2 x}$
a^x	$a^x\cdot\ln a$	$\arcsin x$	$\dfrac{1}{\sqrt{1-x^2}}$
e^x	e^x	$\arccos x$	$-\dfrac{1}{\sqrt{1-x^2}}$
$\log_a\lvert x\rvert$	$\dfrac{1}{x\ln a}$	$\operatorname{arctg} x$	$\dfrac{1}{1+x^2}$
$\ln\lvert x\rvert$	$\dfrac{1}{x}$	$\operatorname{arcctg} x$	$-\dfrac{1}{1+x^2}$

○ **Примеры.**

1. Вычислить производную функции

$$y=\frac{e^x+4x^3}{\ln x}.$$

Применяя правила (см. п. 5.6) и таблицу производных, имеем

$$y'=\frac{(e^x+4x^3)'\cdot\ln x-(e^x+4x^3)\cdot(\ln x)'}{\ln^2 x}=\frac{(e^x+12x^2)\ln x-\dfrac{e^x+4x^3}{x}}{\ln^2 x}.$$

2. Для функции $y=a^x\cdot\operatorname{arctg} x$ ее производная

$$y'=(a^x)'\cdot\operatorname{arctg} x+a^x\cdot(\operatorname{arctg} x)'=a^x\cdot\ln a\cdot\operatorname{arctg} x+\frac{a^x}{1+x^2}.$$ ●

5.8. Производная и дифференциал сложной функции

Если функция $u=g(x)$ дифференцируема в точке x_0, а функция $y=f(u)$ дифференцируема в соответствующей точке $u_0 = g(x_0)$, то сложная функция $y = \Phi(x) = f[g(x)]$ также дифференцируема в точке x_0, причем

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx} \quad \text{или} \quad \Phi'(x_0) = f'(u_0) \cdot g'(x_0).$$

Справедливо равенство $dy = \Phi'(x_0) \cdot dx = f'(u_0) \cdot du$, где $du = g'(x_0)\,dx$, т. е. дифференциал равен произведению производной по некоторой переменной на дифференциал этой переменной независимо от того, является ли эта переменная независимой или функцией другой переменной (инвариантность формы первого дифференциала).

○ **Примеры.**

1. Производная функции $y = 3^{\cos^5 2x}$ равна

$$y' = 3^{\cos^5 2x} \cdot \ln 3 \cdot (\cos^5 2x)' = 3^{\cos^5 2x} \cdot \ln 3 \cdot 5 \cdot \cos^4 2x \times$$
$$\times (\cos 2x)' = 5 \ln 3 \cdot 3^{\cos^5 2x} \cdot \cos^4 2x \cdot (-\sin 2x) \cdot (2x)' =$$
$$= -10 \ln 3 \cdot \sin 2x \cdot \cos^4 2x \cdot 3^{\cos^5 2x}.$$

2. Дифференциал функции $y = \mathrm{tg}^4\, 6x$ равен

$$dy = 4\, \mathrm{tg}^3\, 6x \cdot d(\mathrm{tg}\, 6x) = 4\, \mathrm{tg}^3\, 6x \cdot \frac{1}{\cos^2 6x} \cdot d(6x) =$$
$$= \frac{24\, \mathrm{tg}^3\, 6x}{\cos^2 6x}\, dx = \frac{24 \cdot \sin^3 6x}{\cos^5 6x}\, dx.$$

3. Вычислить производную функции $y = \ln \sqrt[4]{\frac{1+x^2}{1-x}}$ $(x < 1)$.

Целесообразно предварительно преобразовать выражение этой функции к виду

$$y = \frac{1}{4} \ln(1 + x^2) - \frac{1}{4} \ln(1 - x).$$

Тогда

$$y' = \frac{1}{4} \cdot \frac{2x}{1+x^2} - \frac{1}{4} \cdot \frac{-1}{1-x} = \frac{2x + 1 - x^2}{4(1-x)(1+x^2)}.$$ ●

5.9. Логарифмическое дифференцирование

Прием логарифмического дифференцирования используют в том случае, когда функция имеет вид, удобный для логарифмирования, и сводится к следующей схеме:

а) функцию y заменяют на функцию $|y|$;

б) выражение $|y|$ логарифмируют;

в) находят производную от $\ln|y|$ $\left((\ln|y|)'=\frac{y'}{y}\right.$ по правилу нахождения производной сложной функции$\left.\right)$;

г) находят y'.

○ **Примеры.**

1. Для функции $y=(\cos x)^{\sin x}$ $(\cos x>0)$ имеем $|y|=y$ и, следовательно, $\ln y=\sin x\cdot\ln(\cos x)$. Тогда

$$\frac{y'}{y}=\cos x\cdot\ln(\cos x)+\sin x\cdot\frac{-\sin x}{\cos x},$$

откуда

$$y'=(\cos x)^{\sin x}\cdot\left[\cos x\cdot\ln(\cos x)-\frac{\sin^2 x}{\cos x}\right].$$

2. Для функции $y=\sqrt[5]{\frac{\sin 5x}{1-\sin 5x}}$ имеем

$$\ln|y|=\frac{1}{5}\ln|\sin 5x|-\frac{1}{5}\ln|1-\sin 5x|.$$

Тогда $\frac{y'}{y}=\frac{1}{5}\cdot\frac{\cos 5x}{\sin 5x}\cdot 5-\frac{1}{5}\cdot\frac{-5\cos 5x}{1-\sin 5x}=\frac{\operatorname{ctg} 5x}{1-\sin 5x}$,

откуда

$$y'=\sqrt[5]{\frac{\sin 5x}{1-\sin 5x}}\cdot\frac{\operatorname{ctg} 5x}{1-\sin 5x}.$$ ●

5.10. Производные и дифференциалы высших порядков

Если для функции $y=f(x)$ определена производная $y^{(n-1)}$ порядка $(n-1)$, то производную $y^{(n)}$ порядка n (при условии ее существования) определяют как производную от производной порядка $(n-1)$, т. е. $y^{(n)}=(y^{(n-1)})'$.

В частности, $y''=(y')'$ — производная второго порядка, $y'''=(y'')'$ — производная третьего порядка и т. д. Другие обозначения производных высших порядков:

$$\frac{d^n y}{dx^n},\ y^{\mathrm{IV}},\ y''_{xx},\ \ddot{y},\ f^{(n)}(x).$$

При вычислении производных высших порядков используют те же правила, что и для вычисления y'. Например, если $y=e^{x^2}$, то

$$y'=e^{x^2}\cdot 2x,$$
$$y''=(e^{x^2})'\cdot 2x+e^{x^2}\cdot(2x)'=$$
$$=e^{x^2}\cdot 2x\cdot 2x+e^{x^2}\cdot 2=2e^{x^2}\cdot(2x^2+1).$$

Дифференциалы высших порядков функции $y=f(u)$ последовательно определяют таким образом:

$d^2y=d(dy)$ — дифференциал второго порядка,

$d^3y=d(d^2y)$ — дифференциал третьего порядка, ...

Вообще, $d^ny=d(d^{n-1}y)$ — дифференциал n-го порядка.

Таблица производных высших порядков некоторых функций

$f(x)$	$f^{(n)}(x)$
x^{α}	$\alpha(\alpha-1)\cdot(\alpha-2)\cdot\ldots\cdot(\alpha-n+1)\cdot x^{\alpha-n}$
e^x	e^x
e^{kx}	$k^n\cdot e^{kx}$
a^{kx}	$(k\cdot\ln a)^n\cdot a^{kx}$
$\ln x$	$(-1)^{n-1}\cdot\dfrac{(n-1)!}{x^n}$
$\log_a x$	$(-1)^{n-1}\cdot\dfrac{(n-1)!}{x^n\cdot\ln a}$
$\sin kx$	$k^n\cdot\sin\left(kx+n\cdot\dfrac{\pi}{2}\right)$
$\cos kx$	$k^n\cos\left(kx+n\dfrac{\pi}{2}\right)$

При этом если $y=f(u)$ и u — независимая переменная или линейная функция $u=kx+b$ переменной x, то $d^2y=y''(du)^2$; $d^3y=y'''(du)^3$; ...; $d^ny=y^{(n)}\cdot(du)^n$.

Если же $y=f(u)$, где $u=g(x)\neq kx+b$, то $d^2y=f''(u)\cdot(du)^2+f'(u)\cdot d^2u$ и т. д. (свойство инвариантности формы не выполняется). Например, для функции $y=3u^5-4u^2+7$ ее первый дифференциал $dy=(15u^4-8u)\,du$ независимо от того, является ли u независимой переменной или функцией другой переменной;

$d^2y=(60u^3-8)(du)^2$, если u — независимая переменная;

$d^2y=d(15u^4-8u)\,du+(15u^4-8u)\,d(du)=(60u^3-8)\times\times(du)^2+(15u^4-8u)\,d^2u$, если u — функция другой переменной.

5.11. Производная обратной функции

Если дифференцируемая функция $y=f(x)$ $(a<x<b)$ имеет непрерывную обратную функцию $x=g(y)$ и $y'_x\neq 0$, то существует производная обратной функции x'_y и имеет

место равенство

$$x'_y = \frac{1}{y'_x}.$$

Дифференцируя последнее равенство по y и предполагая существование y''_{xx}, найдем

$$x''_{yy} = -\frac{y''_{xx}}{(y'_x)^2} x'_y = -\frac{y''_{xx}}{(y'_x)^3}.$$

При соответствующих предположениях аналогично можно определить производные любого порядка обратной функции.

Например, для функции $y = a^x$ $(a > 0,\ a \neq 1,\ y > 0)$ обратной является функция $x = \log_a y$. Ее производная

$$x'_y = (\log_a y)' = \frac{1}{y'_x} = \frac{1}{(a^x)'_x} = \frac{1}{a^x \ln a} = \frac{1}{y \ln a}.$$

Кроме того, так как $y''_{xx} = a^x (\ln a)^2$, то

$$x''_{yy} = -\frac{a^x (\ln a)^2}{(a^x \ln a)^3} = -\frac{1}{y^2 \ln a}.$$

5.12. Производная параметрически заданной функции

Рассмотрим функцию $y = f(x)$. Систему соотношений $x = \varphi(t)$, $y = \psi(t)$, где $\alpha < t < \beta$, называют параметрическим представлением функции $y = f(x)$, если $\psi(t) = f[\varphi(t)]$ для всех $t \in]\alpha, \beta[$. Переменная t называется в этом случае параметром. Если функции $\varphi(t)$ и $\psi(t)$ — дифференцируемые и $\varphi'(t) \neq 0$, то существует производная y'_x параметрически заданной функции и

$$y'_x = \frac{y'_t}{x'_t} = \frac{\psi'(t)}{\varphi'(t)}.$$

Если, кроме того, существуют y''_{tt} и x''_{tt}, то

$$y''_{xx} = \left(\frac{y'_t}{x'_t}\right)'_x = \left(\frac{y'_t}{x'_t}\right)'_t t'_x = \frac{y''_{tt} x'_t - y'_t x''_{tt}}{(x'_t)^3}.$$

При соответствующих предположениях аналогично можно определить производные любого порядка параметрически заданной функции. Например, если функция $y = f(x)$ задана параметрически соотношениями $x = a\cos^3 t$, $y = b\sin^3 t$ $(-\infty < t < \infty)$, где a и b — положительные постоянные, то $y'_t = 3b\sin^2 t \cos t$, $x'_t = -3a\cos^2 t \sin t$. При

$t \neq \pi k/2$ ($k=0, \pm 1, \pm 2, \ldots$) производная $x'_t \neq 0$. Следовательно, при этих значениях t получаем

$$y'_x = \frac{y'_t}{x'_t} = -\frac{3b \sin^2 t \cos t}{3a \cos^2 t \sin t} = -\frac{b}{a} \operatorname{tg} t.$$

Далее,

$$y''_{xx} = \left(-\frac{b}{a} \operatorname{tg} t\right)'_x = \left(-\frac{b}{a} \operatorname{tg} t\right)'_t t'_x = -\frac{b}{a \cos^2 t} \cdot \frac{1}{x'_t} =$$
$$= \frac{b}{3a^2 \cos^4 t \sin t}.$$

5.13. Производная неявно заданной функции

Если дифференцируемая функция $y=f(x)$ задана неявно уравнением $F(x, y)=0$, то, дифференцируя тождество $F[x, f(x)] \equiv 0$ по x (как сложную функцию), можно определить $f'(x)$. Дифференцируя выражение $f'(x)$ по x, можно определить $f''(x)$ и т. д.

Например, если функция $y=f(x)$ задана неявно уравнением $\operatorname{arctg} y - y + x = 0$, то, дифференцируя тождество $\operatorname{arctg} f(x) - f(x) + x \equiv 0$ по x, найдем

$$\frac{f'(x)}{1+y^2} - f'(x) + 1 = 0, \text{ откуда } y' = f'(x) = 1 + y^{-2}.$$

Дифференцируя по x последнее равенство, получаем

$$y'' = f''(x) = -2y^{-3}y' = -\frac{2(1+y^2)}{y^5}.$$

5.14. Теоремы о среднем для дифференцируемых функций

Теорема Ролля. *Если функция $y=f(x)$ непрерывна на отрезке $[a, b]$, дифференцируема на интервале $]a, b[$ и если $f(a)=f(b)$, то на интервале $]a, b[$ найдется хотя бы одна точка c такая, что $f'(c)=0$.*

Геометрически (рис. 5.4), в условиях теоремы, на графике функции $f(x)$ (дуга AB) найдется хотя бы одна точка, в которой касательная к графику параллельна оси Ox (это точки M и N). Условия теоремы существенны. Так, на рис. 5.5 функция $f(x)$ разрывна в точке a_1, а на рис. 5.6 функция $f(x)$ не имеет в точке a_2 производной (ни конечной, ни равной $+\infty$, ни равной $-\infty$). В обоих случаях на отрезке $[a, b]$ не существует ни одной точки, в которой $f'(x)=0$.

Теорема Лагранжа. *Если функция $y=f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема на интервале $]a, b[$,*

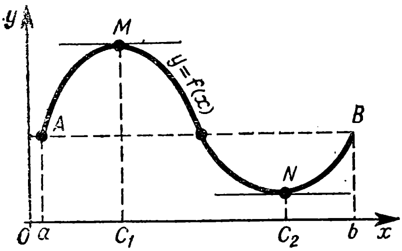


Рис. 5.4

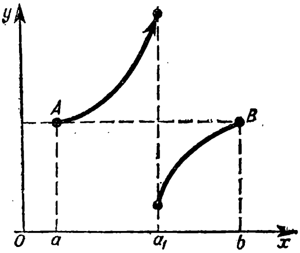


Рис. 5.5

то на интервале $]a, b[$ найдется хотя бы одна точка c такая, что $f(b)-f(a)=f'(c)\cdot(b-a)$.

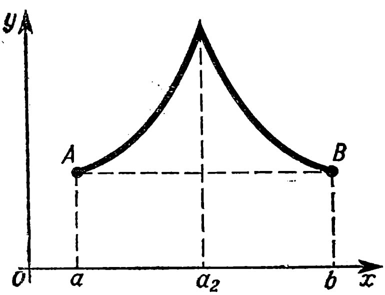


Рис. 5.6

Геометрически (рис. 5.7), в условиях теоремы, на графике функции $f(x)$ (дуга AB) найдется хотя бы одна точка, в которой касательная параллельна хорде AB (это точки M и N).

Замечание. Полагая $b=a+\Delta x$, где $\Delta x \neq 0$, можно точку c представить в виде $c=a+\theta(b-a)=a+\theta\cdot\Delta x$ $(0<\theta<1)$. Тогда формула $f(b)-f(a)=f'(c)(b-a)$ принимает вид

$$f(a+\Delta x)-f(a)=$$
$$=f'(a+\theta\Delta x)\Delta x.$$

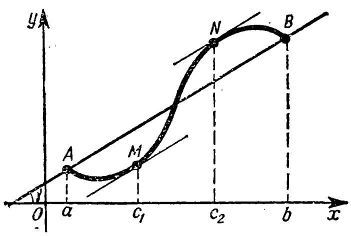


Рис. 5.7

Это *формула конечных приращений Лагранжа.*

Если функция $f(x)$ определена на некотором промежутке, непрерывна в каждом из концов этого промежутка (если он ему принадлежит) и имеет производную, равную нулю, во всех внутренних точках промежутка, то функция $f(x)$ постоянна в промежутке (признак постоянства функции).

Теорема Коши. *Пусть функции $\psi(t)$ и $\varphi(t)$ непре-*

рывны на отрезке $[\alpha, \beta]$ *и дифференцируемы на интервале* $]\alpha, \beta[$. *Если производная* $\varphi'(t) \neq 0$ *при всех* $t \in]\alpha, \beta[$, *то на интервале* $]\alpha, \beta[$ *найдется хотя бы одна точка* ξ *такая, что*

$$\frac{\psi(\beta)-\psi(\alpha)}{\varphi(\beta)-\varphi(\alpha)}=\frac{\psi'(\xi)}{\varphi'(\xi)}.$$

Геометрический смысл теоремы Коши тот же, что и у теоремы Лагранжа. Действительно, если функцию (рис. 5.7) задать параметрически соотношениями $x=\varphi(t)$, $y=\psi(t)$ $(\alpha \leqslant t \leqslant \beta)$ так, что координаты точки A соответствуют значению $t=\alpha$, а координаты точки B — значению $t=\beta$, то в точках M и N угловой коэффициент касательной $\frac{\psi'(\xi)}{\varphi'(\xi)}$ равен угловому коэффициенту $\operatorname{tg}\gamma$ хорды AB, где $\operatorname{tg}\gamma=\frac{\psi(\beta)-\psi(\alpha)}{\varphi(\beta)-\varphi(\alpha)}$.

5.15. Формула Тейлора

Если функция $y=f(x)$ имеет производные до $(n+1)$-го порядка включительно в некоторой окрестности точки $x=a$, то для всех x из этой окрестности справедливо равенство

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots$$
$$\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+R_n(x),$$

где

$$R_n(x)=\frac{f^{(n+1)}[a+\theta(x-a)]}{(n+1)!}\cdot(x-a)^{n+1} \qquad (0<\theta<1)$$

— остаточный член формулы Тейлора в форме Лагранжа.

Замечание. Полагая $x=a+\Delta x$, где $\Delta x \neq 0$, формулу Тейлора можно представить в виде

$$f(a+\Delta x)-f(a)=\frac{f'(a)}{1!}\cdot\Delta x+\frac{f''(a)}{2!}\cdot(\Delta x)^2+\ldots$$
$$\ldots+\frac{f^{(n)}(a)}{n!}(\Delta x)^n+\frac{f^{(n+1)}(a+\theta\Delta x)}{(n+1)!}\cdot(\Delta x)^{n+1}$$
$$(0<\theta<1),$$

обобщающем формулу конечных приращений Лагранжа.

При $a=0$ формула Тейлора принимает вид

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+$$
$$+\frac{f^{(n+1)}(\theta x)}{(n+1)!}x^{n+1} \qquad (0<\theta<1)$$

и называется *формулой Маклорена*.

Формулу Тейлора используют для представления функций многочленами, вычисления приближенных значений функций, при исследовании функций и вычислении пределов.

○ **Примеры.**

1. $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + R_n(x)$.

2. $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n \cdot \frac{x^{2n}}{(2n)!} + R_{2n+1}(x)$.

3. $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^{n-1} \cdot \frac{x^{2n-1}}{(2n-1)!} +$
$+ R_{2n}(x)$.

4. $(1+x)^m = 1 + mx + \frac{m(m-1)}{2!} x^2 + \ldots$
$\ldots + \frac{m(m-1)\cdot\ldots\cdot(m-n+1)}{n!} x^n + R_n(x)$.

5. $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^{n-1}\frac{x^n}{n} +$
$+ R_n(x)$. ●

5.16. Правило Лопиталя раскрытия неопределенностей

Если существует окрестность точки $x=a$, в которой функции $\psi(x)$ и $\varphi(x)$ дифференцируемы, за исключением, быть может, самой точки $x=a$, $\varphi'(x) \neq 0$ и либо $\lim\limits_{x\to a} \psi(x) = \lim\limits_{x\to a} \varphi(x) = 0$, либо $\lim\limits_{x\to a} \psi(x) = \lim\limits_{x\to a} \varphi(x) = \infty$, то при условии существования $\lim\limits_{x\to a} \frac{\psi'(x)}{\varphi'(x)}$ существует и $\lim\limits_{x\to a} \frac{\psi(x)}{\varphi(x)}$, причем имеет место равенство

$$\lim_{x\to a} \frac{\psi(x)}{\varphi(x)} = \lim_{x\to a} \frac{\psi'(x)}{\varphi'(x)}$$

(правило применимо и в случае, когда a бесконечно).

Замечание. Если отношение $\frac{\psi'(x)}{\varphi'(x)}$ в точке $x=a$ также представляет неопределенность вида $\frac{0}{0}$ или $\frac{\infty}{\infty}$, то при выполнении соответствующих условий правило Лопиталя может быть применено и к $\frac{\psi'(x)}{\varphi'(x)}$, так что

$$\lim_{x\to a} \frac{\psi(x)}{\varphi(x)} = \lim_{x\to a} \frac{\psi'(x)}{\varphi'(x)} = \lim_{x\to a} \frac{\psi''(x)}{\varphi''(x)},$$

причем процесс, если это необходимо, можно продолжить.

Неопределенности вида $(0\cdot\infty)$ или $(\infty-\infty)$ приводятся к неопределенностям $\frac{0}{0}$ или $\frac{\infty}{\infty}$ с помощью алгебраических преобразований.

Неопределенности вида (1^∞), (∞^0), (0^0) приводятся к неопределенностям $\frac{0}{0}$ или $\left(\frac{\infty}{\infty}\right)$ с помощью предварительного логарифмирования.

○ **Примеры.**

1. $\lim\limits_{x\to 2}\dfrac{\ln(x^2-3)}{x^2+7x-18}=\left(\dfrac{0}{0}\right)=\lim\limits_{x\to 2}\dfrac{2x/(x^2-3)}{2x+7}=\dfrac{4}{11}.$

2. $\lim\limits_{x\to 4}(x-4)\operatorname{tg}\dfrac{\pi x}{8}=(0\cdot\infty)=\lim\limits_{x\to 4}\dfrac{x-4}{\operatorname{ctg}\dfrac{\pi x}{8}}=$

$$=\left(\frac{0}{0}\right)=\lim_{x\to 4}\frac{1}{-\dfrac{\pi}{8}\cdot\dfrac{1}{\sin^2\dfrac{\pi x}{8}}}=-\frac{8}{\pi}.$$

3. Вычислить предел $\lim\limits_{x\to\pi/2}(\sin x)^{\operatorname{tg}x}=(1^\infty)$.

Сначала найдем предел логарифма данной функции. Получим

$$\lim_{x\to\pi/2}\ln(\sin x)^{\operatorname{tg}x}=\lim_{x\to\pi/2}[\operatorname{tg}x\ln(\sin x)]=(\infty\cdot 0)=$$

$$=\lim_{x\to\pi/2}\frac{\ln(\sin x)}{\operatorname{ctg}x}=\left(\frac{0}{0}\right)=\lim_{x\to\pi/2}\frac{\operatorname{ctg}x}{-\dfrac{1}{\sin^2 x}}=$$

$$=\lim_{x\to\pi/2}(-\sin x\cos x)=0.$$

Отсюда

$$\lim_{x\to\pi/2}(\sin x)^{\operatorname{tg}x}=e^0=1. \ \bullet$$

5.17. Признаки монотонности функции

Пусть функция $y=f(x)$ определена и дифференцируема в интервале $]a, b[$.

Для того чтобы функция $f(x)$ была неубывающей (невозрастающей) в интервале $]a, b[$, необходимо и достаточно, чтобы во всех точках интервала $]a, b[$ производная $f'(x)\geqslant 0$ $(f'(x)\leqslant 0)$.

Для того чтобы функция $f(x)$ была строго возрастающей (строго убывающей) в интервале $]a, b[$, достаточно, чтобы во всех точках интервала $]a, b[$ производная $f'(x)>0$ $(f'(x)<0)$.

○ **Пример.** Для функции $f(x)=x^2e^{-x}$ производная $f'(x)=xe^{-x}(2-x)$. Поэтому $f'(x)<0$, если $x\in$ $\in]-\infty, 0[\cup]2, +\infty[$, и в этих промежутках функция $f(x)$ строго убывает; $f'(x)>0$, если $x\in]0, 2[$, и в этом интервале функция строго возрастает. ●

5.18. Экстремум функции

Если существует такая δ-окрестность точки x_0, что для всех точек $x\neq x_0$ и принадлежащих этой окрестности выполняется неравенство $f(x)\geqslant f(x_0))$, $(f(x)\leqslant f(x_0))$, то точка x_0 называется *точкой нестрогого минимума* (*нестрогого максимума*) функции $f(x)$, а число $f(x_0)$ — *минимумом* (*максимумом*) этой функции.

Если в указанной δ-окрестности выполняется строгое неравенство $f(x)>f(x_0)$ $(f(x)<f(x_0)$, то точку x_0 называют *точкой строгого минимума* (*строгого максимума*).

Точки строгого и нестрогого максимума и минимума функции называют ее *точками экстремума.*

Если x_0 — точка минимума функции $f(x)$, то в указанной δ-окрестности точки x_0 приращение функции $\Delta f(x_0)=$ $=f(x)-f(x_0)\geqslant 0$ $(\Delta f(x_0)>0)$; если же x_0 — точка максимума функции $f(x)$, то $\Delta f(x_0)\leqslant 0$ $(\Delta f(x_0)<0)$ во всех точках δ — окрестности точки x_0.

Пусть точка x_0 является точкой экстремума функции $f(x)$, определенной в некоторой окрестности точки x_0. Тогда либо производная $f'(x_0)$ не существует, либо $f'(x_0)=0$ (*необходимый признак экстремума*).

Точки, в которых производная функции $f(x)$ не существует или обращается в нуль, называют *критическими.*

Например, для функции $f(x)=\sqrt[3]{x}\cdot(x-8)$ ее производная $f'(x)=\frac{4}{3}\cdot\frac{x-2}{\sqrt[3]{x^2}}$. Следовательно, $f'(x)=0$ при $x=2$; $f'(x)$ не существует при $x=0$ $(f'(0)=-\infty)$. Таким образом, точки экстремума функции $f(x)$, если таковые вообще имеются, находятся только среди критических точек $x=0$ и $x=2$.

Достаточные условия строгого экстремума непрерывной функции

1. Пусть функция $f(x)$ дифференцируема в некоторой окрестности $]x_0-\delta, x_0+\delta[$ критической точки x_0, за исключением, быть может, самой точки x_0, в которой тем не менее функция $f(x)$ непрерывна. Если при этом в интервалах $]x_0-\delta, x_0[$ и $]x_0, x_0+\delta[$ производная $f'(x)$ имеет противоположные знаки, то x_0 — точка экстремума, причем:

а) если $f'(x) > 0$ при $x \in]x_0 - \delta, x_0[$ и $f'(x) < 0$ при $x \in]x_0, x_0 + \delta[$, то x_0 — точка строгого максимума функции;

б) если $f'(x) < 0$ при $x \in]x_0 - \delta, x_0[$ и $f'(x) > 0$ при $x \in]x_0, x_0 + \delta[$, то x_0 — точка строгого минимума функции.

Если же $f'(x)$ сохраняет знак при всех $x \neq x_0$, $x \in]x_0 - \delta, x_0 + \delta[$, то точка x_0 не является точкой экстремума функции $f(x)$.

2. Пусть $f'(x_0) = 0$, функция $f(x)$ дважды дифференцируема в некоторой окрестности точки x_0 и $f''(x)$ непрерывна в этой окрестности. Тогда:

а) если $f''(x_0) < 0$, то x_0 — точка строгого максимума функции $f(x)$;

б) если $f''(x_0) > 0$, то x_0 — точка строгого минимума функции $f(x)$;

в) если $f''(x_0) = 0$, то вопрос о наличии экстремума остается открытым.

3. Пусть $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0$, а $f^{(n)}(x_0) \neq 0$. Тогда:

а) если n — четное, то при $f^{(n)}(x_0) < 0$ точка x_0 является точкой строгого максимума, а при $f^{(n)}(x_0) > 0$ — точкой строгого минимума;

б) если n — нечетное, то x_0 не является точкой экстремума.

○ **Примеры.**

1. Дана функция $f(x) = \sqrt[3]{x}(x-8)$. Ее производная $f'(x) = \frac{4}{3} \frac{x-2}{\sqrt[3]{x^2}}$, так что критическими являются точки $x = 0$, $x = 2$. При этом $f'(x) < 0$ как при $x < 0$, так и при $0 < x < 2$ и, следовательно, согласно условию 1, точка $x = 0$ не является точкой экстремума функции $f(x)$. С другой стороны, $f'(x) < 0$ при $0 < x < 2$ и $f'(x) > 0$ при $x > 2$; следовательно, $x = 2$ является точкой строгого минимума, а число $f(2) = -6\sqrt[3]{2}$ — минимумом функции $f(x)$.

2. Рассмотрим функцию $f(x) = x^3 - 3x^2$. Ее производная $f'(x) = 3x^2 - 6x = 3x(x-2)$, так что $f'(x) = 0$ при $x = 0$ и $x = 2$ (точки, в которых $f'(x) = 0$, называют *стационарными*). Вторая производная $f''(x) = 6x - 6$ и, следовательно, $f''(0) = -6 < 0$, а $f''(2) = 12 - 6 = 6 > 0$. Тогда, согласно условию 2, точка $x = 0$ является точкой строгого максимума функции и $f(0) = 0$, а точка $x = 2$ — точкой строгого минимума и $f(2) = -4$.

3. Пусть дана функция $f(x) = x^4$. Для этой функции $f'(x) = 4x^3$, $f''(x) = 12x^2$, $f'''(x) = 24x$, $f^{IV}(x) = 24$, так что

$f'(0)=f''(0)=f'''(0)=0$, а $f^{IV}(0)=24>0$. Тогда, согласно условию 3, точка $x=0$ является точкой строгого минимума и $f(0)=0$. ●

5.19. Наибольшее и наименьшее значения функции на множестве

Пусть функция $y=f(x)$ определена на некотором множестве $V \subseteq \mathbf{R}^1$ и точка $x_0 \in V$.

Если для всех $x \in V$ выполняется неравенство $f(x) \leqslant f(x_0)$, то говорят, что в точке x_0 функция $f(x)$ *принимает свое наибольшее значение* $f(x_0)$ *на множестве* V.

Если для всех $x \in V$ справедливо неравенство $f(x) \geqslant f(x_0)$, то в точке x_0 функция $f(x)$ *принимает свое наименьшее значение* $f(x_0)$ на множестве V.

Если функция $f(x)$ определена и непрерывна на отрезке $[a, b]$, то она достигает на этом отрезке своих наибольшего и наименьшего значений, для определения которых необходимо найти все критические точки функции, принадлежащие отрезку $[a, b]$, добавить к ним концы отрезка (точки $x=a$ и $x=b$), найти значения функции во всех выделенных точках и из полученных значений выбрать самое большое и самое маленькое.

Если же функция $f(x)$ имеет на отрезке $[a, b]$ конечное число точек разрыва, то для решения задачи отыскания наибольшего и наименьшего значений такой функции к указанным точкам необходимо добавить все точки разрыва функции, принадлежащие отрезку $[a, b]$, и исследовать поведение функции в окрестности каждой точки разрыва.

Наконец, если функция $f(x)$ задана на открытом промежутке (например, на интервале $]a, b[$, $a<b$), то помимо исследования ранее указанных точек необходимо исследовать также поведение функции в односторонних окрестностях концов промежутка (при $x \to a+0$ и при $x \to b-0$).

○ **Примеры.**

1. Функция $f(x)=x^3-3x^2$ определена и непрерывна на отрезке $[1, 3]$. Ее производная $f'(x)=3x(x-2)$, так что точки $x=0$, $x=2$ — стационарные. При этом, однако, $x=0 \notin [1, 3]$. Следовательно, необходимо рассмотреть лишь точки $x=1$, $x=2$, $x=3$. Имеем: $f(1)=-2$, $f(2)=-4$, $f(3)=0$. Таким образом, наибольшее значение, равное нулю, функция принимает в точке $x=3$, а наименьшее значение, равное (-4), — в точке $x=2$.

2. Рассмотрим функцию $f(x)=\begin{cases} x^2 & \text{при } -2 \leqslant x < 0, \\ x-1 & \text{при } \quad 0 \leqslant x \leqslant 2. \end{cases}$

Функция определена на отрезке $[-2, 2]$, однако разрывна при $x=0$ $\left(\lim\limits_{x \to -0} x^2 = 0\right.$, в то время как $\left. f(0) = -1\right)$.

Стационарных точек функция $f(x)$ не имеет, так как $f'(x) = 2x$ при $-2 \leqslant x < 0$ и $f'(x) = 1$ при $0 < x \leqslant 2$, т. е. $f'(x) \neq 0$ на отрезке $[-2, 2]$.

Следовательно, необходимо рассмотреть лишь точки $x = -2$, $x = 0$, $x = 2$. Имеем: $f(-2) = 4$, $f(0) = -1$, $\lim\limits_{x \to -0} f(x) = 0$, $f(2) = 1$. Таким образом, наибольшее значение, равное 4, функция принимает при $x = -2$, а наименьшее значение, равное (-1),— в точке $x = 0$.

3. Функция $f(x) = 4 - x^2$ определена и непрерывна в полуинтервале $[-2, 1[$. Ее производная $f'(x) = -2x$, так что точка $x = 0$—стационарная точка, принадлежащая $[-2, 1[$. Необходимо рассмотреть точки $x = -2, x = 0, x = 1$. Имеем: $f(-2) = 0$, $f(0) = 4$, $\lim\limits_{x \to 1-0} f(x) = \lim\limits_{x \to 1-0} (4 - x^2) = 3$. Таким образом, наибольшее значение, равное 4, функция принимает при $x = 0$, а наименьшее значение, равное 0,—в точке $x = -2$. ●

5.20. Направление выпуклости графика функции

График дифференцируемой функции $y = f(x)$ направлен выпуклостью вверх (выпуклостью вниз) на интервале $]a, b[$, если в пределах этого интервала он расположен ниже (выше) касательной, проведенной в любой его точке (рис. 5.8).

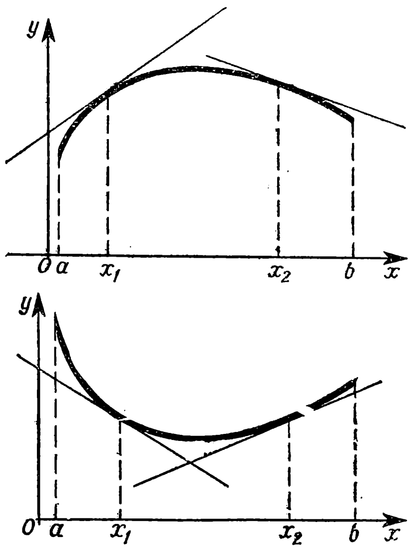


Рис. 5.8

Достаточное условие выпуклости

Пусть функция $y = f(x)$ определена и дважды дифференцируема на интервале $]a, b[$. Тогда:

а) если $f''(x) < 0$ во всех точках интервала $]a, b[$, то график функции $f(x)$ направлен выпуклостью вверх на этом интервале;

б) если $f''(x) > 0$ во всех точках интервала $]a, b[$, то график функции $f(x)$ направлен выпуклостью вниз на этом интервале.

○ **Пример.** Для функции $f(x) = \sqrt[3]{x}\cdot(x-8)$ вторая производная $f''(x) = \frac{4}{9}\cdot\frac{x+4}{\sqrt[3]{x^5}}$. Поэтому $f''(x) < 0$ при $-4 < x < 0$ и, следовательно, на этом интервале график функции направлен выпуклостью вверх; $f''(x) > 0$ на интервалах $]-\infty, -4[$ и $]0, +\infty[$; следовательно, на этих промежутках график функции направлен выпуклостью вниз. ●

5.21. Точки перегиба графика функции

Точка $(x_0; f(x_0))$ называется *точкой перегиба* графика функции $y = f(x)$, если в этой точке существует касательная к графику и в промежутках $]x_0 - \delta, x_0[$ и $]x_0, x_0 + \delta[$, где δ — некоторое положительное число, график функции имеет разное направление выпуклости.

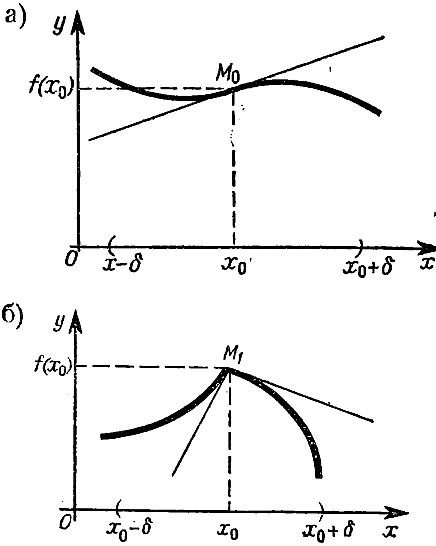


Рис. 5.9

Так, на рис. 5.9, а точка M_0 является точкой перегиба графика функции, а на рис. 5.9, б точка M_1 не является точкой перегиба, хотя в этой точке и происходит изменение направления выпуклости графика (в точке M_1 не существует касательной к графику).

Пусть точка $(x_0, f(x_0))$ является точкой перегиба графика функции $y = f(x)$, определенной в некоторой окрестности точки x_0. Тогда либо вторая производная $f''(x_0)$ не существует, либо $f''(x_0) = 0$ (*необходимый признак точки перегиба*).

Пусть в точке $(x_0; f(x_0))$ существует касательная (хотя бы вертикальная) к графику функции $y = f(x)$ и в некоторой окрестности $]x_0-\delta,\ x_0+\delta[$ $(\delta>0)$ существует $f''(x)$, за исключением, быть может, точки x_0, причем $f''(x_0)=0$ или $f''(x_0)$ не существует. Если при этом в интервалах $]x_0-\delta,\ x_0[$ и $]x_0,\ x_0+\delta[$ производная $f''(x)$ имеет противоположные знаки, то точка $(x_0;\ f(x_0))$ является точкой перегиба графика функции $y=f(x)$ (*достаточные условия точки перегиба*).

○ **Примеры.**

1. Рассмотрим функцию $f(x)=x^3$. Ее производная $f'(x)=3x^2$, а $f''(x)=6x$. Тогда $f''(x)=0$ при $x=0$, причем $f'(0)=0$ (в точке $x=0$ существует касательная к графику). Производная $f''(x)<0$ при $x<0$ и $f''(x)>0$ при $x>0$. Следовательно, точка $(0;\ 0)$ является точкой перегиба графика функции $y=x^3$.

2. Для функции $f(x)=\sqrt[3]{x}(x-8)$ имеем $f'(x)=\frac{4}{3}\frac{x-2}{\sqrt[3]{x^2}}$, $f''(x)=\frac{4}{9}\frac{x+4}{\sqrt[3]{x^5}}$, так что $f''(x)=0$ при $x=-4$ и $f''(x)$ не существует при $x=0$. При этом $f'(-4)=-\frac{4}{\sqrt[3]{2}}$, а $f'(0)=-\infty$ (в точках $x=-4$ и $x=0$ существуют касательные к графику функции, причем при $x=0$ касательная вертикальна). Кроме того, $f''(x)>0$ при $x<-4$, $f''(x)<0$ при $-4<x<0$ и $f''(x)>0$ при $x>0$. Следовательно, точки $(-4;\ 12\sqrt[3]{4})$ и $(0;\ 0)$ являются точками перегиба графика функции. ●

5.22. Общая схема исследования функции

1. Находят область определения, точки разрыва, множество значений функции.
2. Находят асимптоты графика.
3. Исследуют функцию на четность, нечетность, периодичность.
4. Исследуют функцию на монотонность и находят ее экстремумы.
5. Определяют направление выпуклости графика, точки перегиба.
6. Находят точки пересечения с осями координат.
7. Строят график функции.

○ **Пример.** Исследуем функцию

$$y=f(x)=(x+5)\sqrt[3]{x^2}.$$

1. $D(y)=]-\infty,\ +\infty[$. Функция $f(x)$ непрерывна на $D(y)$. Точек разрыва нет.

2. Вертикальных асимптот нет; $k=\lim\limits_{x\to\infty}\frac{x+5}{x}\sqrt[3]{x^2}=\infty$, наклонных асимптот нет.

3. Функция не является ни четной, ни нечетной, ни периодической.

4. $y'=\frac{5}{3}\frac{x+2}{\sqrt[3]{x}}$. Критические точки $x=-2$, $x=0$.

x	$]-\infty, -2[$	-2	$]-2, 0[$	0	$]0, +\infty[$
Знак y'	$+$	$y'=0$	$-$	$y'=\infty$	$+$
Поведение функции	Возрастает	max $3\sqrt[3]{4}$	Убывает	min 0	Возрастает

5. $y''=\frac{10}{9}\cdot\frac{x-1}{\sqrt[3]{x^4}}$; $y''=0$ при $x=1$, y'' не существует при $x=0$.

x	$]-\infty, 0[$	0	$]0, 1[$	1	$]1, +\infty[$
Знак y''	$-$	$y''=-\infty$	$-$	$y''=0$	$+$
Поведение функции	Выпукла вверх	Не является точкой перегиба	Выпукла вверх	Точка перегиба $y=6$	Выпукла вниз

6. $f(x)=0$ при $x=0$ и $x=-5$.

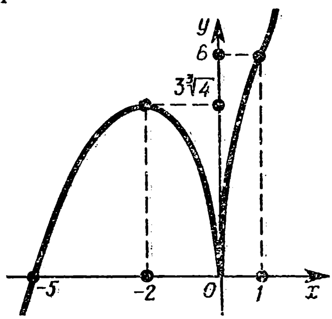


Рис. 5.10

7. График функции изображен на рис. 5.10. ●

РАЗДЕЛ VI. ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

6.1. Частные производные функций нескольких переменных

Пусть функция $f(M)$ определена в некоторой окрестности точки $M_0(x_1^0; \ldots; x_i^0; \ldots; x_n^0)$. Рассмотрим точку $M_i(x_1^0, \ldots, x_i^0+\Delta x_i, \ldots, x_n^0)$. Если существует $\lim\limits_{\Delta x_i \to 0} \frac{f(M_i)-f(M_0)}{\Delta x_i}$, то он называется *частной производной* $\left(\text{обозначение: } \frac{\partial f}{\partial x_i}(M_0)\right)$ функции $f(M)$ в точке M_0, т. е.

$$\frac{\partial f}{\partial x_i}(M_0) = \lim_{\Delta x_i \to 0} \frac{f(M_i)-f(M_0)}{\Delta x_i} =$$

$$= \lim_{\Delta x_i \to 0} \frac{f(x_1^0, \ldots, x_i^0+\Delta x_i, \ldots, x_n^0) - f(x_1^0, \ldots, x_i^0, \ldots, x_n^0)}{\Delta x_i}.$$

Из определения частной производной следует, что для ее нахождения достаточно вычислить обычную производную по x_i, считая $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ постоянными. Например, если $f(M)=x_1^5x_2x_3-x_1x_2^3x_3+2x_2x_3^2$, то

$$\frac{\partial f}{\partial x_1} = (x_1^5)'\,x_2x_3 - x_1'x_2^3x_3 = 5x_1^4x_2x_3 - x_2^3x_3,$$

$$\frac{\partial f}{\partial x_2} = x_1^5x_2'x_3 - x_1(x_2^3)'\,x_3 + 2x_2'x_3^2 = -3x_2^2x_1x_3 + x_1^5x_3 + 2x_3^2,$$

$$\frac{\partial f}{\partial x_3} = x_1^5x_2x_3' - x_1x_2x_3' + 2x_2(x_3^2)' = x_1^5x_2 - x_1x_2^3 + 4x_2x_3.$$

Для того чтобы вычислить частную производную в некоторой фиксированной точке, достаточно найти эту частную производную в любой точке и в найденное выражение подставить вместо неизвестных координаты данной точки.

Найдем, например, частную производную $\frac{\partial f}{\partial y}(M_0)$, где $f(M)=e^{x^2+y^2}$, $M_0(1;\ -1)$.

Так как $\frac{\partial f}{\partial y} = 2ye^{x^2+y^2}$, то

$$\frac{\partial f}{\partial y}(M_0) = 2(-1)e^{1^2+(-1)^2} = -2e^2.$$

6.2. Полное приращение функции нескольких переменных

Пусть функция $f(M)$ определена в некоторой окрестности точки $M_0(x_1^0; \ldots; x_i^0; \ldots; x_n^0)$. Рассмотрим точку $M_\Delta(x_1^0+\Delta x_1; \ldots; x_i^0+\Delta x_i; \ldots; x_n^0+\Delta x_n)$. *Полным приращением* Δf функции $f(M)$ в точке M_0 называется число $f(M_\Delta)-f(M_0)$, т. е.

$$\Delta f = f(M_\Delta)-f(M_0) = \\ = f(x_1^0+\Delta x_1, \ldots, x_i^0+\Delta x_i, \ldots, x_n^0+\Delta x_n) - \\ - f(x_1^0, \ldots, x_i^0, \ldots, x_n^0).$$

○ Найдем, например, полное приращение функции $f(M) = x_1^2 + x_2^2$ в точке $M_0(1, -2)$.

Так как $M_\Delta(1+\Delta x_1; -2+\Delta x_2)$, то

$$\Delta f = f(M_\Delta)-f(M_0) = (1+\Delta x_1)^2+(-2+\Delta x_2)^2-1^2-(-2)^2 = \\ = 2\Delta x_1 - 4\Delta x_2 + (\Delta x_1)^2 + (\Delta x_2)^2. \ ●$$

Полное приращение функции нескольких переменных существует в любой точке, в окрестности которой эта функция определена.

○ Найдем, например, полное приращение функции $f(M) = x_1^3 + x_1x_2 + x_1x_3$ в произвольной точке $M(x_1; x_2; x_3) \in \in \mathbf{R}^3$.

Так как $M_\Delta(x_1+\Delta x_1; x_2+\Delta x_2; x_3+\Delta x_3)$, то

$$\Delta f = f(M_\Delta)-f(M_0) = (x_1+\Delta x_1)^3 + (x_1+\Delta x_1)(x_2+\Delta x_2) + \\ + (x_1+\Delta x_1)(x_3+\Delta x_3) - x_1^3 - x_1x_2 - x_1x_3 = (3x_1^2+x_2+x_3)\Delta x_1 + \\ + x_1\Delta x_2 + x_1\Delta x_3 + 3x_1(\Delta x_1)^2 + (\Delta x_1)^3 + \Delta x_1\cdot\Delta x_2 + \Delta x_1\cdot\Delta x_3 \ ●$$

6.3. Дифференцируемость функций нескольких переменных

Пусть функция $f(M)$ определена в некоторой окрестности точки M_0, $M_0 \in \mathbf{R}^n$. Функция $f(M)$ называется *дифференцируемой в точке* M_0, если полное приращение Δf в этой точке имеет следующий вид:

$$\Delta f = A_1\Delta x_1 + A_2\Delta x_2 + \ldots + A_n\Delta x_n + \\ + \alpha_1\Delta x_1 + \alpha_2\Delta x_2 + \ldots + \alpha_n\Delta x_n,$$

где $A_1, A_2, \ldots, A_n$ — некоторые числа, не зависящие от $\Delta x_1, \Delta x_2, \ldots, \Delta x_n$, а $\alpha_1, \alpha_2, \ldots, \alpha_n \to 0$ при $\Delta x_1 \to 0$, $\Delta x_2 \to 0, \ldots, \Delta x_n \to 0$.

○ **Примеры.**

1. Функция $f(M) = x_1^2 + x_2^2$ дифференцируема в точке $M_0(1; -2)$. В самом деле,

$$\Delta f = 2\Delta x_1 - 4\Delta x_2 + (\Delta x_1)^2 + (\Delta x_2)^2,$$

т. е.

$$\Delta f = A_1\Delta x_1 + A_2\Delta x_2 + \alpha_1\Delta x_1 + \alpha_2\Delta x_2,$$

где $A_1 = 2$, $A_2 = -4$, $\alpha_1 = \Delta x_1$, $\alpha_2 = \Delta x_2$.

2. Функция $f(M) = x_1^3 + x_1x_2 + x_1x_3$ дифференцируема в любой точке $M(x_1; x_2; x_3) \in \mathbf{R}^3$.

Действительно,

$\Delta f = (3x_1^2 + x_2 + x_3)\Delta x_1 + x_1\Delta x_2 + x_1\Delta x_3 + \Delta x_1 (3x_1\Delta x_1 + \Delta x_2 + \Delta x_3 + (\Delta x_1)^2)$, т. е. $\Delta f = A_1\Delta x_1 + A_2\Delta x_2 + A_3\Delta x_3 + \alpha_1\Delta x_1$, где $A_1 = 3x_1^2 + x_2 + x_3$, $A_2 = x_1$, $A_3 = x_1$, $\alpha_1 = 3x_1\Delta x_1 + \Delta x_2 + \Delta x_3 + (\Delta x_1)^2$ и $\alpha_1 \to 0$ при $\Delta x_1 \to 0$, $\Delta x_2 \to 0$, $\Delta x_3 \to 0$. ●

Линейная функция $f(M) = a_1x_1 + a_2x_2 + \ldots + a_nx_n$ и квадратичная функция $f(M) = a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots + 2a_{1n}x_1x_n + 2a_{23}x_2x_3 + \ldots + 2a_{2n}x_2x_n + \ldots + 2a_{n-1\,n}x_{n-1}x_n$ дифференцируемы в любой точке $M(x_1; x_2; \ldots; x_n) \in R^n$.

Свойства дифференцируемых функций

1°. Если функция дифференцируема в некоторой точке, то она в этой точке непрерывна.

2°. Если функция $f(M)$ дифференцируема в точке M_0, то она имеет в этой точке все частные производные, причем

$$\Delta f = \frac{\partial f}{\partial x_1}(M_0)\Delta x_1 + \frac{\partial f}{\partial x_2}(M_0)\Delta x_2 + \ldots + \frac{\partial f}{\partial x_n}(M_0)\Delta x_n + \alpha_1\Delta x_1 + \alpha_2\Delta x_2 + \ldots + \alpha_n\Delta x_n,$$

где $\alpha_1, \alpha_2, \ldots, \alpha_n \to 0$ при $\Delta x_1, \Delta x_2, \ldots, \Delta x_n \to 0$.

3°. Если функция $f(M)$ имеет все частные производные в некоторой окрестности точки M_0, которые непрерывны в самой точке M_0, то функция $f(M)$ дифференцируема в этой точке.

6.4. Дифференциал функции нескольких переменных

Если функция $f(M)$ дифференцируема в точке M_0, то линейная часть приращения функции $f(M)$ в точке M_0 называется ее дифференциалом $\mathrm{d}f(M_0)$ в точке M_0, т. е.

$$df(M_0)=\frac{\partial f}{\partial x_1}(M_0)\Delta x_1+\frac{\partial f}{\partial x_2}(M_0)\Delta x_2+\ldots+\frac{\partial f}{\partial x_n}(M_0)\Delta x_n.$$

Можно считать, что

$$dx_1=\Delta x_1,\ dx_2=\Delta x_2,\ \ldots,\ dx_n=\Delta x_n.$$

Тогда $df(M_0)=\frac{\partial f}{\partial x_1}(M_0)dx_1+\frac{\partial f}{\partial x_2}(M_0)dx_2+\ldots+\frac{\partial f}{\partial x_n}(M_0)dx_n$.

○ **Пример.** Найти дифференциал функции $f(M)=x_1^3x_2+x_2^2x_3+x_3$ в точке $M_0(2;\ 1;\ -3)$.

Так как $\frac{\partial f}{\partial x_1}=3x_1^2x_2$, $\frac{\partial f}{\partial x_2}=x_1^3+2x_2x_3$, $\frac{\partial f}{\partial x_3}=x_2^2+1$, то $\frac{\partial f}{\partial x_1}(M_0)=12$, $\frac{\partial f}{\partial x_2}(M_0)=2$, $\frac{\partial f}{\partial x_3}(M_0)=2$ и $df(M_0)=12dx_1+2dx_2+2dx_3$. ●

Основное свойство дифференциала

Если функция $f(M)$ дифференцируема в точке M_0, то при малых $\Delta x_1,\ \Delta x_2,\ \ldots,\ \Delta x_n$

$$\Delta f(M_0)\approx df(M_0),$$

т. е. $\Delta f(M_0)\approx\frac{\partial f}{\partial x_1}(M_0)\Delta x_1+\frac{\partial f}{\partial x_2}(M_0)\Delta x_2+\ldots+\frac{\partial f}{\partial x_n}(M_0)\Delta x_n$.

6.5. Градиент функции нескольких переменных

Градиентом функции $f(M)$ в точке M_0, $M_0\in\mathbb{R}^n$, называется вектор, координаты которого соответственно равны значениям частных производных функции $f(M)$ в точке M_0, т. е.

$$\operatorname{grad} f|_{M_0}=\left\{\frac{\partial f}{\partial x_1}(M_0),\ \frac{\partial f}{\partial x_2}(M_0),\ \ldots,\ \frac{\partial f}{\partial x_n}(M_0)\right\}.$$

Например, если $f(M)=x_1^3+x_1x_2^2+x_3^3$, то $\operatorname{grad} f=\{3x_1^2+x_2^2,\ 2x_1x_2,\ 3x_3^2\}$ ($\operatorname{grad} f$ — градиент функции $f(M)$ в произвольной точке $M(x_1;\ x_2;\ x_3)\in\mathbb{R}^3$). Если же $f(M)=x_1^5x_2-x_2^6$, $M_0(1;\ -1)$, то $\operatorname{grad} f|_{M_0}=(-5;\ 7)$.

Основное свойство градиента.

Пусть функция $f(M)$ дифференцируема в точке $M_0(x_1^0;\ x_2^0;\ \ldots;\ x_n^0)$, а $\alpha=(a_1,\ a_2,\ \ldots,\ a_n)$ — некоторый n-мерный вектор. Рассмотрим точку $M_t(x_1^0+a_1t,\ x_2^0+a_2t,\ \ldots,\ x_n^0+a_nt)$. Тогда:

1) если скалярное произведение $\operatorname{grad} f|_{M_0}\cdot\alpha<0$, то существует число $T_1>0$ такое, что $f(M_t)<f(M_0)$ для всех t, $0<t<T_1$;

2) если скалярное произведение $\operatorname{grad} f|_{M_0} \cdot \alpha > 0$, то существует число $T_2 > 0$ такое, что $f(M_t) > f(M_0)$ при всех t, $0 < t < T_2$.

Чтобы найти точку, в которой данная функция принимает значение большее, чем в точке $M_0(x_1^0; \ldots; x_n^0)$, можно поступить следующим образом:

1) выбрать направление перемещения, т. е. найти вектор $\alpha = (a_1, a_2, \ldots, a_n)$ такой, что $\operatorname{grad} f|_{M_0} \cdot \alpha > 0$ (если нет дополнительных ограничений, можно положить $\alpha = \operatorname{grad} f|_{M_0}$);

2) рассмотреть точку $M_t(x_1^0 + a_1 t; x_2^0 + a_2 t; \ldots; x_n^0 + a_n t)$ и подобрать параметр $t > 0$ так, чтобы $f(M_t) > f(M_0)$.

○ Найдем точку, в которой значение функции $f(M) = -3x_1^2 - 3x_2^2 + 2x_1x_2 + 10x_1 - 6x_2 + 2$ больше ее значения в точке $M_0(-1; 1)$.

Так как $\operatorname{grad} f = \{-6x_1 + 2x_2 + 10, \ -6x_2 + 2x_1 - 6\}$, то $\operatorname{grad} f|_{M_0} = (18, -14)$. Если $\alpha = (1, -1)$, то $\operatorname{grad} f|_{M_0} \cdot \alpha = 18 + 14 = 32 > 0$.

Рассмотрим точку $M_t(-1 + t; 1 - t)$. Тогда $f(M_t) = -8t^2 + 32t - 22$ и при $t = 2$ $\dfrac{df(M_t)}{dt} = 0$. Значит, при $t = 2$ функция $f(M_t)$ имеет наибольшее значение. Если $t = 2$, то $M_t(1, -1)$ и $f(M_t) = 10$, в то время как $f(M_0) = -22$.

Рассмотрим более сложный пример: найти точку на плоскости $x_1 + 3x_2 + x_3 = 15$, в которой значение функции $f(M) = -x_1^2 - 2x_2^2 - x_3^2$ больше ее значения в точке $M_0(1; 2; 8)$.

Рассмотрим точку $M_t(1 + a_1 t; \ 2 + a_2 t; \ 8 + a_3 t)$. Эта точка должна принадлежать данной плоскости, т. е. $1 + a_1 t + 3(2 + a_2 t) + 8 + a_3 t = 15$ или $a_1 + 3a_2 + a_3 = 0$. Кроме того, вектор $\alpha = (a_1, a_2, a_3)$ должен удовлетворять условию $\operatorname{grad} f|_{M_0} \cdot \alpha > 0$. Так как $\operatorname{grad} f|_{M_0} = (-2; -8; -16)$, то имеем систему

$$\begin{cases} a_1 + 3a_2 + a_3 = 0, \\ -2a_1 - 8a_2 - 16a_3 > 0. \end{cases} \qquad (*)$$

Вектор $\alpha = (-2, 1, -1)$ является решением системы (*). Таким образом, $M_t(1 - 2t; 2 + t; 8 - t)$, а $f(M_t) = -7t^2 + 12t - 73$. Функция $f(M_t)$ имеет наибольшее значение при $t = 6/7$. Если $t = 6/7$, то $M_t(-5/7; 20/7; 50/7)$, а $f(M_t) = -67\frac{6}{7}$, в то время как $f(M_0) = -73$. ●

Основное свойство градиента используют для отыскания экстремумов функций нескольких переменных (п. 9.23; 9.24).

6.6. Частные производные высших порядков

Предположим, что функция $f(M)$ имеет частную производную $\frac{\partial f}{\partial x_j}$ в каждой точке некоторой окрестности точки M_0. Если при этом существует частная производная по x_i от функции $\frac{\partial f}{\partial x_j}$ в самой точке M_0, то она называется *частной производной* $\frac{\partial^2 f}{\partial x_i \partial x_j}$ по x_i и x_j в точке M_0, т. е.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right).$$

Частная производная, взятая по различным переменным, называется *смешанной частной производной*. Кроме того, по определению,

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial^2 f}{\partial x_i \partial x_i}.$$

○ Найдем, например, частные производные второго порядка функции $f(M) = x_1^3 x_2^2 - x_1^2 x_2^3 + 2x_1 x_2$ в произвольной точке $M(x_1;\ x_2)$.

Так как

$$\frac{\partial f}{\partial x_1} = 3x_1^2 x_2^2 - 2x_1 x_2^3 + 2x_2; \quad \frac{\partial f}{\partial x_2} = 2x_1^3 x_2 - 3x_1^2 x_2^2 + 2x_1,$$

то

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_1}\right) = 6x_1 x_2^2 - 2x_2^3;$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right) = 6x_1^2 x_2 - 6x_1 x_2^2 + 2;$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right) = 6x_1^2 x_2 - 6x_1 x_2^2 + 2;$$

$$\frac{\partial^2 f}{\partial x_2^2} = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_2}\right) = 2x_1^3 - 6x_1^2 x_2. \quad ●$$

Аналогично определяются частные производные более высоких порядков.

6.7. Экстремумы функций нескольких переменных

Точка $M_0 \in \mathbf{R}^n$ называется точкой *локального минимума (максимума)* функции $f(M)$, если существует окрестность $S_r(M_0)$ точки M_0 такая, что при всех $M \in S_r(M_0)$ выполняется неравенство $f(M_0) \leqslant f(M)$ $(f(M_0) \geqslant f(M))$.

Точки локального минимума и локального максимума функции $f(M)$ называются *точками экстремума* этой функции.

По определению, точки экстремума функции всегда являются внутренними точками области определения этой функции.

Точка $M_0 \in \mathbf{R}^n$ называется *стационарной точкой функции* $f(M)$, если в этой точке градиент функции $f(M)$ является нулевым вектором, т. е.

$$\operatorname{grad} f|_{M_0} = \theta.$$

○ Найдем, например, стационарные точки функции

$$f(M) = x^2 - xy + y^2 + 9x - 6y + 20.$$

Так как $\operatorname{grad} f = \{2x - y + 9, \ -x + 2y - 6\}$, то для отыскания стационарных точек имеем систему уравнений

$$\begin{cases} 2x - \ y + 9 = 0, \\ -x + 2y - 6 = 0, \end{cases}$$

откуда получим единственную стационарную точку $M(-4;\ 1)$. ●

Необходимое условие экстремума

Если в точке экстремума функции $f(M)$ существуют все частные производные этой функции, то эта точка экстремума является стационарной точкой функции $f(M)$, т. е. $\frac{\partial f(M)}{\partial x_i} = 0$, $i = 1, \ \ldots, \ n$.

Заметим, что стационарная точка функции $f(M)$ может не быть точкой экстремума этой функции. Однако все точки экстремума функции находятся среди стационарных точек этой функции или точек, где не существуют ее частные производные.

6.8. Наименьшее и наибольшее значения функции нескольких переменных

Пусть функция $f(M)$ определена и непрерывна на некотором ограниченном замкнутом множестве V и, за исключением, быть может, отдельных точек, имеет на этом множестве частные производные. В этом случае найдется точка $M_0 \in V$, в которой функция имеет наименьшее (наибольшее) значение на множестве V. Если точка M_0 лежит внутри множества V, то она является точкой экстремума функции и ее следует искать среди стационарных точек или точек, где не существуют частные производные. Однако своего наименьшего (наибольшего) значения функция $f(M)$ может достигать и на границе множества V.

Таким образом, для отыскания наименьшего (наибольшего) значения функции $f(M)$ на множестве V необходимо:

1) определить все точки множества V, где не существуют частные производные функции $f(M)$;

2) найти все стационарные точки функции $f(M)$;

3) во всех найденных точках вычислить значения функции $f(M)$ и сравнить их со значениями функции на границе множества V.

Наименьшее (наибольшее) из этих значений и будет наименьшим (наибольшим) значением $f(M)$ на всем множестве V.

○ Найдем, например, наименьшее и наибольшее значения функции $f(M)=8x^2+y^2-(x^2+y^2)^2$ внутри круга $x^2+y^2\leqslant 3$.

Функция $f(M)$ имеет частные производные во всех точках. Так как $\operatorname{grad} f=\{16x-4x(x^2+y^2), 2y-4y(x^2+y^2)\}$, то для отыскания стационарных точек имеем систему уравнений

$$\begin{cases} 16x-4x(x^2+y^2)=0, \\ 2y-4y(x^2+y^2)=0. \end{cases}$$

Из этой системы уравнений найдем следующие стационарные точки: $M_1(0, 0)$, $M_2(0; 1/\sqrt{2})$, $M_3(0; -1/\sqrt{2})$, $M_4(2; 0)$, $M_5(-2; 0)$. Тогда $f(M_1)=0$, $f(M_2)=f(M_3)=1/4$, M_4, $M_5\notin V$.

Выясним, какие значения принимает функция $f(M)$ в точках окружности $x^2+y^2=3$. Так как $y^2=3-x^2$, $-\sqrt{3}\leqslant x\leqslant\sqrt{3}$, то на окружности $f=8x^2+3-x^2-9=7x^2-6$, $-\sqrt{3}\leqslant x\leqslant\sqrt{3}$. Отсюда следует, что на окружности наименьшее значение равно -6 и достигается в точках $M_6(0; \sqrt{3})$; $M_7(0; -\sqrt{3})$, а наибольшее значение равно 15 и достигается в точках $M_8(\sqrt{3}; 0)$, $M_9(-\sqrt{3}, 0)$.

Искомые наименьшее и наибольшее значения равны 6 и 15 и достигаются соответственно в точках

$$M_4(\sqrt{3}; 0),\ M_5(-\sqrt{3}; 0) \text{ и } M_6(0; \sqrt{3}),\ M_7(0; -\sqrt{3}). \ ●$$

6.9. Системы функциональных уравнений и неравенств

Рассмотрим систему функциональных уравнений и неравенств

$$\begin{cases} \Phi_1(x_1, x_2, \dots, x_n) = 0, \\ \dots\dots\dots\dots \\ \Phi_k(x_1, x_2, \dots, x_n) = 0, \\ \Phi_{k+1}(x_1, x_2, \dots, x_n) \leqslant 0, \\ \dots\dots\dots\dots \\ \Phi_l(x_1, x_2, \dots, x_n) \leqslant 0 \end{cases}$$

(в частности, эта система может содержать только уравнения или только неравенства).

Обозначим через V множество всех решений этой системы, т. е. $V = \{M \in \mathbf{R}^n \mid \Phi_i(M) = 0,\ i = 1, 2, \dots, k;$ $\Phi_i(M) \leqslant 0,\ i = k+1, \dots, l\}$.

Предположим, что функции $\Phi_i(M)$, $i = 1, 2, \dots, k$, $k+1, \dots, l$, непрерывны на всем пространстве $\mathbf{R}^n$. Тогда:

1) множество V замкнуто;

2) если множество V ограничено, а функция $f(M)$ непрерывна на множестве V, то она на этом множестве имеет наименьшее и наибольшее значения.

Рассмотрим множество $Q = \{M \in \mathbf{R}^n \mid \Phi_i(M) \leqslant 0,\ i = 1, 2, \dots, l\}$, где функции $\Phi_i(M)$, $i = 1, 2, \dots, l$, непрерывны на всем пространстве $\mathbf{R}^n$.

Если $\Phi_i(M_0) < 0$, $i = 1, 2, \dots, l$, то M_0 является внутренней точкой множества Q. Если же M_0 — граничная точка множества Q, то $\Phi_i(M_0) = 0$ для некоторого i.

○ Рассмотрим, например, множество

$$Q = \{M(x_1; x_2; x_3) \in \mathbf{R}^3 \mid \Phi_1(M) = x_1^2 + x_2^2 + x_3^2 - 4 \leqslant 0,$$
$$\Phi_2(M) = x_1^2 + x_2^2 - 2x_3 \leqslant 0\}.$$

Точка $M_1(0; 1; 1)$ является внутренней точкой множества Q, так как $\Phi_1(M_1) = -2 < 0$, $\Phi_2(M_1) = -1 < 0$. Точки $M_2(0;$ $0; 2)$ и $M_3(1; 1; 1)$ принадлежат границе множества Q, так как $\Phi_1(M_2) = 0$, $\Phi_2(M_2) = -4 < 0$, $\Phi_1(M_3) = -1 < 0$, $\Phi_2(M_3) = 0$. ●

6.10. Особые точки множества

Точку M_0 из множества

$$V = \{M \in \mathbf{R}^n \mid \Phi_i(M) = 0,\ i = 1, 2, \dots, k,$$
$$\Phi_i(M) \leqslant 0,\ i = k+1, \dots, l\}$$

называют *особой точкой* этого множества, если в точке M_0 линейно зависимы градиенты тех функций $\Phi_i(M)$, которые в ней обращаются в 0.

○ Например, $M_0(2;\ 0;\ 0)$ — особая точка множества

$$V=\{M(x_1;\ x_2;\ x_3)\in \mathbf{R}^3 \mid \Phi_1(M)=x_1^2+x_2^2+x_3^2-4=0,\ \Phi_2(M)=x_1^2+x_2^2-2x_1\leqslant 0\}.$$

Действительно, $\Phi_1(M_0)=0$, $\Phi_2(M_0)=0$, $\operatorname{grad}\Phi_1|_{M_0}=(4;\ 0;\ 0)$, $\operatorname{grad}\Phi_2|_{M_0}=(2;\ 0;\ 0)$. Так как $\operatorname{grad}\Phi_1|_{M_0}=2\operatorname{grad}\Phi_2|_{M_0}$, то эти векторы образуют линейно зависимую систему. ●

Точка $M_0\in\mathbf{R}^n$ является особой точкой множества $P=\{M\in\mathbf{R}^n \mid \Phi_i(M)=0,\ i=1,\ 2,\ \ldots,\ k\}$ тогда и только тогда, когда

$$\Phi_i(M_0)=0,\ i=1,\ 2,\ \ldots,\ k,$$

$$\sum_{i=1}^{k}\mu_i\operatorname{grad}\Phi_i|_{M_0}=\theta\ (\mu_1,\ \ldots,\ \mu_k\text{ — ненулевой набор чисел}).$$

○ Найдем, например, особые точки множества

$$P=\{M(x_1,\ x_2,\ x_3)\in\mathbf{R}^3 \mid \Phi_1(M)=x_1^2+16x_2^2+4x_3^2-16=0,\ \Phi_2(M)=x_1^2+x_2^2+x_3^2-4=0\}.$$

Точка M_0 является особой точкой множества P тогда и только тогда, когда

$$\begin{cases}\Phi_1(M_0)=\Phi_2(M_0)=0,\\ \mu_1\operatorname{grad}\Phi_1|_{M_0}+\mu_2\operatorname{grad}\Phi_2|_{M_0}=\theta.\end{cases}$$

Следовательно, имеем систему уравнений

$$\begin{cases}x_1^2+16x_2^2+4x_3^2-16=0,\\ x_1^2+\ \ x_2^2+\ \ x_3^2-\ \ 4=0,\\ \mu_1(2x_1)+\mu_2(2x_1)=0,\\ \mu_1(32x_2)+\mu_2(2x_2)=0,\\ \mu_1(8x_3)+\mu_2(2x_3)=0.\end{cases}$$

Решив эту систему, найдем две особые точки $M_1(0;\ 0;\ -2)$, $M_2(0;\ 0;\ 2)$ множества P. ●

Точка $M_0\in\mathbf{R}^n$ является особой точкой множества $Q=\{M\in\mathbf{R}^n \mid \Phi_i(M)\leqslant 0,\ i=1,\ 2,\ \ldots,\ k\}$ тогда и только тогда, когда

$$\begin{cases}\Phi_i(M_0)\leqslant 0,\ i=1,\ 2,\ \ldots,\ k,\\ \sum_{i=1}^{k}\mu_i\operatorname{grad}\Phi_i\mid M_0=\theta,\\ \mu_i\Phi_i(M_0)=0,\ i=1,\ 2,\ \ldots,\ k\ (\mu_1,\ \mu_2,\ \ldots,\ \mu_k\text{ — ненулевой набор чисел}).\end{cases}$$

6.11. Условные экстремумы функций нескольких переменных

Пусть функция $f(M)$ определена на множестве $V \subseteq \mathbf{R}^n$. Точка $M_0 \in V$ называется *точкой условного локального минимума (максимума)* функции $f(M)$ на множестве V, если существует окрестность $S_r(M_0)$ точки M_0 такая, что для всех точек $M \in S_r(M_0) \cap V$ выполняется неравенство $f(M_0) \leqslant f(M)$ $(f(M_0) \geqslant f(M))$.

Точки условного локального минимума и условного локального максимума функции $f(M)$ на множестве V называются *точками условного экстремума функции* $f(M)$ на множестве V.

Точка экстремума функции всегда является точкой условного экстремума. Если же точка условного экстремума функции $f(M)$ на множестве V является внутренней точкой этого множества, то она — точка экстремума функции $f(M)$.

Рассмотрим множество

$$V = \{M \in \mathbf{R}^n \mid \Phi_i(M) = 0,\ i = 1,\ 2,\ \ldots,\ k;\\ \Phi_i(M) \leqslant 0,\ i = k+1,\ \ldots,\ l\}.$$

Точка $M_0 \in V$ называется *условно стационарной точкой* функции $f(M)$ на множестве V, если в ней градиент $f(M)$ раскладывается по градиентам тех функций $\Phi_i(M)$, которые в точке M_0 обращаются в 0.

○ Например, точка $M_0(4;\ 6)$ является условно стационарной точкой функции $f(M) = (x-2)^2 + (y-3)^2$ на множестве $Q = \{M(x, y) \in \mathbf{R}^2 \mid \Phi(M) = x^2 + y^2 - 52 \leqslant 0\}$. Действительно, $\Phi(M_0) = 0$, $\operatorname{grad} f|_{M_0} = (4, 6)$, $\operatorname{grad} \Phi|_{M_0} = (8, 12)$, т. е. $\operatorname{grad} f|_{M_0} = \frac{1}{2} \operatorname{grad} \Phi|_{M_0}$. ●

Предположим, что функция $f(M)$ определена на множестве $P = \{M(x_1,\ x_2,\ \ldots,\ x_n) \mid \Phi_i(x_1,\ x_2,\ \ldots,\ x_n) = 0,\ i = 1,\ \ldots,\ k\}$. В этом случае функция

$$L(x_1,\ x_2,\ \ldots,\ x_n,\ y_1,\ y_2,\ \ldots,\ y_k) = \\ = f(x_1,\ x_2,\ \ldots,\ x_n) - \sum_{i=1}^{k} y_i \Phi_i(x_1,\ x_2,\ \ldots,\ x_n)$$

называется *функцией Лагранжа*. Точка $M_0 \in \mathbf{R}^n$ является условно стационарной точкой функции $f(M)$ на множестве P тогда и только тогда, когда

$$\begin{cases} \dfrac{\partial L}{\partial x_j}=0, \ j=1, 2, \ldots, n, \\ \dfrac{\partial L}{\partial y_i}=0, \ i=1, 2, \ldots, k. \end{cases}$$

Например, найдем условно стационарные точки функции $f(x_1, x_2)=3x_1^2+2x_2^2-3x_1+1$ на множестве $P=\{M(x_1, x_2)\,|\,\Phi(M)=x_1^2+x_2^2-4=0\}$. Функция Лагранжа в этом случае имеет вид

$$L(x_1, x_2, y)=3x_1^2+2x_2^2-3x_1+1-y(x_1^2+x_2^2-4).$$

Тогда

$$\frac{\partial L}{\partial x_1}=6x_1-3-2yx_1=0; \ \frac{\partial L}{\partial x_2}=4x_2-2yx_2=0;$$
$$\frac{\partial L}{\partial y}=-(x_1^2+x_2^2-4)=0,$$

откуда найдем следующие четыре условно стационарные точки:

$M_1(3/2; \sqrt{7}/2)$, $M_2(3/2; -\sqrt{7}/2)$, $M_3(2; 0)$; $M_4(-2; 0)$. ●

Необходимое условие условного экстремума функций на множестве

$$V=\{M\in \mathbf{R}^n\,|\,\Phi_i(M)=0, \ i=1, \ldots, k; \\ \Phi_i(M)\leqslant 0, \ i=k+1, \ldots, l\}.$$

Предположим, что M_0—неособая точка условного экстремума функции $f(M)$ на множестве V. Если функции $f(M)$ и $\Phi_i(M)$, $i=1, 2, \ldots, k, k+1, \ldots, l$, дифференцируемы в точке M_0, то эта точка является условно стационарной точкой функции $f(M)$ на множестве V.

Условно стационарная точка функции $f(M)$ на множестве V может не быть точкой условного экстремума. Однако все неособые точки условного экстремума находятся среди условно стационарных точек (при условии дифференцируемости функций $f(M)$ и $\Phi_i(M)$, $i=1, 2, \ldots, k, k+1, \ldots, l$).

6.12. Наименьшее и наибольшее значения функции на множестве решений системы уравнений и неравенств

Рассмотрим функцию $f(M)$ и множество

$$V=\{M\in \mathbf{R}^n\,|\,\Phi_i(M)=0, \ i=1, 2, \ldots, k; \\ \Phi_i(M)\leqslant 0, \ i=k+1, \ldots, l\}.$$

Предположим, что функции $\Phi_i(M)$, $i=1, 2, \ldots, k, k+1, \ldots, l$, дифференцируемы в любой точке $M \in \mathbf{R}^n$, а функция $f(M)$ дифференцируема на множестве V.

Отыскать наименьшее и наибольшее значения функции $f(M)$ на множестве V можно следующим образом:

1) выявить все особые точки множества V;

2) найти все условно стационарные точки функции $f(M)$ на множестве V;

3) в каждой из найденных точек вычислить значения функции $f(M)$ и выбрать среди них две точки, в которых функция имеет наибольшее и наименьшее значения.

○ Найдем, например, наименьшее и наибольшее значения функции $f(M)=(x-2)^2+(y-3)^2$ на множестве $Q=\{M(x, \;\; y)\,|\,\Phi(M)=x^2+y^2-52\leqslant 0\}$. Особых точек у множества Q нет. Поэтому найдем условно стационарные точки $f(M)$ на множестве Q. Для этого составим систему

$$\begin{cases} \operatorname{grad} f=\lambda \operatorname{grad}\Phi, \\ \lambda\Phi(M)=0, \end{cases}$$

откуда получим

$$\begin{cases} 2(x-2)=\lambda(2x), \\ 2(y-3)=\lambda(2y), \\ \lambda(x^2+y^2-52)=0, \\ x^2+y^2-52\leqslant 0. \end{cases}$$

Решив эту систему, найдем точки $M_1(2;\ 3)$, $M_2(4;\ 6)$, $M_3(-4;\ -6)$. Так как $f(M_1)=0$, $f(M_2)=13$, $f(M_3)=117$, то наименьшее значение $f(M)$ равно 0 и достигается в точке $M_1(2;\ 3)$, а наибольшее значение равно 117 и достигается в точке $M_3(-4;\ -6)$.

6.13. Экстремумы выпуклых и вогнутых функций

Пусть V — некоторое выпуклое множество n-мерных точек, а $f(M)$ — функция, определенная на множестве V.

Если $f(M)$ — вогнутая (выпуклая) функция на множестве V, то в любой точке условного локального максимума (минимума) она достигает своего наибольшего (наименьшего) значения.

Любая стационарная точка дифференцируемой вогнутой (выпуклой) функции $f(M)$ является точкой локального максимума (минимума) этой функции.

Отсюда, в частности, следует, что если стационарная точка дифференцируемой вогнутой (выпуклой) функции $f(M)$ принадлежит множеству V, то в этой точке функция $f(M)$ достигает своего наибольшего (наименьшего) значения.

○ Рассмотрим, например, вогнутую функцию $f(M) = 10x_1 + 16x_2 - x_1^2 - x_2^2$ на выпуклом множестве $V = \{M(x_1, x_2) \mid x_1 + 2x_2 \leqslant 21,\ 5x_1 + 2x_2 \leqslant 42\}$.

Точка $M_0(5;\ 8)$ принадлежит множеству V и является стационарной точкой функции $f(M)$, так как $\operatorname{grad} f|_{M_0} = \theta$. Значит, функция $f(M)$ достигает в точке $M_0(5;\ 8)$ своего наибольшего значения $f(M_0) = 89$. ●

РАЗДЕЛ VII. ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ И ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

7.1. Неопределенный интеграл

Функция $F(x)$ называется *первообразной функцией* для данной функции $y=f(x)$ на отрезке $[a, b]$, если в каждой точке этого отрезка ее производная равна $f(x)$:

$$F'(x)=f(x).$$

○ **Пример.** Производная от функции $x^3/3$ равна x^2. Поэтому, по определению, функция $x^3/3$ является первообразной для функции x^2. ●

Теорема о существовании первообразной. *Каждая непрерывная функция имеет бесконечное множество первообразных функций, отличающихся друг от друга на постоянную величину.*

Общее выражение $F(x)+C$ для всех первообразных функций от данной функции $f(x)$ называется *неопределенным интегралом* от этой функции и обозначается

$$\int f(x)\,\mathrm{d}x,$$

где $\int$ — знак интеграла, $f(x)$ — подынтегральная функция, $f(x)\,\mathrm{d}x$ — подынтегральное выражение.

7.2. Таблица основных интегралов

1) $\int x^\alpha\,\mathrm{d}x=\frac{x^{\alpha+1}}{\alpha+1}+C \quad (\alpha\neq-1);$

2) $\int \frac{\mathrm{d}x}{x}=\ln|x|+C, \quad \int\frac{\mathrm{d}x}{x+a}=\ln|x+a|+C;$

3) $\int \mathrm{e}^x\,\mathrm{d}x=\mathrm{e}^x+C;$

4) $\int a^x\,\mathrm{d}x=\frac{a^x}{\ln a}+C;$

5) $\int \sin x \, dx = -\cos x + C$;

6) $\int \cos x \, dx = \sin x + C$;

7) $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$;

8) $\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$;

9) $\int \operatorname{tg} x \, dx = -\ln|\cos x| + C$;

10) $\int \operatorname{ctg} x \, dx = \ln|\sin x| + C$;

11) $\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin \frac{x}{a} + C$;

12) $\int \frac{dx}{a^2 + x^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C$;

13) $\int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| + C$;

14) $\int \frac{dx}{\sqrt{x^2 + a^2}} = \ln \left| x + \sqrt{x^2 + a^2} \right| + C$.

7.3. Свойства неопределенного интеграла

1°. Производная от неопределенного интеграла равна подынтегральной функции:

$$\left(\int f(x) \, dx \right)' = f(x).$$

2°. Дифференциал от неопределенного интеграла равен подынтегральному выражению:

$$d \left(\int f(x) \, dx \right) = f(x) \, dx.$$

3°. Неопределенный интеграл от дифференциала непрерывно дифференцируемой функции равен самой этой функции с точностью до постоянного слагаемого:

$$\int d\varphi(x) = \varphi(x) + C.$$

4°. Постоянный множитель можно выносить за знак неопределенного интеграла:

$$\int Af(x) \, dx = A \int f(x) \, dx.$$

5°. Неопределенный интеграл от алгебраической суммы конечного числа непрерывных функций равен сумме неопределенных интегралов от этих функций:

$$\int [f_1(x)+f_2(x)+\ldots+f_n(x)]\,dx=\int f_1(x)\,dx+\ldots+\int f_n(x)\,dx.$$

7.4. Методы интегрирования

Метод разложения. Метод основан на разложении подынтегральной функции на сумму функций, каждая из которых является табличной.

○ **Примеры.**

1. $\int (x+\sqrt{x})^2\,dx=\int (x^2+2x\sqrt{x}+x)\,dx=$
$$=\int x^2\,dx+2\int x^{3/2}\,dx+\int x\,dx=\frac{x^3}{3}+2\frac{x^{5/2}}{5/2}+\frac{x^2}{2}=$$
$$=\frac{x^3}{3}+\frac{4}{5}x^2\sqrt{x}+\frac{x^2}{2}+C.$$

2. $\int \frac{dx}{\sin^2 x\cos^2 x}=\int \frac{\sin^2 x+\cos^2 x}{\sin^2 x\cos^2 x}\,dx=$
$$=\int \frac{dx}{\cos^2 x}+\int \frac{dx}{\sin^2 x}=\operatorname{tg} x-\operatorname{ctg} x+C. \bullet$$

Метод замены переменной. Вводят новую переменную с помощью соотношения $x=\varphi(t))$ и данный интеграл преобразуют следующим образом:

$$\int f(x)\,dx=\int f[\varphi(t)\,\varphi'(t)\,dt],$$

где $\varphi(t)$ — дифференцируемая функция.

○ **Примеры.**

1. Вычислить $\int \sin 5x\,dx$. Делаем замену переменной: $x=t/5$, $dx=dt/5$. Далее имеем

$$\int \sin 5x\,dx=\frac{1}{5}\int \sin t\,dt=-\frac{1}{5}\cos t+C=-\frac{1}{5}\cos 5x+C.$$

2. Вычислить $\int x\sqrt{x-2}\,dx$.

Обозначим $\sqrt{x-2}=t$. Тогда $x-2=t^2$, $x=t^2+2$, $dx=2t\,dt$. Далее имеем

$$\int x\sqrt{x-2}\,dx=\int (t^2+2)\,t\,2t\,dt=2\left[\int t^4\,dt+2\int t^2\,dt\right]=$$
$$=\frac{2t^5}{5}+\frac{4t^3}{3}+C=\frac{2}{5}(x-2)^{5/2}+\frac{4}{3}(x-2)^{3/2}+C. \bullet$$

Метод интегрирования по частям. Интегрирование осуществляют с помощью формулы

$$\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u,$$

где u, v — дифференцируемые функции.

Для применения этой формулы подынтегральное выражение разбивают на две части, одну из которых принимают за u, а другую — за $\mathrm{d}v$ так, чтобы интеграл $\int v\,\mathrm{d}u$ вычислялся проще, чем исходный.

○ **Пример.** Вычислить $\int x \ln x\,\mathrm{d}x$.

Обозначим $u = \ln x$, $\mathrm{d}v = x\,\mathrm{d}x$. Тогда $\mathrm{d}u = \frac{1}{x}\mathrm{d}x$, $v = \int x\,\mathrm{d}x = \frac{x^2}{2}$. Далее имеем

$$\int x \ln x\,\mathrm{d}x = \frac{x^2}{2}\ln x - \frac{1}{2}\int \frac{x^2\,\mathrm{d}x}{x} = \frac{x^2}{2}\ln x - \frac{1}{4}x^2 + C. \quad ●$$

Интегрирование правильных рациональных дробей. Интегрирование правильных рациональных дробей $\frac{f(x)}{\varphi(x)}$ начинают с разложения их на простейшие рациональные дроби $\frac{A}{(x-a)^n}$, $\frac{Ax+B}{(x^2+px+q)^n}$, где $n = 1, 2, 3, \ldots$ — натуральные числа; x^2+px+q не разлагается на действительные множители, т. е. имеет только комплексно-сопряженные корни.

Простейшие рациональные дроби интегрируются с помощью следующих формул:

$$\int \frac{A}{x-a}\,\mathrm{d}x = A \ln|x-a| + C,$$

$$\int \frac{A}{(x-a)^n}\,\mathrm{d}x = \frac{A}{(1-n)(x-a)^{n-1}} + C,$$

$$\int \frac{Ax+B}{x^2+px+q}\,\mathrm{d}x = \frac{A}{2}\ln|x^2+px+q| + \frac{2B-Ap}{\sqrt{4q-p^2}}\,\mathrm{arctg}\,\frac{2x+p}{\sqrt{4q-p^2}} + C.$$

Дроби вида $\frac{Ax+B}{(x^2+px+q)^n}$ интегрируют с помощью формулы понижения степени

$$\int \frac{Ax+B}{(x^2+px+q)^n}\,\mathrm{d}x = \frac{\left(B-\frac{Ap}{2}\right)\left(x+\frac{p}{2}\right)-A\left(q-\frac{p^2}{4}\right)}{2(n-1)\left(q-\frac{p^2}{4}\right)(x^2+px+q)^{n-1}} + \frac{\left(B-\frac{Ap}{2}\right)(2n-3)}{2(n-1)\left(q-\frac{p^2}{4}\right)}\int \frac{\mathrm{d}x}{(x^2+px+q)^{n-1}}.$$

Разложение рациональных дробей на простейшие основано на том, что любой многочлен $\varphi(x)=a_0x^n+a_1x^{n-1}+\ldots$ $\ldots+a_n$ может быть записан в виде произведения

$$\varphi(x)=a_0(x-x_1)(x-x_2)\ldots(x-x_n), \qquad (7.1)$$

где $x_1, x_2, \ldots, x_n$ — корни уравнения $\varphi(x)=0$.

Эти корни могут быть действительными и комплексными. Значения корней в разложении (7.1) могут повторяться (кратные корни). Вместе с комплексным корнем $x_j=\alpha_j+i\beta_j$ в выражении (7.1) имеется комплексно-сопряженный корень $\bar{x}_j=\alpha_j-i\beta_j$. Произведение линейных множителей $(x-x_j)(x-\bar{x}_j)$, содержащих комплексно-сопряженные корни, может быть записано в виде x^2+px+q. Многочлен (7.1) в этом случае принимает следующий вид:

$$\varphi(x)=a_0(x-x_1)^{m_1}(x-x_2)^{m_2}\ldots(x-x_k)^{m_k}\ldots$$
$$\ldots(x^2+p_1x+q_1)^{m_{k+1}}\ldots(x^2+p_lx+q_l)^{m_{k+l}}, \qquad (7.2)$$

где $x_1, x_2, \ldots, x_k$ — действительные числа; m_j — кратности простых и комплексно-сопряженных корней ($j=1, 2, \ldots$ $\ldots, k+l$).

Правильную рациональную дробь со знаменателем, имеющим вид (7.2), записывают как сумму простейших рациональных дробей:

$$\frac{f(x)}{\varphi(x)}=\frac{f(x)}{a_0(x-x_1)^{m_1}\ldots(x^2+p_lx+q_l)^{m_{k+l}}}=$$
$$=\frac{A_1}{(x-x_1)^{m_1}}+\frac{A_2}{(x-x_1)^{m_1-1}}+\ldots+\frac{A_{m_1}}{x-x_1}+\ldots$$
$$\ldots+\frac{B_1x+C_1}{(x^2+p_1x+q_1)^{m_{k+1}}}+\ldots+\frac{B_{m_{k+l}}x+C_{m_{k+l}}}{x^2+p_lx+q_l}. \qquad (7.3)$$

Приводя правую часть выражения (7.3) к общему знаменателю и приравнивая коэффициенты при одинаковых степенях переменной x полученного выражения и исходной дроби, находят значения A_1, A_2, ..., $B_{m_{k+l}}$, $C_{m_{k+l}}$.

○ **Пример.** Вычислить $\int \frac{2x^2-x+1}{(x-1)(x^2-2x+3)}\,dx$.

Имеем:

$$\frac{2x^2-x+1}{(x-1)(x^2-2x+3)}=\frac{A}{x-1}+\frac{Bx+C}{x^2-2x+3}=$$
$$=\frac{A(x^2-2x+3)+(x-1)(Bx+C)}{(x-1)(x^2-2x+3)}=$$
$$=\frac{(A+B)x^2+(-2A-B+C)x+3A-C}{(x-1)(x^2-2x+3)};$$

$$\begin{cases} A+B=2, \\ -2A-B+C=-1, \\ 3A-C=1, \end{cases} \Rightarrow A=B=1,\ C=2;$$

$$\int\frac{2x^2-x+1}{(x-1)(x^2-2x+3)}\,dx=\int\frac{dx}{x-1}+\int\frac{x+2}{x^2-2x+3}=\ln|x-1|+$$
$$+\int\frac{(1/2)(2x-2)+3}{x^2-2x+3}\,dx=\ln|x-1|+\frac{1}{2}\ln|x^2+2x+3|+$$
$$+3\int\frac{dx}{x^2-2x+3}=\ln|x-1|+\frac{1}{2}\ln|x^2+2x+3|+$$
$$+3\int\frac{dx}{(x-1)^2+2}=\ln|x-1|+\frac{1}{2}\ln|x^2-2x+3|+$$
$$+\frac{3}{\sqrt{2}}\operatorname{arctg}\frac{x-1}{\sqrt{2}}+C. \quad ●$$

Интегрирование тригонометрических функций. Интегралы вида $\int \cos^m x \sin^n x\,dx$, где m, n — натуральные числа, вычисляют в зависимости от четности степеней m и n следующим образом.

1) Если m или n — нечетное, то используют замену переменного

$$t=\sin x \text{ при } m \text{ нечетном},$$
$$t=\cos x \text{ при } n \text{ нечетном}.$$

○ **Пример.** $\int \cos x \sin^2 x\,dx = |\,t=\sin x,\quad dt=\cos x\,dx\,| =$
$$=\int t^2\,dt=\frac{t^3}{3}+C=\frac{\sin^3 x}{3}+C. \quad ●$$

2) Если m и n — четные, то используют формулы понижения степеней

$$\sin^2 x=\frac{1-\cos 2x}{2},\quad \cos^2 x=\frac{1+\cos 2x}{2},\quad \sin x\cos x=\frac{\sin 2x}{2}.$$

○ **Пример.** $\int \cos^2 x \sin^2 x \, dx = \int (\cos x \sin x)^2 \, dx =$

$$= \int \frac{\sin^2 2x}{4} dx = \frac{1}{4} \int \frac{1 - \cos 4x}{2} dx =$$

$$= \frac{1}{8}\left(x - \frac{1}{4}\sin 4x\right) + C. \quad \bullet$$

Интегралы вида $\int R(\sin x, \cos x)\, dx$, где $R(\sin x, \cos x)$ — рациональная функция от $\sin x$, $\cos x$, приводят к интегралам от рациональных функций переменной t с помощью следующей подстановки: $t = \operatorname{tg} \frac{x}{2}$. Тогда

$$dx = \frac{2dt}{1+t^2}, \quad \sin x = \frac{2t}{1+t^2}, \quad \cos x = \frac{1-t^2}{1+t^2}.$$

○ **Пример.** $\int \frac{dx}{\sin x} = \int \frac{2dt\,(1+t^2)}{(1+t^2)\,2t} = \int \frac{dt}{t} = \ln|t| + C =$

$$= \ln\left|\operatorname{tg}\frac{x}{2}\right| + C. \quad \bullet$$

Интегрирование иррациональных функций. Некоторые интегралы от иррациональных функций могут быть приведены к интегралам от рациональных функций с помощью следующих подстановок:

Интеграл	Подстановка
$\int R(x, \sqrt[n]{ax+b})\, dx$,	$ax + b = t^n$,
$\int R(x^m, \sqrt[n]{ax^m+b})\, x^{m-1}\, dx$,	$ax^m + b = t^n$,
$\int R(x, \sqrt{a^2 - x^2})\, dx$,	$x = a \sin t$,
$\int R(x, \sqrt{a^2 + x^2})\, dx$,	$x = a \operatorname{tg} t$.

○ **Пример.**

$$\int x\sqrt{1-x^2}\, dx = |x = \sin t, \ dx = \cos t\, dt| =$$

$$= \int \sin t \cos^2 t\, dt = |\cos t = u, \ -\sin t\, dt = du| =$$

$$= -\int u^2 du = -\frac{u^3}{3} + C = -\frac{\cos^3 t}{3} + C =$$

$$= \frac{-(\sqrt{1-\sin^2 t})^3}{3} + C = -\frac{(\sqrt{1-x^2})^3}{3} + C. \quad \bullet$$

7.5. Определенный интеграл. Основные понятия

Интегральной суммой для функции $y=f(x)$ на отрезке $[a, b]$ называется выражение $\sum_{i=1}^{n} f(\bar{x}_i)\Delta x_i$, где n — число «элементарных» отрезков, на которые разбивается отрезок $[a, b]$; $\bar{x}_i$ — произвольная точка внутри отрезка $[x_{i-1}, x_i]$, длина которого равна Δx_i (рис. 7.1).

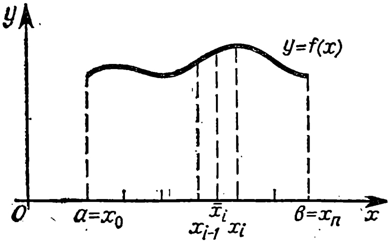


Рис. 7.1

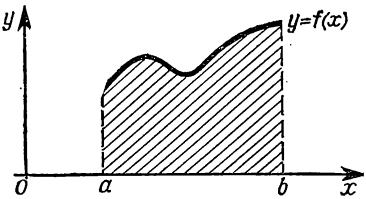


Рис. 7.2

Определенным интегралом функции $y=f(x)$ на отрезке $[a, b]$ называется предел интегральной суммы при стремлении к нулю длины наибольшего «элементарного» отрезка

$$\int_a^b f(x)\,dx = \lim_{\max|\Delta x_i| \to 0} \sum_{i=1}^{n} f(\bar{x}_i)\Delta x_i.$$

Число a называют нижним *пределом интегрирования*, b — верхним пределом.

Определенный интеграл имеет геометрический смысл. Интеграл $\int_a^b f(x)\,dx$ численно равен площади криволинейной трапеции, ограниченной графиком функции $y=f(x)\geqslant 0$, осью Ox и прямыми $x=a$, $x=b$ (рис. 7.2).

7.6. Основные свойства определенного интеграла

1°. При перестановке пределов интегрирования интеграл меняет знак на обратный:

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

2°. Каковы бы ни были числа a, b, c, имеет место равенство

$$\int_a^b f(x)\,\mathrm{d}x=\int_a^c f(x)\,\mathrm{d}x+\int_c^b f(x)\,\mathrm{d}x.$$

3°. Постоянный множитель можно выносить за знак определенного интеграла:

$$\int_a^b Af(x)\,\mathrm{d}x=A\int_a^b f(x)\,\mathrm{d}x.$$

4°. Определенный интеграл от алгебраической суммы конечного числа непрерывных функций равен такой же алгебраической сумме определенных интегралов от этих функций:

$$\int_a^b [f(x)+g(x)-h(x)]\,\mathrm{d}x=\int_a^b f(x)\,\mathrm{d}x+\int_a^b g(x)\,\mathrm{d}x-\int_a^b h(x)\,\mathrm{d}x.$$

5°. **Теорема о среднем значении.** *Определенный интеграл от непрерывной функции равен произведению длины промежутка интегрирования на значение подынтегральной функции при некотором промежуточном значении аргумента*

$$\int_a^b f(x)\,\mathrm{d}x=(b-a)f(c),$$

где $c\in\,]a,\,b[$.

6°. Если $F(x)$ — какая-либо первообразная от непрерывной функции $f(x)$, то справедлива формула Ньютона—Лейбница

$$\int_a^b f(x)\,\mathrm{d}x=F(b)-F(a).$$

7.7. Вычисление определенных интегралов

Основным способом вычисления определенных интегралов является определение первообразной для подынтегральной функции и использование формулы Ньютона — Лейбница, которая может быть записана в виде

$$\int_a^b f(x)\,\mathrm{d}x=F(x)\Big|_a^b=F(b)-F(a).$$

Определение первообразной для многих функций может быть сложным процессом: не все функции имеют первообразные в виде элементарных функций. Поэтому для вычисления определенных интегралов используют приближенные формулы.

Разбивают отрезок интегрирования $[a, b]$ на n равных частей длиной $h=(b-a)/n$ и используют одну из следующих формул:

1) формула прямоугольников

$$\int_a^b y\,dx \simeq h(y_0+y_1+\ldots+y_{n-1});$$

2) формула трапеций

$$\int_a^b y\,dx \simeq h\left(\frac{y_0+y_n}{2}+y_1+y_2+\ldots+y_{n-1}\right);$$

3) формула парабол (Симпсона) (n—четное)

$$\int_a^b y\,dx \simeq \frac{h}{3}(y_0+4y_1+2y_2+4y_3+\ldots+2y_{n-2}+4y_{n-1}+y_n)$$

(чем больше n, тем точнее результат вычисления определенного интеграла).

7.8. Геометрические приложения определенного интеграла

Площадь криволинейной трапеции, ограниченной кривой $y=f(x)$, осью Ox, прямыми $x=a$, $x=b$, находят по формуле

$$S=\int_a^b f(x)\,dx.$$

Объем тела, образованного вращением кривой $y=f(x)$, ограниченной прямыми $x=a$, $x=b$ при $a<x<b$, вокруг оси Ox, равен

$$V_x=\pi\int_a^b y^2\,dx.$$

Объем тела, образованного вращением кривой $x=\varphi(y)$, ограниченной прямыми $y=c$, $y=d$ при $c<y<d$, вокруг

оси Oy, равен

$$V_y = \pi \int_c^d x^2 \, dy.$$

Длину дуги плоской кривой $y = f(x)$, ограниченной прямыми $x = a$, $x = b$, определяют по формуле

$$l = \int_a^b \sqrt{1 + (y')^2} \, dx.$$

Площадь поверхности, образованной вращением кривой $y = f(x)$, ограниченной прямыми $x = a$, $x = b$, вокруг оси Ox, равна

$$S_x = 2\pi \int_a^b y \sqrt{1 + (y')^2} \, dx.$$

Площадь поверхности, образованной вращением кривой $x = \varphi(y)$, ограниченной прямыми $y = c$, $y = d$, вокруг оси Oy, равна

$$S_y = 2\pi \int_c^d x \sqrt{1 + (x')^2} \, dy.$$

7.9. Несобственные интегралы

Пусть функция $f(x)$ определена и непрерывна на промежутке $[a, +\infty[$ (рис. 7.3). Рассмотрим интеграл $\int_a^b f(x) \, dx$. Предел

$$\lim_{b \to +\infty} \int_a^b f(x) \, dx$$

называют *несобственным интегралом первого рода* от функции $f(x)$ на промежутке $[a, +\infty[$ и обозначают $\int_a^{+\infty} f(x) \, dx$, т. е.

$$\int_a^{+\infty} f(x) \, dx = \lim_{b \to +\infty} \int_a^b f(x) \, dx.$$

Если указанный предел конечен, то говорят, что несобственный интеграл $\int_a^{+\infty} f(x) \, dx$ сходится; если бесконечен или не существует, то расходится.

Аналогичным способом определяют несобственный интеграл первого рода для промежутка $]-\infty, b]$ (рис. 7.4):

$$\int_{-\infty}^{b} f(x)\,\mathrm{d}x = \lim_{a \to -\infty} \int_{a}^{b} f(x)\,\mathrm{d}x.$$

Пусть функция $f(x)$ определена и непрерывна на интервале $]-\infty, +\infty[$ и пусть точка $c \in]-\infty, +\infty[$. Тогда сумму

$$\int_{-\infty}^{c} f(x)\,\mathrm{d}x + \int_{c}^{+\infty} f(x)\,\mathrm{d}x \qquad (7.4)$$

называют *несобственным интегралом первого рода* от функции $f(x)$ на интервале $]-\infty, +\infty[$ и обозначают

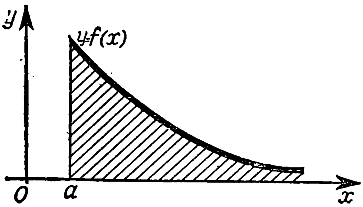


Рис. 7.3

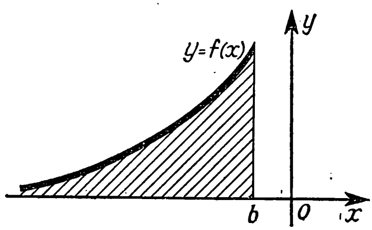


Рис. 7.4

$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$. Этот интеграл сходится, если оба интеграла $\int_{-\infty}^{c} f(x)\,\mathrm{d}x$, $\int_{c}^{+\infty} f(x)\,\mathrm{d}x$ сходятся. В этом случае сумма (7.4) не зависит от выбора точки c.

○ **Примеры.**

1. $\int_{1}^{+\infty} \frac{\mathrm{d}x}{x} = \lim_{b \to +\infty} \int_{1}^{b} \frac{\mathrm{d}x}{x} = \lim_{b \to +\infty} \ln b$ (интеграл расходится).

2. $\int_{-\infty}^{+\infty} \frac{\mathrm{d}x}{1+x^2} = \lim_{\substack{a \to -\infty \\ b \to +\infty}} \int_{a}^{b} \frac{\mathrm{d}x}{1+x^2} =$

$= \lim_{\substack{a \to -\infty \\ b \to +\infty}} (\operatorname{arctg} b - \operatorname{arctg} a) =$

$= \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi$ (интеграл сходится). ●

Пусть функция $f(x)$ определена и непрерывна при $a \leqslant x < b$ и не ограничена в любой окрестности точки

$x=b$ (рис. 7.5). Предел

$$\lim_{\varepsilon\to 0}\int_a^{b-\varepsilon} f(x)\,dx$$

называют *несобственным интегралом второго рода* от функции $f(x)$ на промежутке $[a, b[$.

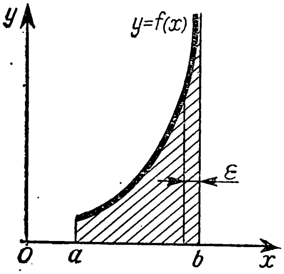


Рис. 7.5

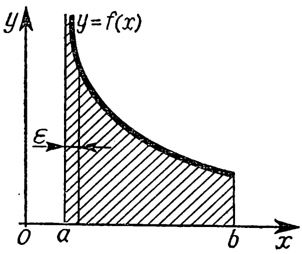


Рис. 7.6

Если этот предел конечен, то несобственный интеграл

$$\int_a^b f(x)\,dx=\lim_{\varepsilon\to 0}\int_a^{b-\varepsilon} f(x)\,dx$$

называют *сходящимся*; если бесконечен или не существует, то *расходящимся*.

Аналогично определяют несобственные интегралы от функций, определенных и непрерывных при $a<x\leqslant b$ (рис. 7.6):

$$\int_a^b f(x)\,dx=\lim_{\varepsilon\to 0}\int_{a+\varepsilon}^{b} f(x)\,dx.$$

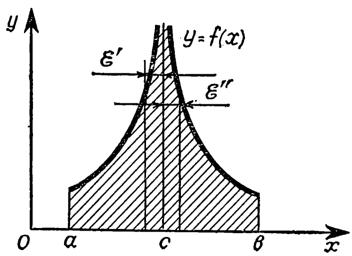


Рис. 7.7

Пусть функция $f(x)$ определена и непрерывна на отрезке $[a, b]$, за исключением точки $c\in\,]a, b[$, в любой окрестности которой она не ограничена (рис. 7.7). Тогда несобственный интеграл от этой функции определяется как сумма двух несобственных интегралов на промежутках $[a, c[$ и $]c, b]$:

$$\int_a^b f(x)\,dx=\int_a^c f(x)\,dx+\int_c^b f(x)\,dx.$$

Этот интеграл сходится, если оба слагаемых сходятся.

○ **Пример.**

$$\int_0^2 \frac{dx}{(x-1)^2} = \lim_{\varepsilon'\to 0} \int_0^{1-\varepsilon'} \frac{dx}{(x-1)^2} + \lim_{\varepsilon''\to 0} \int_{1+\varepsilon''}^{2} \frac{dx}{(x+1)^2} =$$

$$= \lim_{\varepsilon'\to 0} \left(-\frac{1}{x-1}\right)\Big|_0^{1-\varepsilon'} + \lim_{\varepsilon''\leftarrow 0}\left(-\frac{1}{x-1}\right)\Big|_{1+\varepsilon''}^{2} =$$

$$= \lim_{\varepsilon'\to 0}\left(\frac{1}{\varepsilon'}-1\right) + \lim_{\varepsilon''\to 0}\left(-1+\frac{1}{\varepsilon''}\right)$$

(интеграл расходится). ●

7.10. Обыкновенные дифференциальные уравнения

Обыкновенным дифференциальным уравнением называется выражение, связывающее независимую переменную x, функцию y и ее производные.

Порядком дифференциального уравнения называется наивысший порядок производной, входящей в это уравнение.

Дифференциальное уравнение n-го порядка вида

$$y^{(n)} = f(x,\ y,\ y',\ \ldots,\ y^{(n-1)}) \tag{7.5}$$

называется *разрешенным относительно высшей производной*.

Решением дифференциального уравнения n-го порядка называется всякая функция $y=\varphi(x)$, определенная для значений x на конечном или бесконечном интервале, имеющая производные до n-го порядка включительно и такая, что подстановка этой функции и ее производных в дифференциальное уравнение обращает последнее в тождество по x.

Нахождение решений дифференциального уравнения называется *интегрированием* этого дифференциального уравнения.

○ **Пример.** Решением дифференциального уравнения второго порядка $y''+y=0$ является, например, функция $y=\sin x$.

В самом деле, $y'=\cos x$, $y''=-\sin x$. Подставив выражения для y и y'' в уравнение, получаем тождественное равенство $-\sin x+\sin x=0$. Функции $y=C_1\sin x$, $y=C_2\cos x$, где C_1 и C_2 — произвольные постоянные, также являются решением данного уравнения. ●

Во многих случаях требуется находить решение дифференциального уравнения, удовлетворяющего некоторым дополнительным условиям. Например, задача Коши состоит в отыскании решения дифференциального уравнения (7.5), определенного в некоторой окрестности точки x_0 и такого, что

$$y(x_0)=y_0,\ y'(x_0)=y_1,\ \ldots,\ y^{(n-1)}(x_0)=y_{n-1},$$

где $y_0,\ y_1,\ \ldots,\ y_{n-1}$ — заданные числа.

○ **Пример 1.** Найти решение дифференциального уравнения $y'=x$, удовлетворяющего начальным условиям $y(x_0)=y_0$.

Интегрируя левую и правую часть уравнения, находим

$$y=\frac{x^2}{2}+C,$$

где C — произвольная постоянная. Подставляя начальные условия, определяем C: $C=y_0-\frac{x_0^2}{2}$. Искомое решение примет вид

$$y=\frac{x^2}{2}+y_0-\frac{x_0^2}{2}.$$

Это решение определено на всей числовой оси.

Пример 2. Решение дифференциального уравнения $y'+y^2=0$, удовлетворяющего условию $y(0)=1$, имеет вид

$$y(x)=\frac{1}{1+x}.$$

Это решение определено на полуоси $]-1,\ +\infty[$. ●

Кроме задачи Коши для дифференциального уравнения (7.5) решаются также краевые задачи. Например, для дифференциального уравнения второго порядка $y''=f(x,\ y,\ y')$ отыскивают решение на отрезке $[x_0,\ x_1]$ такое, что выполняются граничные (краевые) условия $y(x_0)=y_0,\ y(x_1)=y_1$.

○ **Пример 3.** Найти решение уравнения $y''=4x$, удовлетворяющего граничным условиям $y(0)=1,\ y(1)=\frac{5}{3}$.

Последовательно интегрируя уравнение, находим $y'=2x^2+C_1,\ y=\frac{2}{3}x^3+C_1x+C_2$. Подставим в выражение для y граничные условия. Получим $C_1=0,\ C_2=1$. Искомое решение таково:

$$y=\frac{2}{3}x^3+1.\ ●$$

7.11. Системы обыкновенных дифференциальных уравнений первого порядка

Система n уравнений первого порядка с n неизвестными функциями, разрешенная относительно производных, имеет вид

$$\begin{cases} \dfrac{dy_1}{dx}=f_1(x,\ y_1,\ y_2,\ \ldots,\ y_n), \\ \dfrac{dy_2}{dx}=f_2(x,\ y_1,\ y_2,\ \ldots,\ y_n), \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{dy_n}{dx}=f_n(x,\ y_1,\ y_2,\ \ldots,\ y_n), \end{cases} \tag{7.6}$$

где x — независимая переменная, $y_1,\ y_2,\ \ldots,\ y_n$ — неизвестные функции.

Решением системы (7.6) называется всякая система функций $y_1=\varphi_1(x),\ y_2=\varphi_2(x),\ \ldots,\ y_n=\varphi_n(x)$, определенных на конечном или бесконечном интервале изменения аргумента x, имеющих производные первого порядка и обращающих уравнения системы (7.6) в тождества по x.

Задача Коши для системы (7.6) заключается в определении ее решения $y_1=\varphi_1(x),\ y_2=\varphi_2(x),\ \ldots,\ y_n=\varphi_n(x)$, удовлетворяющего начальным условиям $y_1(x_0)=y_1^0$, $y_2(x_0)=y_2^0,\ \ldots,\ y_n(x_0)=y_n^0$, где $y_1^0,\ y_2^0,\ \ldots,\ y_n^0$ — заданные числа.

○ **Пример.** Решением системы

$$\begin{cases} \dfrac{dy_1}{dx}=-2y_1-3y_2, \\ \dfrac{dy_2}{dx}=-y_1 \end{cases}$$

с начальными условиями $y_1(0)=1$, $y_2(0)=-1$ являются функции $y_1=e^x$, $y_2=-e^x$. В самом деле,

$$\frac{dy_1}{dx}=e^x,\qquad \frac{dy_2}{dx}=-e^x.$$

Подставляя выражения y_1, y_2, $\frac{dy_1}{dx}$, $\frac{dy_2}{dx}$ в систему уравнений, получаем два тождественных равенства:

$$\begin{cases} e^x=-2e^x+3e^x, \\ -e^x=-e^x. \end{cases}$$ ●

7.12. Линейные дифференциальные уравнения

Линейным дифференциальным уравнением n-го *порядка* называется уравнение вида

$$y^{(n)}+a_1(x)\,y^{(n-1)}+\ldots+a_n(x)\,y=b(x), \qquad (7.7)$$

где $a_1(x), \ldots, a_n(x)$, $b(x)$ — известные функции от x, y — искомая функция.

Функции $a_1(x), \ldots, a_n(x)$ называются *коэффициентами дифференциального уравнения* (7.7).

Уравнение (7.7), в котором $b(x)\neq 0$, называется *неоднородным*.

Наряду с каждым неоднородным уравнением (7.7) можно рассмотреть соответствующее ему однородное уравнение

$$y^{(n)}+a_1(x)\,y^{(n-1)}+\ldots+a_n(x)\,y=0. \qquad (7.8)$$

Если $y_1=\varphi_1(x)$, $y_2=\varphi_2(x)$, $\ldots$, $y_k=\varphi_k(x)$ — решения однородного уравнения (7.8), то любая их линейная комбинация

$$C_1y_1+C_2y_2+\ldots+C_ky_k,$$

где $C_1, C_2, \ldots, C_k$ — постоянные, также решение этого однородного уравнения.

Система функций называется *линейно независимой на интервале* $]a, b[$, если ни одна из этих функций не может быть выражена в виде линейной комбинации остальных функций.

Фундаментальный набор решений — это набор линейно независимых решений уравнения (7.8), содержащий столько функций, каков порядок дифференциального уравнения.

Теорема. *Для того чтобы решения* $y_1=\varphi_1(x)$, $y_2=\varphi_2(x)$, $\ldots$, $y_n=\varphi_n(x)$ *линейного однородного дифференциального уравнения с непрерывными на отрезке* $[a, b]$ *коэффициентами были линейно независимыми на интервале* $]a, b[$, *необходимо и достаточно, чтобы определитель Вронского*

$$W(x)=\begin{vmatrix} \varphi_1(x) & \varphi_2(x) & \ldots & \varphi_n(x) \\ \varphi_1'(x) & \varphi_2'(x) & \ldots & \varphi_n'(x) \\ \ldots & \ldots & \ldots & \ldots \\ \varphi_1^{(n-1)}(x) & \varphi_2^{(n-1)}(x) & \ldots & \varphi_n^{(n-1)}(x) \end{vmatrix}$$

был отличен от нуля при любом x *из* $[a, b]$.

Любое решение однородного уравнения можно представить в виде линейной комбинации фундаментального набора решений

$$\bar{y}=\sum_{i=1}^{n} C_i y_i, \tag{7.9}$$

где C_i ($i=1, 2, \ldots, n$) — произвольные постоянные.

Выражение (7.9) называется *общим решением однородного дифференциального уравнения* (7.8).

○ **Пример.** Уравнение $y''+y=0$ имеет решения

$$y=\sin x,\ y=2\sin x,\ y=\cos x.$$

Легко убедиться, что первое и второе решения не образуют фундаментальную систему, а первое и третье — образуют. Следовательно, общее решение данного уравнения можно представить в виде

$$y=C_1\sin x+C_2\cos x,$$

где C_1, C_2 — произвольные постоянные. ●

Пусть $\hat{y}$ — некоторое решение неоднородного уравнения (7.7), а $\bar{y}$ — общее решение однородного уравнения (7.8). Тогда $y=\bar{y}+\hat{y}$ — *общее решение неоднородного уравнения* (7.7).

Знaя общее решение неоднородного уравнения (7.7), легко найти любое его частное решение.

7.13. Линейные дифференциальные уравнения с постоянными коэффициентами

Линейное дифференциальное уравнение n-го порядка с постоянными коэффициентами имеет вид

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_ny=b(x), \tag{7.10}$$

где a_1, a_2, $\ldots$, a_n — постоянные.

Уравнению (7.10) соответствует однородное уравнение

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_ny=0. \tag{7.11}$$

Уравнению (7.11) можно сопоставить многочлен относительно переменной λ

$$M(\lambda)=\lambda^n+a_1\lambda^{n-1}+\ldots+a_n, \tag{7.12}$$

называемый *характеристическим многочленом* уравнения (7.11), и соответствующее характеристическое уравнение

$$\lambda^n+a_1\lambda^{n-1}+\ldots+a_n=0. \tag{7.13}$$

○ **Пример.** Уравнению

$$y'' + 2y' - 5y = 0$$

соответствует характеристическое уравнение $\lambda^2+2\lambda-5=0$.●

Если λ_0 — корень характеристического уравнения (7.13), то $y = e^{\lambda_0 x}$ — решение однородного дифференциального уравнения (7.10).

Характеристический многочлен (7.12) можно представить (см. (7.2)) в виде $M(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\ldots(\lambda-\lambda_k)^{m_k}(\lambda^2+p_1\lambda+q_1)^{m_{k+1}}\ldots$ $\ldots(\lambda^2+p_l\lambda+q_l)^{m_{k+l}}$, где $\lambda_1 \neq \lambda_2 \neq \ldots \neq \lambda_k$ — различные действительные числа; $\lambda^2 + p_j\lambda + q_j$ — квадратные трехчлены, не имеющие действительных корней. В этом случае фундаментальный набор решений уравнения (7.11) можно построить следующим образом.

Каждому действительному корню λ_j кратности m_j $(j \leqslant k)$ сопоставляют набор линейно независимых решений $e^{\lambda_j x}$, $xe^{\lambda_j x}$, ..., $x^{m_j-1}e^{\lambda_j x}$. Для каждой пары комплексно-сопряженных корней $\lambda_j = \alpha_j \pm i\beta_j$ $(j > k)$ кратности m_j составляют набор решений:

$$e^{\alpha_j x}\cos\beta_j x,\ xe^{\alpha_j x}\cos\beta_j x,\ \ldots,\ x^{m_j-1}e^{\alpha_j x}\cos\beta_j x,$$
$$e^{\alpha_j x}\sin\beta_j x,\ xe^{\alpha_j x}\sin\beta_j x,\ \ldots,\ x^{m_j-1}e^{\alpha_j x}\sin\beta_j x.$$

○ **Пример.** Уравнению

$$y^{(4)} - 2y^{(3)} + 2y'' - 2y' + 1 = 0 \qquad (7.14)$$

соответствует характеристическое уравнение

$$\lambda^4 - 2\lambda^3 + 2\lambda^2 - 2\lambda + 1 = 0 \text{ или } (\lambda-1)^2(\lambda^2+1) = 0,$$

которое имеет корни $\lambda_{1,2} = 1$, $\lambda_{3,4} = \pm i$. Корню $\lambda = 1$ кратности 2 соответствуют решения $y_1 = e^x$, $y_2 = xe^x$. Корням $\lambda_{3,4} = \pm i$ соответствуют решения $y_3 = \cos x$, $y_4 = \sin x$.

Общее решение дифференциального уравнения (7.14) можно записать в виде

$$\bar{y} = C_1e^x + C_2xe^x + C_3\cos x + C_4\sin x,$$

где C_1, C_2, C_3, C_4 — произвольные постоянные. ●

Для определения общего решения неоднородного уравнения кроме фундаментального набора решений соответствующего однородного уравнения необходимо найти некоторое (частное) решение неоднородного уравнения. Обычно вид этого частного решения определяется формой правой части неоднородного уравнения.

○ **Пример.** Решить краевую задачу для дифференциального уравнения

$$y''-5y'+6y=e^x,\ y(0)=\frac{1}{2},\ y(1)=0.$$

Однородному уравнению $y''-5y'+6y=0$ соответствует характеристическое уравнение $\lambda^2-5\lambda+6=0$, корнями которого являются числа $\lambda_1=3$, $\lambda_2=2$. Им соответствуют решения $y_1=e^{3x}$, $y_2=e^{2x}$.

Общее решение однородного уравнения можно представить в виде $\bar{y}=C_1e^{3x}+C_2e^{2x}$.

Частное решение неоднородного уравнения ищем в форме

$$\hat{y}=Ae^x,$$

где A — постоянное число. Дифференцируя последнее выражение, находим

$$\hat{y}'=Ae^x,\ \hat{y}''=Ae^x.$$

Подставляя $\hat{y}$, $\hat{y}'$, $\hat{y}''$ в исходное уравнение $Ae^x-5Ae^x+{}$ $+6Ae^x=e^x$, находим $A=1/2$.

Таким образом, общее решение неоднородного уравнения имеет вид

$$y=\bar{y}+\hat{y}=C_1e^{3x}+C_2e^{2x}+\frac{1}{2}e^x.$$

Подставим в общее решение граничное условие $y(0)=1/2$. Получим $C_1=-C_2$. Подставляя в решение второе граничное условие и учитывая, что $C_1=-C_2$, имеем $C_1=$ $=-C_2=\frac{1}{2e(1-e)}$. Искомое решение принимает вид

$$y=\frac{1}{2e(1-e)}e^{3x}-\frac{1}{2e(1-e)}e^{2x}+\frac{1}{2}e^x. \quad ●$$

7.14. Разностные методы решения обыкновенных дифференциальных уравнений

Разностные методы решения дифференциальных уравнений — это способы вычисления значений искомого решения $y(x)$ на некоторой сетке значений аргумента.

Разностные методы позволяют находить только конкретное (частное) решение, например решение задачи Коши. Но эти методы в настоящее время являются основными при решении дифференциальных уравнений с помощью ЭВМ.

Одним из простейших разностных методов является *метод ломаных* или *метод Эйлера*.

Пусть требуется решить задачу Коши для уравнения первого порядка

$$y'=f(x, y),\ y(x_0)=y_0$$

на отрезке $[x_0, x_N]$.

На данном отрезке выбирают некоторую сетку значений аргумента $x_0, x_1, \ldots, x_N$, для которых вычисляют значения функции y по схеме

$$y_{n+1}=y_n+h_n f(x_n, y_n),\ h_n=x_{n+1}-x_n,$$

где $n=0, 1, \ldots, N-1$.

Этот метод дает хорошее приближение к решению только для достаточно малых h_n.

Модификации этого метода определяются следующими формулами:

$$y_{n+1}=y_n+h_n f\left[x_n+\frac{h_n}{2},\ y_n+\frac{f(x_n, y_n)}{2}h_n\right],$$

$$y_{n+1}=y_n+\frac{h_n}{6}\{f(x_n, y_n)+f[x_{n+1}, y_n+f(x_n, y_n)h_n]\}.$$

Более высокую точность обеспечивает *метод Рунге — Кутта*. Наиболее употребительной является следующая схема указанного метода:

$$y_{n+1}=y_n+\frac{h_n}{6}(k_1+2k_2+2k_3+k_4), \qquad (7.15)$$

где

$$k_1=f(x_n, y_n),\ k_2=f\left(x_n+\frac{h_n}{2},\ y_n+\frac{h_n}{2}k_1\right),$$

$$k_3=f\left(x_n+\frac{h_n}{2},\ y_n+\frac{h_n}{2}k_2\right),\ k_4=f(x_n+h_n, y_n+h_n k_3).$$

При решении конкретных задач используют также и другие разностные методы решения дифференциальных уравнений.

○ **Пример.** Решить задачу Коши методом Рунге — Кутта для дифференциального уравнения $y'=x^2+y^2$, $y(0)=1$ на отрезке $[0,\ 0{,}7]$.

Выберем шаг $h=0{,}1$. Используя формулы (7.15), получаем следующие значения функции y на сетке значений x:

x	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7
y	1	1,11	1,25	1,44	1,7	2,07	2,64	3,65

●

РАЗДЕЛ VIII. РЯДЫ

8.1. Сумма числового ряда

Выражение

$$a_1+a_2+\ldots+a_n+\ldots=\sum_{n=1}^{\infty} a_n,$$

где $a_1, a_2, \ldots, a_n, \ldots$ — некоторые числа, называют *числовым рядом*. Числа $a_1, a_2, \ldots, a_n, \ldots$ — *члены ряда*.

Для каждого числового ряда $\sum\limits_{n=1}^{\infty} a_n$ можно построить последовательность его частичных сумм S_n:

$$S_n=a_1+a_2+\ldots+a_n,\ n=1,\ 2,\ 3,\ \ldots.$$

○ **Пример.** Для ряда

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{n\,(n+1)}+\ldots=\sum_{n=1}^{\infty}\frac{1}{n\cdot(n+1)}$$

$$S_1=\frac{1}{1\cdot 2}=1-\frac{1}{2};\ S_2=\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}=1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}=$$
$$=1-\frac{1}{3};\ \ldots,\ S_n=\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{n\,(n+1)}=$$
$$=1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\ldots+\frac{1}{n}-\frac{1}{n+1}=1-\frac{1}{n+1};\ \ldots. \text{ ●}$$

Конечный предел S последовательности частичных сумм ряда $\sum\limits_{n=1}^{\infty} a_n$ называют *суммой* этого ряда.

○ **Пример.** $\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{n\,(n+1)}+\ldots=1$, так как

$$\lim_{n\to\infty} S_n=\lim_{n\to\infty}\left(1-\frac{1}{n+1}\right)=1. \text{ ●}$$

Если S — сумма ряда $\sum\limits_{n=1}^{\infty} a_n$, то число $r_n=S-S_n$ называют *остатком ряда*. Так как $\lim\limits_{n\to\infty} r_n=0$, то при достаточно большом n

$$S \approx S_n \ (S_n \text{— частичная сумма ряда}).$$

Числовой ряд называют *сходящимся*, если он имеет сумму, и *расходящимся* в противном случае.

○ **Примеры.**

1. Гармонический ряд $1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}+\ldots$ расходится.

2. Геометрическая прогрессия $a+aq+aq^2+\ldots+aq^{n-1}+\ldots$ $(a \neq 0)$, сходится при $|q|<1$ и расходится при $|q| \geqslant 1$. Если $|q|<1$, то $a+aq+aq^2+\ldots+aq^{n-1}+\ldots=\frac{a}{1-q}$.

3. Обобщенно гармонический ряд $\frac{1}{1^\alpha}+\frac{1}{2^\alpha}+\ldots+\frac{1}{n^\alpha}+\ldots$ сходится при $\alpha>1$ и расходится при $\alpha \leqslant 1$.

4. $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots+(-1)^{n-1}\frac{1}{n}+\ldots=\ln 2$.

5. $1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\ldots+(-1)^{n-1}\frac{1}{2n-1}+\ldots=\frac{\pi}{4}$.

6. $1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{n^2}+\ldots=\frac{\pi^2}{6}$.

7. $1-\frac{1}{2^2}+\frac{1}{3^2}-\frac{1}{4^2}+\ldots+(-1)^{n-1}\frac{1}{n^2}+\ldots=\frac{\pi^2}{12}$.

8. $\frac{1}{1!}+\frac{1}{2!}+\ldots+\frac{1}{n!}+\ldots=e-1$. ●

8.2. Основные свойства сходящихся числовых рядов

1°. Сходимость числового ряда не нарушится, если изменить конечное число его членов (сумма изменится).

2°. Если члены сходящегося ряда умножить на одно и то же число λ, то его сходимость не нарушится (сумма лишь умножится на λ).

3°. Два сходящихся ряда

$$a_1+a_2+\ldots+a_n+\ldots=A, \quad b_1+b_2+\ldots+b_n+\ldots=B$$

можно почленно складывать (или вычитать), так что ряд

$$(a_1 \pm b_1)+(a_2 \pm b_2)+\ldots+(a_n \pm b_n)+\ldots$$

также сходится и его сумма равна соответственно $A \pm B$.

4°. Если числовой ряд $a_1+a_2+\ldots+a_n+\ldots$ сходится, то $\lim\limits_{n\to\infty} a_n=0$ (*необходимый признак сходимости*).

Таким образом, если $\lim\limits_{n\to\infty} a_n \neq 0$ или не существует, то ряд $\sum\limits_{n=1}^{\infty} a_n$ расходится, однако условие $\lim\limits_{n\to\infty} a_n=0$ не обеспечивает сходимость этого ряда.

○ **Примеры.**

1. Ряд $\frac{1}{2}+\frac{2}{3}+\frac{3}{4}+\ldots+\frac{n}{n+1}+\ldots$ расходится, так как $\lim\limits_{n\to\infty} a_n=\lim\limits_{n\to\infty}\frac{n}{n+1}=\lim\limits_{n\to\infty}\frac{1}{1+1/n}=1\neq 0.$

2. Ряд $1-1+1-1+\ldots+(-1)^{n-1}+\ldots$ расходится, так как $\lim\limits_{n\to\infty}(-1)^{n-1}$ не существует. ●

8.3. Признаки сходимости положительных числовых рядов

Признаки сравнения. Если все члены рядов

$$a_1+a_2+\ldots+a_n+\ldots, \tag{1}$$

$$b_1+b_2+\ldots+b_n+\ldots \tag{2}$$

не отрицательны и $a_n \leqslant b_n$, $n=1, 2, 3\ldots$, то из сходимости ряда (2) следует сходимость ряда (1), а из расходимости ряда (1) следует расходимость ряда (2).

Если все члены рядов (1) и (2) положительны и существует $\lim\limits_{n\to\infty}\frac{a_n}{b_n}=k$, $0<k<+\infty$, то эти ряды сходятся или расходятся одновременно.

○ **Пример.** Ряд $\frac{1}{2}+\frac{1}{4}+\frac{1}{6}+\ldots+\frac{1}{2n}+\ldots$ расходится, так как гармонический ряд $1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}+$ расходится и

$$\lim_{n\to\infty}\frac{1/(2n)}{1/n}=\lim_{n\to\infty} n/(2n)=1/2. \text{ ●}$$

Признак Коши. Если все члены ряда

$$a_1+a_2+\ldots+a_n+\ldots$$

не отрицательны и существует $\lim\limits_{n\to\infty}\sqrt[n]{a_n}=q$, то при $q<1$ этот ряд сходится, а при $q>1$ расходится. (При $q=1$ признак Коши не дает возможности судить о поведении ряда.)

○ **Пример.** Ряд $\frac{2}{1}+\frac{2^2}{1^2}+\ldots+\frac{2^n}{n^n}+\ldots$ сходится, так как

$$\lim_{n\to\infty}\sqrt[n]{a_n}=\lim_{n\to\infty}\sqrt[n]{\frac{2^n}{n^n}}=\lim_{n\to\infty}\frac{2}{n}=0<1. \bullet$$

Признак Даламбера. Если все члены ряда

$$a_1+a_2+\ldots+a_n+\ldots$$

положительны и существует $\lim\limits_{n\to\infty}\frac{a_{n+1}}{a_n}=d$, то при $d<1$ этот ряд сходится, а при $d>1$ ряд расходится. (При $d=1$ признак Даламбера не дает возможности судить о поведении ряда.)

○ **Пример.** Ряд $\frac{3}{1!}+\frac{3^2}{2!}+\ldots+\frac{3^n}{n!}+\ldots$ сходится, так как

$$a_n=\frac{3^n}{n!},\ a_{n+1}=\frac{3^{n+1}}{(n+1)!} \text{ и } \lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\frac{3}{n+1}=0<1. \bullet$$

Интегральный признак сходимости Коши. Если $a_n=f(n)$, $n=1,\ 2,\ \ldots$, где $f(n)$ — значение при $x=n$ некоторой функции $f(x)$, непрерывной, положительной и невозрастающей при $x\geqslant 1$, то ряд $a_1+a_2+\ldots$ $\ldots+a_n+\ldots$. сходится или расходится в зависимости от того, существует или нет конечный $\lim\limits_{b\to\infty}\int\limits_1^b f(x)\,dx$.

○ **Пример.** Ряд $\frac{1}{1+1^2}+\frac{2}{1+2^2}+\ldots+\frac{n}{1+n^2}+\ldots$ расходится, так как функция $f(x)=\frac{x}{1+x^2}$ является положительной, непрерывной и невозрастающей при $x\geqslant 1$ и

$$\lim_{b\to+\infty}\int\limits_1^b f(x)\,dx=\lim_{b\to+\infty}\int\limits_1^b\frac{x\,dx}{1+x^2}=$$

$$=\frac{1}{2}\lim_{b\to+\infty}[\ln(1+b^2)-\ln 2]=+\infty. \bullet$$

8.4. Абсолютная и условная сходимость рядов

Числовой ряд

$$u_1+u_2+\ldots+u_n+\ldots \tag{8.1}$$

называется *абсолютно сходящимся*, если сходится ряд, составленный из абсолютных величин его членов:

$$|u_1|+|u_2|+\ldots+|u_n|+\ldots. \tag{8.2}$$

Абсолютно сходящийся ряд всегда сходится. Если сам ряд (8.1) сходится, а ряд (8.2) расходится, то говорят, что ряд (8.1) *сходится условно*.

Теорема Лейбница. *Ряд*

$$c_1 - c_2 + c_3 - c_4 + \ldots + (-1)^{n-1} c_n + \ldots,$$

где все $c_n > 0$, *сходится, если последовательность* $\{c_n\}$ *невозрастающая и* $\lim\limits_{n \to \infty} c_n = 0$.

В этом случае для остатка ряда справедлива оценка

$$|r_n| \leqslant c_{n+1}, \quad n = 1,\ 2,\ 3, \ldots .$$

Свойства абсолютно и условно сходящихся рядов

1°. Если ряд сходится абсолютно, то новый ряд, полученный из него перестановкой членов, также сходится и имеет ту же сумму, что и исходный ряд.

2°. Если ряд сходится условно, то, какое бы число B ни взять, можно так переставить члены в этом ряде, чтобы сумма преобразованного ряда была равна именно B.

3°. Если ряд сходится условно, то можно так переставить члены в этом ряде, что новый ряд будет расходиться.

8.5. Сходимость функциональных рядов

Выражение вида

$$f_1(x) + f_2(x) + \ldots + f_n(x) + \ldots = \sum_{n=1}^{\infty} f_n(x), \qquad (8.3)$$

где $f_1(x)$, $f_2(x)$, ..., $f_n(x)$, ... — некоторые функции, определенные на одном и том же множестве M, называется *функциональным рядом*.

Множество Ω $(\Omega \subseteq M)$ всех значений x, при которых функциональный ряд (8.3) сходится (как числовой ряд), называется *областью сходимости* этого ряда.

Функция $S(x)$, $x \in \Omega$ является *суммой ряда* (8.3), если

$$S(x) = \lim_{n \to \infty} S_n(x),$$

где $S_n(x) = f_1(x) + f_2(x) + \ldots + f_n(x)$.

Если функция $S(x)$, $x \in L$, $(L \subseteq \Omega)$ является суммой ряда (8.3), то говорят, что функциональный ряд (8.3) *сходится на множестве* L *к функции* $S(x)$.

Функциональный ряд называется *равномерно сходящимся* на множестве L к функции $S(x)$, если для любого числа $\varepsilon > 0$ существует номер N такой, что при $n \geqslant N$ сразу для всех $x \in L$ выполняется неравенство

$$|S(x) - S_n(x)| < \varepsilon.$$

Если функциональный ряд сходится на множестве L, то на этом множестве сходимость не обязана быть равномерной, однако на некотором подмножестве множества L сходимость может оказаться уже равномерной.

Признак равномерной сходимости Вейерштрасса. Если члены функционального ряда $f_1(x) + f_2(x) + \ldots + f_n(x) + \ldots$ удовлетворяют на множестве L неравенствам

$$|f_n(x)| \leqslant c_n \quad (n = 1, 2, \ldots),$$

где c_n — члены сходящегося числового ряда $c_1 + c_2 + \ldots + c_n + \ldots$, то функциональный ряд сходится на множестве L равномерно.

○ **Пример.** Ряд $\frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \ldots + \frac{\sin nx}{n^2} + \ldots$ сходится на $L =]-\infty; +\infty[$ равномерно, так как всегда $\left|\frac{\sin nx}{n^2}\right| \leqslant \frac{1}{n^2}$ и ряд $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ сходится. ●

8.6. Функциональные свойства суммы ряда

Если функции $f_n(x)$ непрерывны на $[a, b]$, а составленный из них ряд $f_1(x) + f_2(x) + \ldots + f_n(x) + \ldots$ сходится равномерно на этом отрезке к функции $f(x)$, то:

1°. Функция $f(x)$ на отрезке $[a, b]$ непрерывна.

2°. $\int\limits_a^b f(x)\,dx = \int\limits_a^b f_1(x)\,dx + \int\limits_a^b f_2(x)\,dx + \ldots + \int\limits_a^b f_n(x)\,dx + \ldots$.

○ **Пример.** Ряд $1 + x + x^2 + \ldots + x^{n-1} + \ldots$ на отрезке $[0, 1/2]$ сходится равномерно к функции $\frac{1}{1-x}$. Тогда

$$\int\limits_0^{1/2} 1 \cdot dx + \int\limits_0^{1/2} x\,dx + \ldots + \int\limits_0^{1/2} x^{n-1}\,dx + \ldots = \int\limits_0^{1/2} \frac{dx}{1-x}$$

или

$$1/2 + \frac{1}{2^2 \cdot 2} + \frac{1}{2^3 \cdot 3} + \ldots + \frac{1}{2^n \cdot n} + \ldots = \ln 2. \; ●$$

Если функции $f_n(x)$ имеют непрерывные производные на отрезке $[a, b]$ и на этом отрезке:

а) ряд $f_1(x)+f_2(x)+\ldots+f_n(x)+\ldots$ сходится к функции $f(x)$;

б) ряд $f_1'(x)+f_2'(x)+\ldots+f_n'(x)+\ldots$ сходится равномерно, то $f(x)$ имеет на $[a, b]$ непрерывную производную и

$$f'(x)=f_1'(x)+f_2'(x)+\ldots+f_n'(x)+\ldots .$$

8.7. Степенные ряды

Функциональный ряд

$$a_0+a_1(x-c)+a_2(x-c)^2+\ldots+a_n(x-c)^n+\ldots=$$
$$=\sum_{n=0}^{\infty} a_n(x-c)^n, \qquad (8.4)$$

где a_n $(n=0, 1, 2, \ldots)$ и c — некоторые числа, называют *степенным рядом с центром в точке c*.

Возможны лишь следующие три случая:

1) степенной ряд (8.4) сходится только при $x=c$ (всюду расходящийся ряд);

2) степенной ряд (8.4) сходится (при том абсолютно) при любом значении x (всюду сходящийся ряд);

3) существует число $R>0$ такое, что ряд (8.4) сходится абсолютно при $|x-c|<R$ и расходится при $|x-c|>R$ (R — радиус сходимости ряда).

Кроме того, считают: $R=0$ для всюду расходящегося ряда и $R=+\infty$ для всюду сходящегося ряда.

Интервал $]c-R, c+R[$, $R>0$, называют *интервалом сходимости* степенного ряда (8.4). При этом на концах интервала сходимости степенной ряд может как сходиться, так и расходиться.

● **Пример.** Найдем область сходимости степенного ряда

$$\frac{x}{1\cdot 2}+\frac{x^2}{2\cdot 2^2}+\ldots+\frac{x^n}{n\cdot 2^n}+\ldots .$$

Положим $u_n=\dfrac{|x|^n}{n\cdot 2^n}$, $u_{n+1}=\dfrac{|x|^{n+1}}{(n+1)2^{n+1}}$. Тогда

$$\lim_{n\to\infty}\frac{u_{n+1}}{u_n}=\lim_{n\to\infty}\frac{|x|^{n+1}\cdot n\cdot 2^n}{(n+1)2^{n+1}|x|^n}=\frac{|x|}{2}\lim_{n\to\infty}\frac{n}{n+1}=\frac{|x|}{2}.$$

По признаку Даламбера степенной ряд сходится абсолютно при $|x|<2$, а при $|x|>2$ абсолютной сходимости у него нет. Следовательно, радиус сходимости ряда $R=2$. Исследуем сходимость ряда на концах интервала сходимости: при $x=2$ ряд $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}+\ldots$ расходится, а при $x=-2$ ряд $-\frac{1}{1}+\frac{1}{2}-\frac{1}{3}+\ldots+(-1)^n\frac{1}{n}+\ldots$ сходится. Таким образом, область сходимости степенно́го ряда $\Omega = [-2, 2[$. ●

Основные свойства степенных рядов

1°. Если степенной ряд не является всюду расходящимся, то его сумма непрерывна в каждой точке области сходимости.

2°. Степенной ряд внутри его области сходимости можно интегрировать почленно, так что если

$$a_0+a_1(x-c)+a_2(x-c)^2+\ldots+a_n(x-c)^n+\ldots=f(x), \quad x\in\Omega,$$

то

$$a_0(x-c)+a_1\frac{(x-c)^2}{2}+\ldots+a_n\frac{(x-c)^{n+1}}{n+1}+\ldots=\int_c^x f(x)\,\mathrm{d}x.$$

3°. Степенной ряд внутри его интервала сходимости можно дифференцировать почленно, так что если

$$a_0+a_1(x-c)+a_2(x-c)^2+\ldots+a_n(x-c)^n+\ldots=f(x),$$
$$x\in]c-R,\ c+R[,\ R>0,$$

то

$$a_1+2a_2(x-c)+\ldots+na_n(x-c)^{n-1}+\ldots=f'(x),$$
$$x\in]c-R,\ c+R[.$$

Это утверждение сохраняет силу и для конца интервала сходимости, если только последний ряд на этом конце сходится.

4°. Если степенной ряд

$$a_0+a_1(x-c)+a_2(x-c)^2+\ldots+a_n(x-c)^n+\ldots$$

не является всюду расходящимся, то его сумма $f(x)$ имеет внутри интервала сходимости производные всех порядков. При этом

$$a_0=f(c),\ a_1=f'(c),\ a_2=\frac{f''(c)}{2!},\ \ldots,\ a_n=\frac{f^{(n)}(c)}{n!},\ \ldots .$$

8.8. Разложение функций в степенные ряды

Если функция $f(x)$ имеет производные всех порядков при $x=c$, то степенной ряд

$$f(c)+\frac{f'(c)}{1!}(x-c)+\frac{f''(c)}{2!}(x-c)^2+\ldots+\frac{f^n(c)}{n!}(x-c)^n+\ldots \tag{8.5}$$

называют *рядом Тейлора* для функции $f(x)$. При $c=0$ ряд (8.5) называется *рядом Маклорена.*

Для того чтобы ряд (8.5) сходился к функции $f(x)$, необходимо и достаточно, чтобы

$$\lim_{n\to\infty} R_n(x)=0,$$

где

$$R_n(x)=\frac{(x-c)^{n+1}}{(n+1)!}f^{(n+1)}[c+\theta\cdot(x-c)],\ 0<\theta<1.$$

Таблица разложений в степенной ряд некоторых функций

$$1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}+\ldots=e^x,\ -\infty<x<+\infty;$$

$$\frac{x}{1!}-\frac{x^3}{3!}+\frac{x^5}{5!}-\ldots(-1)^n\frac{x^{2n+1}}{(2n+1)!}+\ldots=\sin x,$$
$$-\infty<x<+\infty;$$

$$1-\frac{x^2}{2!}+\frac{x^4}{4!}-\ldots(-1)^n\frac{x^{2n}}{(2n)!}+\ldots=\cos x,$$
$$-\infty<x<+\infty;$$

$$x-\frac{x^3}{3}+\frac{x^5}{5}-\ldots(-1)^n\frac{x^{2n+1}}{2n+1}+\ldots=\operatorname{arctg} x,\ |x|\leqslant 1;$$

$$x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots+(-1)^{n-1}\frac{x^n}{n}+\ldots=\ln(1+x),$$
$$-1<x\leqslant 1;$$

$$1+mx+\frac{m(m-1)}{2!}x^2+\ldots+\frac{m(m-1)\ldots(m-n+1)}{n!}x^n+\ldots=$$
$$=(1+x)^m,\ |x|<1.$$

8.9. Тригонометрические ряды

Функциональный ряд

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx), \tag{8.6}$$

где a_0, a_n, b_n, $(n=1, 2, \ldots)$ — некоторые числа, называется *тригонометрическим рядом*.

Свойства тригонометрических рядов

1°. Сумма тригонометрического ряда (8.6) является функцией периодической, ее период $T=2\pi$.

2°. Если числовой ряд $\sum_{n=1}^{\infty} c_n$ сходится абсолютно, то тригонометрические ряды $\frac{a_0}{2}+\sum_{n=1}^{\infty} c_n \cos nx$ и $\sum_{n=1}^{\infty} c_n \sin nx$ сходятся равномерно на всей числовой оси.

3°. Если тригонометрический ряд (8.6) сходится равномерно на отрезке $[-\pi, \pi]$ к функции $f(x)$, то

$$a_0=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx, \quad a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx \quad (n=1, 2, 3, \ldots).$$

8.10. Ряды Фурье

Пусть $f(x)$ — некоторая периодическая функция периода $T=2\pi$. Тригонометрический ряд

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

называется *рядом Фурье для функции* $f(x)$, если

$$a_0=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx, \quad a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx,$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx \quad (n=1, 2, 3, \ldots).$$

Свойства рядов Фурье

1°. Если функция $f(x)$ — четная, то

$$b_n=0 \ (n=1, 2, 3, \ldots), \quad a_n=\frac{2}{\pi}\int_{0}^{\pi} f(x)\cos nx\,dx$$

$$(n=0, 1, 2, \ldots).$$

2°. Если функция $f(x)$ — нечетная, то

$$a_n = 0 \ (n = 0, 1, 2, 3, \ldots), \quad b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx \, dx$$

$$(n = 1, 2, \ldots).$$

3°. В точке x_0, где функция $f(x)$ дифференцируема или по крайней мере имеет конечные односторонние пределы

$$\lim_{t \to +0} \frac{f(x_0 + t) - f(x_0)}{t} \quad \text{и} \quad \lim_{t \to -0} \frac{f(x_0 + t) - f(x_0)}{t},$$

ряд Фурье для функции $f(x)$ сходится, причем его сумма равна $f(x_0)$. В противном случае сумма этого ряда Фурье может быть отличной от $f(x_0)$.

4°. Если функция $f(x)$ периодическая периода $T = 2l$, то ее ряд Фурье имеет вид

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{l} + b_n \sin \frac{n\pi x}{l} \right),$$

где $\quad a_0 = \frac{1}{l} \int_{-l}^{l} f(x) \, dx, \quad a_n = \frac{1}{l} \int_{-l}^{l} f(x) \cos \frac{n\pi x}{l} \, dx,$

$$b_n = \frac{1}{l} \int_{-l}^{l} f(x) \sin \frac{n\pi x}{l} \, dx \ (n = 1, 2, 3, \ldots).$$

5°. Если функция $f(x)$ задана на конечном интервале $]-l, l[$, то для того, чтобы ее разложить в ряд Фурье, необходимо предварительно построить функцию $\varphi(x)$ периода $T = 2l$ такую, что

$$\varphi(x) = f(x) \ \text{при} \ -l < x < l.$$

Если функцию $\varphi(x)$ можно разложить в ряд Фурье, то на интервале $]-l, l[$ этот ряд будет рядом Фурье для функции $f(x)$.

○ **Пример.** Разложить в ряд Фурье периодическую функцию $f(x)$ периода $T = 2l$, если

$$f(x) = \begin{cases} x, & 0 \leqslant x \leqslant l, \\ -x, & -l \leqslant x < 0. \end{cases}$$

Так как функция $f(x)$ является четной, то $b_n=0$ $(n=1,\ 2,\ 3,\dots)$,

$$a_0=\frac{2}{l}\int_0^l f(x)\,dx=\frac{2}{l}\int_0^l x\,dx=\frac{2}{l}\cdot\frac{x^2}{2}\Big|_0^l=l;$$

$$a_n=\frac{2}{l}\int_0^l f(x)\cos\frac{n\pi x}{l}\,dx=\frac{2}{l}\int_0^l x\cos\frac{n\pi x}{l}\,dx=$$

$$=\frac{2}{l}\,\frac{l}{n\pi}\int_0^l x\,d\sin\frac{n\pi x}{l}=\frac{2}{n\pi}\left[x\sin\frac{n\pi x}{l}\Big|_0^l-\int_0^l\sin\frac{n\pi x}{l}\,dx\right]=$$

$$=\frac{2}{n\pi}\cdot\frac{l}{n\pi}\cos\frac{n\pi x}{l}\Big|_0^l=\frac{2l}{(n\pi)^2}[(-1)^n-1].$$

Отсюда

$$a_n=\begin{cases}0, & \text{если } n\text{ — четное;}\\ -4l/(n\pi)^2, & \text{если } n\text{ — нечетное.}\end{cases}$$

Таким образом,

$$f(x)=\frac{l}{2}-\frac{4l}{\pi^2}\left[\frac{\cos\frac{\pi x}{l}}{1^2}+\frac{\cos\frac{3\pi x}{l}}{3^2}+\frac{\cos\frac{5\pi x}{l}}{5^2}+\dots\right].\ \bullet$$

8.11. Приложения рядов

1. С помощью рядов можно приближенно вычислять различные постоянные величины.

Для того чтобы вычислить некоторую постоянную S, необходимо ее представить в виде суммы числового ряда $S=a_1+a_2+\dots+a_n+\dots$ и положить

$$S\approx S_n=a_1+a_2+\dots+a_n.$$

При достаточно большом n остаток $r_n=S-S_n$ станет сколь угодно малым, так что S_n воспроизведет S с любой наперед заданной точностью.

○ **Пример.** Вычислить $1/\sqrt{e}$ с точностью 0,005.

Так как

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\dots,\quad -\infty<x<+\infty,$$

то

$$\frac{1}{\sqrt{e}}=e^{-\frac{1}{2}}=1-\frac{1}{2}+\frac{1}{2^2\cdot 2!}-\frac{1}{2^3\cdot 3!}+\frac{1}{2^4\cdot 4!}-\dots.$$

Полученный числовой ряд удовлетворяет условиям теоремы Лейбница. Поэтому для остатка справедлива оценка $|r_n| < a_{n+1}$. Значит, $|r_3| < \frac{1}{2^4 \cdot 4!} < 0{,}003$. Поэтому с нужной точностью имеем

$$\frac{1}{\sqrt{e}} = 1 - \frac{1}{2} + \frac{1}{2^2 \cdot 2!} - \frac{1}{2^3 \cdot 3!} = 0{,}604. \quad \bullet$$

2. Ряды применяют для приближенных вычислений определенных интегралов. Для вычисления $\int\limits_0^t f(x)\,dx$ необходимо подынтегральную функцию $f(x)$ разложить в степенной ряд. Если

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots, \qquad -R < x < R,$$

то при $-R < t < R$ степенной ряд можно интегрировать почленно. Получим

$$\int\limits_0^t f(x)\,dx = a_0 t + a_1 \frac{t^2}{2} + a_2 \frac{t^3}{3} + \ldots + a_n \frac{t^{n+1}}{n+1} + \ldots,$$

откуда $\int\limits_0^t f(x)\,dx$ можно вычислить с любой наперед заданной точностью.

○ **Пример.** Рассмотрим интеграл вероятностей

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-t}^{t} e^{-\frac{x^2}{2}}\,dx = \frac{2}{\sqrt{2\pi}} \int\limits_0^t e^{-\frac{x^2}{2}}\,dx.$$

Так как

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \quad -\infty < x < +\infty,$$

то

$$e^{-x^2/2} = 1 - \frac{x^2}{2} + \frac{x^4}{2^2 \cdot 2!} - \frac{x^6}{2^3 \cdot 3!} + \ldots.$$

Тогда

$$\Phi(t) = \frac{2}{\sqrt{2\pi}} \int\limits_0^t e^{-\frac{x^2}{2}}\,dx =$$

$$= \frac{2}{\sqrt{2\pi}} \left[t - \frac{t^3}{3 \cdot 2} + \frac{t^5}{5 \cdot 2^2 \cdot 2!} - \frac{t^7}{7 \cdot 2^3 \cdot 3!} + \ldots \right]. \quad \bullet$$

3. С помощью рядов можно интегрировать некоторые дифференциальные уравнения. Если, например, необхо-

димо найти решение дифференциального уравнения $y' = f(x, y)$ такое, что $y(x_0) = y_0$, то это решение можно искать в виде ряда Тейлора

$$y(x) = y(x_0) + y'(x_0)(x - x_0) + \frac{y''(x_0)}{2!}(x - x_0)^2 + \ldots,$$

где $y(x_0) = y_0$, $y'(x_0) = f(x_0, y_0)$, а дальнейшие производные $y^{(n)}(x_0)$ последовательно находят с помощью дифференцирования уравнения $y' = f(x, y)$ и цодстановки вместо x числа x_0.

4. Ряды Фурье позволяют выделить периодические (сезонные) колебания, свойственные многим экономическим явлениям. Для изучения периодических колебаний некоторого экономического показателя $f(t)$, зависящего от времени, функцию $f(t)$ раскладывают в ряд Фурье $\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt)$ (для практических целей достаточно рассмотреть лишь несколько первых членов этого ряда).

Чаще всего сама функция $f(t)$ неизвестна, известен лишь конечный набор ее значений: $f(t_1)$, $f(t_2), \ldots$. В этом случае приходится вычислять коэффициенты ряда Фурье приближенно. Для приближенных вычислений коэффициентов ряда Фурье существует большое количество различных методов.

РАЗДЕЛ IX. МЕТОДЫ ОПТИМИЗАЦИИ

9.1. Оптимизационные задачи

Пусть $f(x)$ — функция, определенная на множестве V, а Ω — некоторое подмножество множества V.

Оптимизационная задача задается тройкой (V, f, Ω). При этом функция $f(x)$ называется *целевой функцией*, а Ω — *допустимым множеством* (*множеством допустимых решений*) оптимизационной задачи.

Оптимизационные задачи бывают двух типов: задачи минимизации и задачи максимизации. *Задача минимизации* (*максимизации*) (V, f, Ω) состоит в отыскании наименьшего (наибольшего) значения целевой функции $f(x)$ на допустимом множестве Ω.

Для того чтобы решить задачу минимизации (максимизации) (V, f, Ω), достаточно найти ее оптимальное решение, т. е. указать $x_0 \in \Omega$ такое, что $f(x_0) \leqslant f(x)$ $(f(x_0) \geqslant f(x))$ при любом $x \in \Omega$.

Оптимизационная задача называется *неразрешимой*, если она не имеет оптимального решения. В частности, задача минимизации (максимизации) (V, f, Ω) будет неразрешимой, если целевая функция $f(x)$ не ограничена снизу (сверху) на допустимом множестве Ω.

Решить оптимизационную задачу — значит либо найти ее оптимальное решение, либо установить неразрешимость этой задачи.

Любая задача максимизации (V, f, Ω) сводится к задаче минимизации $(V, -f, \Omega)$: эти задачи либо обе неразрешимы, либо имеют одно и то же оптимальное решение.

Две задачи минимизации (максимизации) называются *эквивалентными*, если они имеют одно и то же множество допустимых решений и на любом допустимом решении значения целевых функций этих задач совпадают. Эквивалентные оптимизационные задачи либо обе неразрешимы, либо имеют одно и то же оптимальное решение. В частности, задачи (V, f, Ω) и (V, f', Ω) будут эквивалентными, если при любом $x \in \Omega$ справедливо равенство

$$f'(x) = \lambda f(x) + \mu,$$

где λ, μ — некоторые числа и $\lambda > 0$.

Методы решения оптимизационных задач зависят как от вида целевой функции, так и от строения допустимого множества Ω.

Методы решения оптимизационных задач, в которых целевая функция является функцией n переменных, часто называют *методами математического программирования*. (Термин «программирование» в данном случае обусловлен тем, что в задачах ищется некоторая программа действий.) В математическом программировании традиционно выделяют следующие основные разделы: линейное программирование, целочисленное программирование, выпуклое программирование.

Методы решения оптимизационных задач, в которых целевая функция представляет собой функционал на некотором множестве функций или вектор-функций, рассматриваются в вариационном исчислении и в теории оптимального управления.

9.2. Задачи линейного программирования

Оптимизационная задача (V, f, Ω), в которой целевая функция является линейной функцией на $V = \mathbf{R}^n$, а Ω — множеством решений некоторой системы линейных уравнений и линейных неравенств от n неизвестных, называется *задачей линейного программирования*. При этом система линейных уравнений и линейных неравенств, определяющая допустимое множество Ω, называется *системой ограничений* задачи линейного программирования.

Задача линейного программирования будет поставлена, если:

а) указана линейная целевая функция $f = \sum_{j=1}^{n} \gamma_j x_j$;

б) записана система ограничений

$$\sum_{j=1}^{n} a_{ij} x_j = b_i,\ i \in I,\ I \subseteq M = \{1, 2, \ldots, m\},$$

$$\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i,\ i \in M \setminus I,$$

$$x_j \geqslant 0,\ j \in J,\ J \subseteq N = \{1, 2, \ldots, n\}$$

(часто бывает полезно в системе ограничений особо выделить неравенста вида $x_j \geqslant 0$);

в) определено, к какому типу (максимизации или минимизации) принадлежит данная задача. (Любую задачу максимизации можно свести к задаче минимизации, поменяв знаки у коэффициентов целевой функции.)

Любую задачу линейного программирования можно записать в следующем виде:

$$f = \sum_{j=1}^{n} \gamma_j x_j \ (\min), \tag{9.1}$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \ i \in I, \ I \subseteq M = \{1, 2, \ldots, m\}, \tag{9.2}$$

$$\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i, \ i \in M \setminus I, \tag{9.3}$$

$$\bar{x}_j \geqslant 0, \ j \in J, \ I \subseteq N = \{1, 2, \ldots, n\}. \tag{9.4}$$

В частности, если $I = \varnothing$ $(I = M)$, то система (9.2)—(9.3) состоит только из линейных неравенств (уравнений).

Матрица A размера $m \times n$, у которой на месте (i, j) стоит коэффициент при x_j в i-м ограничении системы (9.2)—(9.3), т. е.

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

называется *матрицей условий*, а столбцы этой матрицы

$$A_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{pmatrix}, \ \ldots, \quad A_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{pmatrix}$$

— *векторами* условий задачи (9.1)—(9.4). Вектор B называется *вектором ограничений* задачи (9.1)—(9.4)

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{pmatrix}$$

Свойства задач линейного программирования

1°. Допустимое множество задачи линейного программирования либо пусто, либо является выпуклым подмножеством пространства $\mathbf{R}^n$.

2°. Если допустимое множество задачи линейного программирования не пусто, а целевая функция ограничена снизу (для задачи минимизации) на этом множестве, то задача линейного программирования имеет оптимальное решение.

3°. Оптимальные решения задачи линейного программирования (если они существуют) всегда находятся на границе допустимого множества и образуют выпуклое подмножество пространства $\mathbf{R}^n$.

Рассмотрим примеры экономических задач, приводимых к задаче линейного программирования.

Задача о рационе. В хозяйстве имеется n видов кормов, каждый из которых содержит m разновидностей питательных веществ. Известно, что одна единица j-го вида кормов ($j=1, 2, \ldots, n$) содержит a_{ij} единиц i-го питательного вещества ($i=1, 2, \ldots, m$) и имеет стоимость c_j. Требуется составить такой рацион, который бы удовлетворил всем потребностям b_i ($i=1, 2, \ldots, m$) в питательных веществах и имел бы наименьшую стоимость.

Обозначим количество j-го корма в рационе x_j, $x_j \geqslant 0$ ($j=1, 2, \ldots, n$), а f—стоимость этого рациона. Тогда $f = \sum_{j=1}^{n} c_j x_j$.

Рацион должен удовлетворять всем потребностям в питательных веществах. Значит,

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

Таким образом, мы получим задачу линейного программирования:

$$f=\sum_{j=1}^{n} c_j x_j \,(\min),$$

$$\sum_{j=1}^{n} a_{ij}x_j \geqslant b_i, \quad i=1,\ 2,\ \ldots,\ m,$$

$$x_j \geqslant 0, \quad j=1,\ 2,\ \ldots,\ n.$$

Простейшая задача планирования производства. Предприятие располагает m видами ресурсов и может выпускать некоторую однородную продукцию n различными технологическими способами. За единицу времени использования j-го технологического способа ($j=1,\ 2,\ \ldots,\ n$) расходуется a_{ij} единиц i-го ресурса ($i=1,\ 2,\ \ldots,\ m$) и выпускается c_j единиц продукции. Требуется спланировать работу предприятия так, чтобы оно выпускало наибольшее количество продукции при условии, что ресурса i-го вида нельзя израсходовать более чем b_i единиц ($i=1,\ 2,\ \ldots,\ m$).

Пусть x_j, $x_j \geqslant 0$ ($j=1,\ 2,\ \ldots,\ n$) — время использования предприятием j-го *технологического* процесса. Если при этом f — количество выпущенной продукции, то

$$f=\sum_{j=1}^{n} c_j x_j.$$

Работая по этому плану, предприятие израсходует $\sum\limits_{j=1}^{n} a_{ij}x_j$ единиц i-го ресурса. Значит,

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \quad i=1,\ 2,\ \ldots,\ m,$$

и мы снова имеем задачу линейного программирования:

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

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \quad i=1,\ 2,\ \ldots,\ m,$$

$$x_j \geqslant 0, \quad j=1,\ 2,\ \ldots,\ n.$$

Важным примером задачи линейного программирования является транспортная задача (см. п. 9.12).

9.3. Графический метод решения двумерных задач линейного программирования

Дана задача линейного программирования

$$f=\gamma_1 x_1+\gamma_2 x_2 \ (\max); \tag{9.5}$$

$$\begin{cases} a_{11}x_1+a_{12}x_2 \leqslant b_1, \\ a_{21}x_1+a_{22}x_2 \leqslant b_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{m1}x_1+a_{m2}x_2 \leqslant b_m, \end{cases} \tag{9.6}$$

в которую входят только два неизвестных: x_1 и x_2.

Каждое из неравенств системы (9.6) определяет на координатной плоскости (x_1, x_2) некоторую полуплоскость. Следовательно, допустимое множество Ω задачи (9.5)—(9.6) есть пересечение конечного числа полуплоскостей, т. е. некоторая многоугольная область на плоскости (x_1, x_2).

Для решения задачи (9.5)—(9.6) графическим методом прежде всего необходимо построить многоугольную область Ω, а затем перпендикулярно вектору $\Gamma=(\gamma_1, \gamma_2)$ провести прямую l так, чтобы она пересекала область Ω.

Прямая l перемещается параллельно самой себе в направлении вектора Γ до тех пор, пока она не перестанет пересекать область Ω (для задачи минимизации прямую l необходимо перемещать в противоположном направлении).

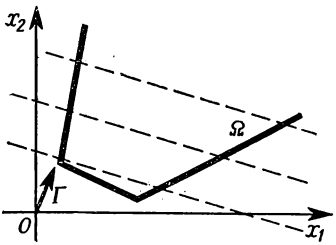


Рис. 9.1

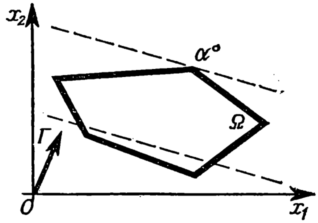


Рис. 9.2

Если при таком перемещении прямая l все время будет пересекать область Ω, то целевая функция не ограничена сверху на допустимом множестве и задача (9.5)—(9.6) не имеет оптимального решения (рис. 9.1).

В противном случае пересечение области Ω с прямой l в том ее положении, когда дальнейшее перемещение дает пустое пересечение с Ω, является множеством оптимальных решений задачи (9.5)—(9.6) (рис. 9.2).

○ **Пример.** Предприятие располагает тремя видами сырья и может выпускать одну и ту же продукцию двумя способами. При этом за 1 ч работы первым способом выпускается 20 единиц продукции, а вторым способом — 30 единиц продукции. Количество сырья (кг) того или иного вида, расходуемого за 1 ч при различных способах производства, и запасы сырья (кг) приведены в таблице. Требуется найти план производства, при котором будет выпущено наибольшее количество продукции.

Способ производства	Сырье		
	1	2	3
Первый	10	20	15
Второй	20	10	15
Запасы сырья	100	100	90

Обозначим через x_1 и x_2 время (ч) использования соответственно первого и второго способов производства. Имеем задачу линейного программирования $f = 20x_1 + 30x_2\,(\max)$;

$$\begin{cases} 10x_1 + 20x_2 \leqslant 100, \\ 20x_1 + 10x_2 \leqslant 100, \\ 15x_1 + 15x_2 \leqslant 90, \\ \quad x_1 \geqslant 0,\ x_2 \geqslant 0, \end{cases}$$

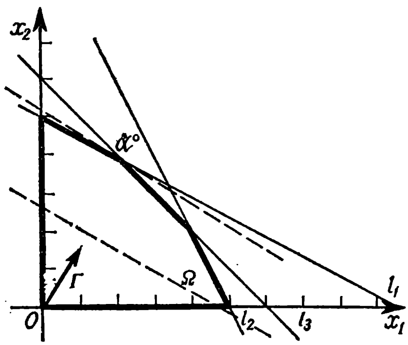


Рис. 9.3

которую можно решать графическим способом. На рис. 9.3 изображены допустимое множество Ω и оптимальное решение α^0 этой задачи.

Любая точка из допустимого множества Ω является планом работы предприятия, для реализации которого хватит имеющихся запасов сырья. Оптимальное решение α^0 — это план из допустимого множества Ω, при котором будет выпущено наибольшее количество продукции.

Очевидно, что α^0 является точкой пересечения прямых l_1 и l_3, имеющих уравнения $10x_1 + 20x_2 = 100$ и $15x_1 + 15x_2 = 90$ соответственно.

Решая систему этих двух уравнений, получаем $x_1=2$, $x_2=4$.

Таким образом, для производства наибольшего количества продукции при имеющихся запасах сырья необходимо 2 ч применять первый способ производства и 4 ч— второй способ.

При этом будет изготовлено $f(\alpha^0)=20\cdot 2+30\times 4=160$ единиц продукции. ●

9.4. Каноническая форма задачи линейного программирования

Задача линейного программирования имеет каноническую форму, если в ее систему ограничений входят только линейные уравнения и условия неотрицательности для всех неизвестных, т. е. в задаче (9.1)—(9.4) $I=M=\{1, 2, \ldots, m\}$, а $J=N=\{1, 2, \ldots, n\}$.

Задача линейного программирования в канонической форме имеет следующий вид:

$$f=\sum_{j=1}^{n}\gamma_j x_j\ (\min), \tag{9.7}$$

$$\sum_{j=1}^{n}a_j x_j=b_i,\ i=1,\ 2,\ \ldots,\ m, \tag{9.8}$$

$$x_j\geqslant 0,\ j=1,\ 2,\ \ldots,\ n. \tag{9.9}$$

Задачу линейного программирования (9.7)—(9.9) в канонической форме можно записать в виде

$$f=\sum_{j=1}^{n}\gamma_j x_j\ (\min),$$

$$\sum_{j=1}^{n}A_j x_j=B,\ x_j\geqslant 0,\ j=1,\ 2\ \ldots,\ n,$$

где $A_1, A_2, \ldots, A_j, \ldots, A_n$—векторы условий, а B—вектор ограничений этой задачи.

Любую задачу линейного программирования можно свести к задаче линейного программирования в канонической форме. В самом деле, рассмотрим задачу (9.1)—(9.4).

Каждое неизвестное x_j, $j\in N\setminus J$, заменим на разность y_j-z_j двух новых неотрицательных неизвестных y_j, z_j.

После этого в неравенства системы (9.3) введем по новому неизвестному $t_i\ (i\in M\setminus I)$, полагая

$$\sum_{j\in J}a_{ij}x_j+\sum_{j\in N\setminus J}a_{ij}(y_j-z_j)+t_i=b_i,\quad i\in M\setminus I.$$

Получим задачу линейного программирования в канонической форме:

$$f' = \sum_{j \in J} \gamma_j x_j + \sum_{j \in N \setminus J} \gamma_j (y_j - z_j) \quad (\min), \tag{9.10}$$

$$\sum_{j \in J} a_{ij} x_j + \sum_{j \in N \setminus J} a_{ij} (y_j - z_j) = b_i, \ i \in I, \tag{9.11}$$

$$\sum_{j \in J} a_{ij} x_j + \sum_{j \in N \setminus J} a_{ij} (y_j - z_j) + t_i = b_i, \ i \in M \setminus I, \tag{9.12}$$

$$x_j \geqslant 0, \ j \in J; \ y_j, \ z_j \geqslant 0, \ j \in N \setminus J; \ t_i \geqslant 0, \ i \in M \setminus I. \tag{9.13}$$

При этом задача (9.10)—(9.13) будет иметь оптимальное решение тогда и только тогда, когда оптимальное решение было и у исходной задачи (9.1)—(9.4).

Кроме того, зная координаты оптимального решения задачи (9.10)—(9.13) x_j^0, $j \in J$, y_j^0, z_j^0, $j \in N \setminus J$, t_i^0, $i \in M \setminus I$, легко найти оптимальное решение исходной задачи, полагая $x_j = x_j^0$ при $j \in J$ и $x_j = y_j^0 - z_j^0$ при $j \in N \setminus J$.

9.5. Опорные решения задачи линейного программирования в канонической форме

Допустимое решение α задачи (9.7)—(9.9) в канонической форме называется *опорным решением* этой задачи, если векторы условий $A_{i_1} A_{i_2}, \ldots, A_{i_k}$, где $i_1, i_2, \ldots, i_k$ — номера всех ненулевых координат α, образуют линейно независимую систему векторов.

○ Например, векторы $\alpha_1 = (0; 0; 1/2; 5/2)$ и $\alpha_2 = (1; 0; 1; 1)$ являются допустимыми решениями задачи:

$$f = 3x_1 + 4x_2 - x_3 + x_4 \quad (\min),$$

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

$$x_j \geqslant 0, \ j = 1, \ 2, \ 3, \ 4.$$

Векторы условий $A_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $A_4 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ образуют, очевидно, линейно независимую систему. Значит, α_1 является опорным решением данной задачи. Векторы $A_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $A_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $A_4 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ линейно зависимы. Поэтому α_2 не является опорным решением.

Свойства опорных решений

1°. Если допустимое множество задачи (9.7)—(9.9) в канонической форме не пусто, то эта задача имеет опорное решение.

2°. Опорные решения задачи (9.7)—(9.9) являются крайними точками допустимого множества этой задачи. (Допустимое множество всегда выпукло.)

3°. Задача (9.7)—(9.9) в канонической форме имеет лишь конечное число различных опорных решений (либо не имеет их вовсе).

Чтобы найти некоторое опорное решение задачи (9.7)—(9.9), достаточно выбрать базис системы $A_1, A_2, \ldots, A_n$ векторов условий этой задачи так, чтобы вектор ограничений B раскладывался по нему с неотрицательными коэффициентами.

Если $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ — такой базис и $B = A_{i_1}d_{i_1} + A_{i_2}d_{i_2} + \ldots + A_{i_r}d_{i_r}$, $d_{i_1} \geqslant 0$, $d_{i_2} \geqslant 0$, $\ldots$, $d_{i_r} \geqslant 0$, то $\alpha = (0; \ldots; 0; d_{i_1}; 0; \ldots; 0; d_{i_2}; 0; \ldots; 0; d_{i_r}; 0; \ldots; 0)$ является опорным решением задачи (9.7)—(9.9).

Базис $A_{i_1}, A_{i_2}, \ldots, A_{i_r}$ системы векторов условий $A_1, A_2, \ldots, A_n$ задачи (9.7)—(9.9) называется *базисом опорного решения* $\alpha = (d_1; d_2; \ldots; d_n)$ этой задачи, если $d_i = 0$ при $i \neq i_1, i_2, \ldots, i_r$.

○ Рассмотрим, например, опорное решение $\alpha = (1; 0; 0; 1)$ задачи

$$f = x_1 + x_2 - x_3 + x_4 \ (\min),$$

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

$$x_j \geqslant 0, \ j = 1, 2, 3, 4.$$

Здесь $A_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $A_3 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $A_4 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

и A_1, A_4; A_1, A_2 — базисы системы. Так как вторая и третья координаты вектора α равны 0, то A_1, A_4 является базисом опорного решения α. С другой стороны, четвертая координата α отлична от нуля. Следовательно, A_1, A_2 не будет базисом α. ●

У любого опорного решения задачи (9.7)—(9.9) не может быть более чем r ненулевых (положительных) координат, где $r = r(A_1, A_2, \ldots, A_n)$ ($r(A_1, A_2, \ldots, A_n)$ — ранг системы векторов условий). Опорное решение называется *невырожденным*, если число его ненулевых координат точно равно r и *вырожденным* — в противном случае.

Любое опорное решение имеет базис. При этом у невырожденного опорного решения базис только один, а вырожденное опорное решение может иметь несколько различных базисов.

Опорные решения играют важную роль при решении задачи линейного программирования в канонической форме, так как если эта задача имеет оптимальное решение, то одно из ее оптимальных решений обязательно будет ее опорным решением. Таким образом, оптимальное решение задачи линейного программирования в канонической форме можно искать только среди ее опорных решений (а их лишь конечное число).

9.6. Признак оптимальности опорного решения задачи линейного программирования в канонической форме. Условие неограниченности целевой функции

По задаче линейного программирования в канонической форме (9.7)—(9.9) всегда можно составить таблицу (табл. 9.1).

Таблица 9.1

x_1	x_2	...	...	x_n	
a_{11}	a_{12}	...	...	a_{1n}	b_1
a_{21}	a_{22}	...	...	a_{2n}	b_2
.	.			.	.
.	.			.	.
.	.			.	.
a_{m1}	a_{m2}			a_{mn}	b_m
$-\gamma_1$	$-\gamma_2$	...	...	$-\gamma_n$	0

Таблица (9.1) называется симплекс-таблицей для задачи (9.7)—(9.9),

Предположим, что $\alpha=(d_1;\ d_2;\ \ldots;\ d_n)$ — некоторое опорное решение задачи (9.7)—(9.9), а векторы A_{i_1}, A_{i_2}, ..., A_{i_r} образуют его базис. Тогда таблицу (9.1) можно преобразовать методом Гаусса в табл. 9.2.

Прибавим к последней строке таблицы (9.2) первую строку, умноженную на γ_{i_1}, вторую строку, умноженную на γ_{i_2}, r-ю строку, умноженную на γ_{i_r}. В результате получим новую табл. 9.3, где d_{i_1}, d_{i_2}, ..., d_{i_r} — координаты опорного решения, соответствующие векторам базиса A_{i_1}, A_{i_2}, ..., A_{i_r}.

Таблица 9.2

x_1	...	x_{i_1}	...	x_{i_2}	...	x_{i_r}	...	...	...	x_s	...	x_n	
a'_{11}		1		0		0				a'_{1s}		a'_{1n}	d_{i_1}
a'_{21}		0		1		0				a'_{2s}		a'_{2n}	d_{i_2}
.		.		.		.				.		.	.
.		.		.		.				.		.	.
.		.		.		.				.		.	.
a'_{r1}		0		0		1				a'_{rs}		a'_{rn}	d_{i_r}
$-\gamma_1$	...	$-\gamma_{i_1}$	...	$-\gamma_{i_2}$	...	$-\gamma_{i_r}$	...	...	...	$-\gamma_s$		$-\gamma_n$	0

Таблица 9.3

x_1	...	x_{i_1}	...	x_{i_2}	...	x_{i_r}	...	...	...	x_s	...	x_n	
a'_{11}		1		0		0				a_{1s}		a'_{1n}	d_{i_1}
a'_{21}		0		1		0				a'_{2s}		a'_{2n}	d_{i_2}
.		.		.		.				.		.	.
.		.		.		.				.		.	.
.		.		.		.				.		.	.
a'_{r1}		0		0		1				a'_{r}		a'_{rn}	d_{i_r}
δ_1	...	δ_{i_1}	...	δ_{i_2}	...	δ_{i_r}	...	...	...	δ_s	...	δ_n	δ_0

где $\delta_{i_1} = \delta_{i_2} = \ldots = \delta_{i_r} = 0$, а $\delta_0 = f(\alpha)$ (значение целевой функции на опорном решении α).

Таблица (9.3), полученная указанным выше способом, называется *симплекс-таблицей, приведенной к базису* $A_{i_1}, A_{i_2}, \ldots, A_{i_r}$ опорного решения α, а числа $\delta_1, \ldots$ $\ldots, \delta_{i_1}, \ldots)), \delta_{i_2}, \ldots, \delta_{i_r}, \ldots, \delta_s, \ldots, \delta_n$ — оценками этого базиса.

Имеют место следующие утверждения (случай задачи минимизации):

1. Если все оценки некоторого базиса опорного решения не положительны, то оно является оптимальным решением задачи линейного программирования в канонической форме. (Признак оптимальности.)

2. Для оптимального опорного решения задачи линейного программирования в канонической форме всегда существует базис, все оценки которого не положительны.

3. Предположим, что симплекс-таблица для задачи линейного программирования в канонической форме при-

ведена к базису некоторого опорного решения. Если среди оценок этого базиса имеется положительная оценка δ_s, а все остальные элементы s-го столбца таблицы не положительны, то целевая функция не ограничена снизу на допустимом множестве. (Условие неограниченности целевой функции.)

○ **Пример.** Рассмотрим опорное решение $\alpha = (1;\ 0;\ 0;\ 0)$ задачи:

$$f = -10x_1 + x_2 - 9x_3 + 6x_1 \ (\min),$$
$$\begin{cases} x_1 + x_2 + 3x_3 = 1, \\ x_1 - x_2 - x_3 + 2x_4 = 1, \end{cases}$$
$$x_j \geqslant 0, \ j = 1, 2, 3, 4.$$

Приведем симплекс-таблицу для этой задачи к базису A_1, A_2 опорного решения α:

x_1	x_2	x_3	x_4	
1	1	3	0	1
1	−1	−1	2	1
10	−1	9	−6	0

$\longrightarrow$

x_1	x_2	x_3	x_4	
1	1	3	0	1
0	−2	−4	2	0
10	−1	9	−6	0

$\longrightarrow$

x_1	x_2	x_3	x_4	
1	0	1	1	1
0	1	2	−1	0
10	−1	9	−6	0

$\longrightarrow$

x_1	x_2	x_3	x_4	
1	0	1	1	1
0	1	2	−1	0
0	0	1	−17	−10

Среди оценок базиса A_1, A_2 есть положительная оценка $\delta_3 = 1$. Значит, утверждать на этом этапе, что α — оптимальное решение, мы не можем.

С другой стороны, возьмем базис A_1, A_3 того же опорного решения α. Приведя симплекс-таблицу к этому базису, получим

x_1	x_2	x_3	x_4	
1	−1/2	0	3/2	1
0	1/2	1	−1/2	0
0	−1/2	0	−33/2	−10

Все оценки базиса A_1, A_3 не положительны. Следовательно, α — оптимальное решение данной задачи. ●

9.7. Решение задачи линейного программирования симплекс-методом

Если известно некоторое опорное решение задачи линейного программирования в канонической форме, то ее можно решать симплекс-методом. Симплекс-метод — это направленный перебор опорных решений, при котором значение целевой функции на каждом последующем опорном решении меньше, чем на предыдущем (для задачи минимизации).

Для решения симплекс-методом задачи линейного программирования в канонической форме необходимо выполнить ряд последовательных шагов. На каждом шаге либо возникает базис нового опорного решения, причем на новом опорном решении значение целевой функции обязательно меньше, чем на предыдущем (для задачи минимизации), либо меняется базис исходного опорного решения.

Если на очередном шаге удается установить оптимальность опорного решения или неограниченность целевой функции на допустимом множестве, то следующий шаг не выполняется.

Предположим, что перед выполнением очередного шага симплекс-таблица для задачи (9.7)—(9.9) приведена к базису

$$A_{i_1}, \ \ldots, \ A_{i_{k-1}}, \ A_{i_k}, \ A_{i_{k+1}}, \ \ldots, \ A_{i_r} \tag{9.14}$$

некоторого опорного решения α. Можно считать, что среди оценок этого базиса есть положительные оценки (в противном случае α — оптимальное решение).

Выберем любую из положительных оценок, например $\delta_s > 0$. В симплекс-таблице, приведенной к базису (9.14), рассмотрим элементы

$$a'_{1s}, \ \ldots, \ a'_{ls}, \ \ldots, \ a'_{ks}, \ \ldots, \ a'_{rs} \tag{9.15}$$

s-го столбца и соответствующие свободные члены

$$\mathrm{d}_1, \ \ldots, \ \mathrm{d}_l, \ \ldots, \ \mathrm{d}_k, \ \ldots, \ \mathrm{d}_r. \tag{9.16}$$

Можно считать, что среди элементов (9.15) есть положительные элементы (в противном случае целевая функция не ограничена снизу на допустимом множестве).

В этом случае необходимо рассмотреть все отношения вида

$$\frac{\mathrm{d}_l}{a'_{ls}}, \quad \text{где} \quad a'_{ls} > 0, \tag{9.17}$$

т. е. отношения свободного члена к соответствующему элементу s-го столбца при условии, что последний положителен.

Среди отношений (9.17) выбираем наименьшее. Если

$$\theta = \frac{d_k}{d'_{ks}} = \min\left\{\frac{d_l}{a'_{ls}}\,\middle|\, a'_{ls} > 0\right\},$$

то симплекс-таблицу приводим к новому базису

$$A_{i_1}, \ldots, A_{i_{k-1}}, A_s, A_{i_{k+1}}, \ldots, A_{i_r}, \tag{9.18}$$

т. е. из базиса (9.14) исключаем вектор A_{i_k}, а вместо него вводим вектор A_s.

Если $\theta > 0$ (так заведомо случится, если исходное опорное решение не вырождено), то базис (9.18) является базисом нового опорного решения β такого, что $f(\beta) < f(\alpha)$.

Если же $\theta = 0$, то (9.18) является новым базисом исходного опорного решения α.

Конечность симплекс-метода. Если задача линейного программирования в канонической форме не имеет вырожденных опорных решений, то через конечное число шагов симплекс-метода она будет решена.

При решении задачи, у которой есть вырожденные опорные решения, может случиться, что, перебирая базисы одного и того же опорного решения, через несколько шагов мы вернемся к исходному базису, т. е. произойдет зацикливание.

Для предотвращения зацикливания необходимо уточнить правило перехода к новому базису. Это сделать можно, например, следующим образом.

Среди оценок базиса (9.14) выбираем положительную оценку с наименьшим номером. Если δ_s — такая оценка, а наименьшее из отношений (9.17) определено неоднозначно, берем наименьшее отношение $\frac{d_k}{a'_{ks}}$ с наименьшим номером k.

○ **Пример.** Для изготовления продукции используют три вида сырья. При этом можно применять любой из четырех способов производства. Запасы сырья, расход и количества производимой продукции за 1 ч работы по каждому способу приведены в табл. 9.4.

Требуется найти план производства, при котором будет выпущено наибольшее количество продукции.

Обозначаем через x_j время использования j-го способа производства ($j = 1, 2, 3, 4$), получаем задачу линейного

Таблица 9.4

Сырье \ Способ производства	1	2	3	4	Запас сырья
1	1	2	1	0	18
2	1	1	2	1	30
3	1	3	3	2	40
Выпуск продукции	12	7	18	10	

программирования:

$$f = 12x_1 + 7x_2 + 18x_3 + 10x_4 \ (\max),$$

$$\begin{cases} x_1 + 2x_2 + x_3 \leqslant 18, \\ x_1 + x_2 + 2x_3 + x_4 \leqslant 30, \\ x_1 + 3x_2 + 3x_3 + 2x_4 \leqslant 40, \end{cases}$$

$$x_j \geqslant 0, \quad j = 1, 2, 3, 4.$$

Эту задачу сведем к канонической задаче минимизации:

$$f' = -12x_1 - 7x_2 - 18x_3 - 10x_4 \ (\min),$$

$$\begin{cases} x_1 + 2x_2 + x_3 + x_5 = 18, \\ x_1 + x_2 + 2x_3 + x_4 + x_6 = 30, \\ x_1 + 3x_2 + 3x_3 + 2x_4 + x_7 = 40, \end{cases}$$

$$x_j \geqslant 0, \quad j = 1, 2, 3, 4, 5, 6, 7,$$

и составим симплекс-таблицу. Симплекс-таблица оказывается приведенной к базису A_5, A_6, A_7 опорного решения $\alpha_1 = (0;\ 0;\ 0;\ 0;\ 18,\ 30,\ 40)$.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	
1	2	1	0	1	0	0	18
1	1	2	1	0	1	0	30
1	3	3	2	0	0	1	40
12	7	18	10	0	0	0	0

Выбираем положительную оценку $\delta_1 = 12$ и составляем следующие отношения: 18/1, 30/1, 40/1. Так как наименьшее среди них 18/1, то необходимо перейти к базису A_1, A_6, A_7. Для этого достаточно выполнить жорданово преобразование всей таблицы с ведущим элементом $a_{11}=1$. Имеем

x_1	x_2	x_3	x_4	x_5	x_6	x_7	
1	2	1	0	1	0	0	18
0	—1	1	1	—1	1	0	12
0	1	2	2	—1	0	1	22
0	—17	6	10	—12	0	0	—216

Базис A_1, A_6, A_7 является базисом опорного решения $\alpha_2=(18;\ 0;\ 0;\ 0;\ 0;\ 12;\ 22)$. При этом $f'(\alpha_2)=-216$, в то время как $f'(\alpha_1)=0$.

Выбираем оценку $\delta_4=10>0$ и составляем отношения: 12/1, 22/2. Наименьшим среди них является 22/2. Следовательно, переходим к базису A_1, A_4, A_6. Получим

x_1	x_2	x_3	x_4	x_5	x_6	x_7	
1	2	1	0	1	0	0	18
0	—3/2	0	0	—1/2	1	—1/2	1
0	1/2	1	1	—1/2	0	1/2	11
0	—22	—4	0	—7	0	—5	—326

Все оценки базиса A_1, A_4, A_6 не положительны. Следовательно, $\alpha_3=(18;\ 0,\ 0,\ 11,\ 0;\ 1;\ 0)$ — оптимальное решение канонической задачи минимизации. Поэтому $\beta=(18;\ 0;\ 0;\ 11)$ — оптимальное решение исходной задачи. При этом $f(\beta)=326$.

Таким образом, для того чтобы выпустить наибольшее количество продукции при имеющихся запасах сырья, необходимо в течение 18 ч использовать первый способ производства и в течение 11 ч — четвертый. В результате будет произведено 326 единиц продукции. ●

9.8. Метод искусственного базиса для отыскания начального опорного решения

Для того чтобы решить задачу линейного программирования в канонической форме, необходимо предварительно найти некоторое начальное опорное решение этой задачи. Сделать это можно методом искусственного базиса.

Рассмотрим задачу линейного программирования в канонической форме:

$$f=\sum_{j=1}^{n}\gamma_j x_j \ (\min), \tag{9.19}$$

$$\sum_{j=1}^{n}a_{ij}x_j=b_i, \ i=1, 2, \ldots, m, \tag{9.20}$$

$$x_j\geqslant 0, \ j=1, 2, \ldots, n. \tag{9.21}$$

Без ограничения общности можно считать, что $b_i\geqslant 0$, $i=1, 2, \ldots, m$. Для отыскания опорного решения задачи (9.19)—(9.21) строим вспомогательную задачу:

$$\varphi=\sum_{i=1}^{m}y_i \ (\min), \tag{9.22}$$

$$\sum_{j=1}^{n}a_{ij}x_j+y_i=b_i, \ i=1, 2, \ldots, m, \tag{9.23}$$

$$x_j\geqslant 0, \ j=1, 2, \ldots, n; \ \ y_i\geqslant 0, \ \ i=1, 2. \ldots, m. \tag{9.24}$$

Свойства вспомогательной задачи.

1°. Вспомогательная задача (9.22)—(9.24) всегда имеет оптимальное решение.

2°. Вектор $\beta=(0; 0; \ldots; 0; b_1; b_2; \ldots; b_m)$ является опорным решением задачи (9.22)—(9.24).

Таким образом, принимая вектор β за начальное опорное решение, вспомогательную задачу можно решить симплекс-методом. Пусть $\beta^0=(x_1^0; x_2^0; \ldots; x_n^0; y_1^0; y_2^0; \ldots; y_m^0)$ — оптимальное опорное решение задачи (9.22)—(9.24).

Если $y_1^0=y_2^0=\ldots=y_m^0=0$, то $\alpha=(x_1^0; x_2^0; \ldots; x_n^0)$ — опорное решение исходной задачи (9.19)—(9.21). Если же среди чисел $y_1^0, y_2^0, \ldots, y_m^0$ есть положительные, то исходная задача (9.19)—(9.21) имеет пустое допустимое множество.

Таким образом, всегда можно либо найти опорное решение исходной задачи, либо установить ее неразрешимость.

9.9. Взаимно двойственные задачи линейного программирования

Взаимно двойственные задачи линейного программирования имеют следующий вид:

$$f=\sum_{j=1}^{n}\gamma_j x_j \ (\max), \qquad (9.25)$$

$$\sum_{j=1}^{n} a_{ij}x_j=b_i,\ i\in I,\ I\subseteq M=\{1,2,\ldots,m\}, \qquad (9.26)$$

$$\sum_{j=1}^{n} a_{ij}x_j\leqslant b_i,\ i\in M\setminus I, \qquad (9.27)$$

$$x_j\geqslant 0,\ j\in J,\ J\subseteq N=\{1,2,\ldots,n\}, \qquad (9.28)$$

$$\varphi=\sum_{i=1}^{m} b_i y_i \ (\min), \qquad (9.29)$$

$$\sum_{i=1}^{m} a_{ij}y_i=\gamma_j,\ j\in N\setminus J, \qquad (9.30)$$

$$\sum_{i=1}^{m} a_{ij}y_i\geqslant \gamma_j,\ j\in J, \qquad (9.31)$$

$$y_i\geqslant 0,\ i\in M\setminus I. \qquad (9.32)$$

При этом задача максимизации (9.25)—(9.28) называется прямой (или исходной) задачей, а задача минимизации (9.29)—(9.32)—двойственной к ней.

Во взаимно двойственных задачах всегда:

1) одна из задач является задачей максимизации, а другая—задачей минимизации, в системе ограничений задачи максимизации неравенства записаны со знаком $\leqslant$, а в системе ограничений задачи минимизации—со знаком $\geqslant$;

2) каждому ограничению одной задачи соответствует переменная другой задачи: номер переменной совпадает с номером ограничения, при этом ограничению, записанному в виде неравенства, соответствует переменная, связанная условием неотрицательности;

3) матрица условий одной задачи получается из матрицы условий другой задачи с помощью транспонирования;

4) коэффициенты целевой функции одной задачи соответственно равны свободным членам системы ограничений другой задачи.

○ Взаимно двойственными являются следующие задачи:

$$f=5x_1-x_2+8x_3-x_4\ (\max),$$
$$\begin{cases} 2x_1+5x_2-x_3+7x_4=2,\\ x_1-x_2+5x_3-x_4\leqslant 3,\\ x_1-x_2+3x_3+7x_4=5,\end{cases}$$
$$x_1\geqslant 0,\ x_3\geqslant 0;$$

$$\varphi=2y_1+3y_2+5y_3\ (\min),$$
$$\begin{cases} 2y_1+y_2+y_3\geqslant 5,\\ 5y_1-y_2-y_3=-1,\\ -y_1+5y_2+3y_3\geqslant 8,\\ 7y_1-y_2+7y_3=-1,\end{cases}$$
$$y_2\geqslant 0. \ ●$$

Наиболее часто встречаются следующие частные случаи взаимно двойственных задач:

1. Если $I=M=\{1, 2, \ldots, m\}$, а $J=N=\{1, 2, \ldots, n\}$, то имеем несимметричную пару взаимно двойственных задач:

$$f=\sum_{j=1}^{n}\gamma_j x_j \text{ (max)}, \qquad \varphi=\sum_{i=1}^{m}b_i y_i \text{ (min)},$$

$$\sum_{j=1}^{n}a_{ij}x_j=b_i,\ i=1, 2, \ldots, m, \qquad \sum_{i=1}^{m}a_{ij}y_i\geqslant\gamma_j,\ j=1, 2, \ldots, n.$$

$$x_j\geqslant 0,\ j=1, 2, \ldots, n;$$

2. Если $I=\varnothing$, а $J=N=\{1, 2, \ldots, n\}$, то получим симметричную пару взаимно двойственных задач:

$$f=\sum_{j=1}^{n}\gamma_j x_j \text{ (min)}, \qquad \varphi=\sum_{i=1}^{m}b_i y_i \text{ (min)},$$

$$\sum_{j=1}^{n}a_{ij}x_j\leqslant b_i,\ i=1, 2, \ldots, m, \qquad \sum_{i=1}^{m}a_{ij}y_i\geqslant\gamma_j,\ j=1, 2, \ldots, n,$$

$$x_j\geqslant 0,\ j=1, 2, \ldots, n; \qquad y_i\geqslant 0,\ i=1, 2, \ldots, m.$$

Простейшие свойства взаимно двойственных задач

1^0. Если α — допустимое решение прямой задачи (9.25)—(9.28), а β — допустимое решение двойственной задачи (9.29)—(9.32), то $f(\alpha)\leqslant\varphi(\beta)$.

2^0. Если α и β — допустимые решения соответственно прямой и двойственной задач и $f(\alpha)=\varphi(\beta)$, то α и β — оптимальные решения этих задач.

9.10. Теоремы двойственности в линейном программировании

Основные утверждения о взаимно двойственных задачах содержатся в двух следующих теоремах.

Первая теорема двойственности. Для взаимно двойственных задач (9.25)—(9.28) и (9.29)—(9.32) имеет место один из взаимоисключающих случаев.

1. В прямой и двойственной задачах имеются оптимальные решения, причем значения целевых функций на оптимальных решениях совпадают.

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

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

4. Обе из рассматриваемых задач имеют пустые допустимые множества.

Вторая теорема двойственности. Пусть $\alpha^0=(x_1^0; x_2^0; \ldots; x_j^0; \ldots; x_n^0)$ — допустимое решение прямой задачи (9.25) — (9.28), а $\beta^0=(y_1^0; y_2^0; \ldots; y_i^0; \ldots; y_m^0)$ — допустимое решение двойственной задачи (9.29) — (9.32). Для того чтобы α^0 и β^0 были оптимальными решениями соответственно задач (9.25) — (9.28) и (9.29) — (9.32), необходимо и достаточно, чтобы выполнялись следующие соотношения:

$$x_j^0\left(\sum_{i=1}^{m} a_{ij}y_i^0-\gamma_j\right)=0,\quad j=1, 2, \ldots, n, \tag{9.33}$$

$$y_i^0\left(\sum_{j=1}^{n} a_{ij}x_j^0-b_i\right)=0,\quad i=1, 2, \ldots, m. \tag{9.34}$$

Условия (9.33) — (9.34) позволяют, зная оптимальное решение одной из взаимно двойственных задач, найти оптимальное решение другой задачи. Поэтому для решения некоторой задачи линейного программирования можно вначале решить двойственную ей задачу, а затем определить оптимальное решение исходной задачи. Существуют и другие методы решения задач линейного программирования, опирающиеся на теорию двойственности.

На основании второй теоремы двойственности можно сформулировать критерий оптимальности для допустимого решения задачи линейного программирования.

Критерий оптимальности. *Пусть* $\alpha^0=(x_1^0; \ldots; x_j^0; \ldots; x_n^0)$ — *допустимое решение задачи* (9.25) — (9.28). *Вектор* α^0 *является оптимальным решением этой задачи тогда и только тогда, когда среди решений системы уравнений*

$$\sum_{i=1}^{m} a_{ij}y_i-\gamma_j=0 \text{ при } x_j^0\neq 0, \tag{9.35}$$

$$y_i=0 \qquad \text{при } \sum_{j=1}^{n} a_{ij}x_j^0<b_i \tag{9.36}$$

содержится хотя бы одно допустимое решение задачи (9.29) — (9.32), *двойственной к задаче* (9.25) — (9.28).

○ Вектор $\alpha=(3; 0; 1; 3)$ является допустимым решением задачи

$$f=-2x_1-x_2+x_3+x_4 \text{ (max)},$$
$$\begin{cases} x_1-x_2+2x_3-\ \ x_4=2,\\ \qquad 3x_2-7x_3+3x_4=2,\\ \qquad\ \ x_2+2x_3 \qquad\ =2,\end{cases}$$
$$x_j \geqslant 0,\ j=1,\ 2,\ 3,\ 4.$$

В данном случае соотношения вида (9.36) отсутствуют, а из условий (9.35) имеем

$$\begin{cases} y_1=-2,\\ 2y_1-7y_2+2y_3=1,\\ -y_1+3y_2=1.\end{cases}$$

Отсюда $y_1=-2$, $y_2=-1/3$, $y_3=4/3$. Нетрудно проверить, что $\beta=(-2;\ -1/3;\ 4/3)$ — решение системы ограничений двойственной задачи:

$$\begin{cases} \qquad y_1 \qquad\qquad\quad \geqslant -2,\\ -y_1+3y_2+\ \ y_3 \geqslant -1,\\ \ \ 2y_1+7y_2+2y_3 \geqslant 1,\\ -y_1+3y_2 \qquad\quad \geqslant 1.\end{cases}$$

Следовательно, $\alpha=(3;\ 0;\ 1;\ 3)$ — оптимальное решение нашей задачи (из второй теоремы двойственности сразу следует, что $\beta=(-2;\ -1/3;\ 4/3)$ — оптимальное решение задачи, двойственной к данной). ●

9.11. Экономическая интерпретация двойственности в линейном программировании

Как правило, если исходная задача имеет определенный содержательный смысл, то и двойственная ей задача имеет естественную интерпретацию. При этом теоремы двойственности также получают содержательную трактовку.

Рассмотрим, например, следующую ситуацию. Имеется m видов ресурсов в количестве b_1, b_2, ..., b_m единиц соответственно. Известно, что на основе имеющихся ресурсов можно выпускать продукцию n различными способами. При этом за единицу времени использования j-го способа ($j=1, 2, \ldots, n$) расходуется a_{ij} единиц i-го ресурса ($i=1, 2, \ldots, m$) и выпускается продукция, обладающая ценностью c_j единиц. Как же оценить имеющиеся ресурсы в зависимости от возможностей нашего производства?

Производственную программу в данном случае можно охарактеризовать вектором $\alpha=(x_1; x_2; \ldots; x_n)$, где x_j — время использования j-го способа производства, причем

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \quad i=1, 2, \ldots, m,$$
$$x_j \geqslant 0, \quad j=1, 2, \ldots, n.$$

Пусть $f(\alpha)$ — ценность продукции, выпущенной при использовании производственной программы α. Тогда

$$f(\alpha)=\sum_{j=1}^{n} c_j x_j.$$

Если y_i — некоторая оценка единицы ресурса i-го вида ($i=1, 2, \ldots, m$), то величина $\sum_{i=1}^{m} a_{ij}y_i$ будет оценкой всех ресурсов, расходуемых в единицу времени при использовании j-го способа производства. Эта величина не может быть меньше ценности выпущенной при тех же условиях продукции. (Иначе часть ценности выпускаемой продукции возникает из «ничего».) Следовательно, имеем

$$\sum_{i=1}^{m} a_{ij}y_i \geqslant c_j, \quad j=1, 2, \ldots, n,$$
$$y_i \geqslant 0, \quad i=1, 2, \ldots, m.$$

Величина

$$\varphi(\beta)=\sum_{i=1}^{m} b_i y_i$$

является оценкой всех имеющихся ресурсов при векторе оценок

$$\beta=(y_1; \ldots; y_i; \ldots; y_m).$$

Для любой производственной программы α и при любом векторе оценок β выполняется неравенство

$$f(\alpha) \leqslant \varphi(\beta),$$

т. е. ценность всей выпущенной продукции не превосходит суммарной оценки имеющихся ресурсов (см. свойство 1° п. 9.9).

Значит, величина

$$\Delta(\alpha, \beta)=\varphi(\beta)-f(\alpha)$$

характеризует производственные потери в зависимости от рассматриваемой производственной программы и от выбранных оценок ресурсов.

Производственную программу и вектор оценок следует выбирать так, чтобы величина потерь была наименьшей. Для этого достаточно производственную программу α подобрать так, чтобы $f(\alpha)$ было как можно больше, а вектор оценок β взять таким, чтобы $\varphi(\beta)$ было как можно меньше. Получаем симметричную пару взаимно двойственных задач:

$$f=\sum_{j=1}^{n} c_j x_j \,(\max), \qquad \varphi=\sum_{i=1}^{m} b_i y_i \,(\min),$$

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i,\ i=1, 2, \ldots, m, \qquad \sum_{i=1}^{m} a_{ij}y_i \geqslant c_j,\ j=1, 2,\ldots, n,$$

$$x_j \geqslant 0,\ j=1, 2,\ldots, n; \qquad y_i \geqslant 0,\ i=1, 2, \ldots, m.$$

Из первой теоремы двойственности следует, что при оптимальной производственной программе и при оптимальном векторе оценок ресурсов производственные потери равны 0.

Из второй теоремы двойственности в данном случае следуют такие требования на оптимальную производственную программу $\alpha^0=(x_1^0; \ldots, x_j^0, \ldots, x_n^0)$ и оптимальный вектор оценок $\beta^0=(y_1^0, \ldots, y_i^0, \ldots, y_m^0)$:

$$\left\{\begin{array}{l} \text{если } y_i^0>0, \text{ то } \sum\limits_{j=1}^{n} a_{ij}x_j^0=b_i,\ i=1, 2, \ldots, m, \\ \text{если } \sum\limits_{j=1}^{n} a_{ij}x_j^0<b_i, \text{ то } y_i^0=0,\ i=1, 2, \ldots, m; \end{array}\right. \tag{9.37}$$

$$\left\{\begin{array}{l} \text{если } x_j^0>0, \text{ то } \sum\limits_{i=1}^{m} a_{ij}y_i^0=c_j,\ j=1, 2, \ldots, n, \\ \text{если } \sum\limits_{i=1}^{m} a_{ij}y_i^0>c_j, \text{ то } x_j^0=0,\ j=1, 2, \ldots, n. \end{array}\right. \tag{9.38}$$

Условия (9.37) можно интерпретировать так: если оценка y_i единицы ресурса i-го вида положительна, то при оптимальной производственной программе этот ресурс используется полностью, если же ресурс используется не полностью, то его оценка равна 0.

Из условий (9.38) следует, что если j-й способ используется в производстве, то он в оптимальных оценках неубыточен, если же j-й способ убыточен, то он в производстве не используется.

9.12. Транспортная задача

Формулировка транспортной задачи. Имеется m пунктов $\Pi_1, \Pi_2, \ldots, \Pi_m$, в которых производится некоторый однородный продукт соответственно в количествах $a_1, a_2, \ldots, a_m$ единиц. Этот продукт необходимо доставить в n пунктов потребления $Q_1, Q_2, \ldots, Q_n$, потребности которых в продукте соответственно составляют $b_1, b_2, \ldots, b_n$ единиц. Стоимость перевозки из каждого пункта производства Π_i $(i=1, 2, \ldots, m)$ в каждый пункт потребления Q_j $(j=1, 2, \ldots, n)$ известна и равна c_{ij} единиц. Требуется найти план перевозок, при котором были бы удовлетворены все потребности, а суммарная стоимость всех перевозок была бы наименьшей.

Очевидно, что если $\sum\limits_{i=1}^{m} a_i < \sum\limits_{j=1}^{n} b_j$, то транспортная задача неразрешима.

Можно считать, что $\sum\limits_{i=1}^{m} a_i = \sum\limits_{j=1}^{n} b_j$ $\left(\text{если } \sum\limits_{i=1}^{m} a_i > \sum\limits_{j=1}^{n} b_j\right.$, то вводят дополнительный (фиктивный) пункт потребления с потребностью, равной $\left.\sum\limits_{i=1}^{m} a_i - \sum\limits_{j=1}^{n} b_j \text{ единиц}\right)$.

Обозначим через x_{ij} количество продукта, перевозимого из пункта Π_i $(i=1, 2, \ldots, m)$ в пункт Q_j $(j=1, 2, \ldots, n)$. Если f — стоимость перевозок по этому плану, то

$$f = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}.$$

При этом из пункта Π_i $(i=1, 2, \ldots, m)$ будет вывезено всего $\sum\limits_{j=1}^{n} x_{ij}$ единиц продукта, а в пункт Q_j $(j=1, 2, \ldots, n)$ завезено $\sum\limits_{i=1}^{m} x_{ij}$ единиц продукта.

Значит,

$$\sum_{j=1}^{n} x_{ij} = a_i, \ i=1, 2, \ldots, m; \quad \sum_{i=1}^{m} x_{ij} = b_j, \ j=1, 2, \ldots, n.$$

Таким образом, транспортная задача является задачей линейного программирования в канонической форме:

$$f=f\sum_{i=1}^{m}\sum_{j=1}^{n}c_{ij}x_{ij}\ (\min), \tag{9.39}$$

$$\sum_{j=1}^{n}x_{ij}=a_i,\ i=1, 2, \ldots, m, \tag{9.40}$$

$$\sum_{i=1}^{m}x_{ij}=b_j,\ j=1, 2, \ldots, n, \tag{9.41}$$

$$x_{ij}\geqslant 0,\ i=1, 2, \ldots, m;\ j=1, 2, \ldots, n. \tag{9.42}$$

Свойства транспортной задачи.

1°. Транспортная задача (9.39)—(9.42) имеет оптимальное решение тогда и только тогда, когда $\sum_{i=1}^{m}a_i=\sum_{j=1}^{n}b_j$.

2°. Если все числа $a_1, a_2, \ldots, a_m$ и $b_1, b_2, \ldots, b_n$ — целые, причем $\sum_{i=1}^{m}a_i=\sum_{j=1}^{n}b_j$, то транспортная задача (9.39)—(9.42) имеет оптимальное решение с целыми координатами.

3°. Ранг системы векторов условий транспортной задачи (9.39)—(9.42) равен $m+n-1$ (m — число пунктов производства, n — число пунктов потребления).

Транспортную задачу можно решать симплекс-методом. Однако в силу специфики ее системы ограничений можно значительно упростить все этапы решения.

9.13. Опорные решения транспортной задачи

Условия транспортной задачи удобно записывать в транспортную таблицу, в которой строки соответствуют пунктам производства, а столбцы — пунктам потребления (табл. 9.5).

Таблица 9.5

c_{11}	c_{12}	c_{13}	...	c_{1n}	a_1
c_{21}	c_{22}	c_{23}	...	c_{2n}	a_2
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
c_{m1}	c_{m2}	c_{m3}	...	c_{mn}	a_m
b_1	b_2	b_3	...	b_n	$b_j \backslash a_i$

При этом каждому неизвестному x_{ij} ($i=1, 2, \ldots, m$; $j=1, 2, \ldots, n$) соответствует клетка (i, j) транспортной таблицы.

Циклом в транспортной таблице называют замкнутую ломаную линию, удовлетворяющую следующим трем условиям:

1) все вершины ломаной находятся в клетках таблицы;

2) ребра ломаной расположены по строкам или по столбцам таблицы;

3) к каждой вершине подходят ровно два ребра, причем одно — по строке, а другое — по столбцу (рис. 9.4).

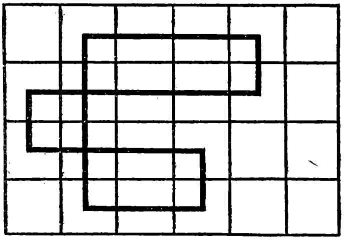


Рис. 9.4

Набор клеток транспортной таблицы называют *ацикличным набором*, если не существует цикла, все вершины которого расположены в клетках этого набора.

Пусть $\alpha=(x_{11}^0; \ldots; x_{ij}^0; \ldots; x_{mn}^0)$ — допустимое решение транспортной задачи (9.39) — (9.42). Ненулевые координаты α запишем в соответствующие клетки транспортной таблицы 9.5. Допустимый вектор α является опорным решением транспортной задачи тогда и только тогда, когда набор заполненных клеток ацикличен.

Опорное решение транспортной задачи можно построить методом «минимального элемента». Для этого среди всех клеток табл. 9.5 выбираем клетку с наименьшим c_{ij} (если таких клеток несколько, то берем любую из них).

Пусть (r, s) — такая клетка. Полагаем $x_{rs}=\min\{a_r, b_s\}$. Если $x_{rs}=a_r$, то, заменив b_s на $b_s'=b_s-a_r$, вычеркиваем r-ю строку транспортной таблицы. При $x_{rs}=b_s$ заменяем число a_r на $a_r'=a_r-b_s$ и вычеркиваем s-й столбец (если $x_{rs}=a_r=b_s$, то вычеркиваем r-ю строку и s-й столбец одновременно). В результате получаем новую таблицу меньшего размера, для которой повторяем указанную процедуру. Через конечное число шагов будет построено опорное решение.

Предположим, что $\alpha=(x_{11}^0; \ldots; x_{ij}^0; \ldots; x_{mn}^0)$ — некоторое опорное решение транспортной задачи. Ненулевые координаты α запишем в соответствующие клетки транспортной таблицы. Если заполненных клеток окажется меньше, чем $m+n-1$ клетка, то дополнительно в некоторые клетки допишем нули так, чтобы в результате образовался ацикличный набор из $m+n-1$ заполненных клеток (в этом случае векторы условий, соответствующие заполненным клеткам, составляют базис опорного решения α).

Потенциалами опорного решения α называют числа u_i $(i=1, 2, \ldots, m)$ и v_j $(j=1, 2, \ldots, n)$ такие, что

$$u_i+v_j=c_{ij} \tag{9.43}$$

для всех заполненных клеток (i, j).

Соотношения (9.43) представляют собой систему $m+n-1$ линейных уравнений с $m+n$ неизвестными. Эта система всегда совместна, причем одно из неизвестных можно положить равным любому числу и тогда все остальные неизвестные определятся однозначно.

Достаточное условие оптимальности опорного решения. Пусть u_i $(i=1, 2, \ldots, m)$ и v_j $(j=1, 2, \ldots, n)$—потенциалы опорного решения α транспортной задачи (в транспортной таблице заполнены клетки, образующие ацикличный набор).

Если для всех незаполненных клеток (i, j) $u_i+v_j \leqslant c_{ij}$, то α—оптимальное решение транспортной задачи.

○ **Пример.** Рассмотрим опорное решение α, ненулевые координаты которого записаны в транспортную таблицу (табл. 9.6).

Таблица 9.6

u_i \ v_j	2	4	9	3
0	2 90	4	10	10
—1	1 20	3 80	10	4
4	6	8 20	13 80	7 40

Для отыскания потенциалов данного опорного решения необходимо найти некоторое решение системы линейных уравнений:

$$u_1+v_1=2,\quad u_3+v_2=8,$$
$$u_2+v_1=1,\quad u_3+v_3=13,$$
$$u_2+v_2=3,\quad u_3+v_4=7.$$

Полагая $u_1=0$, получаем $v_1=2$, $u_2=-1$, $v_2=4$, $u_3=4$, $v_3=9$, $v_4=3$. Проверка показывает, что для всех незаполненных клеток $(i,\ j)$ $u_i+v_j\leqslant c_{ij}$.
Следовательно, данное опорное решение оптимально.

9.14. Решение транспортной задачи методом потенциалов

Решение транспортной задачи методом потенциалов проводится по следующей схеме:

1) находят начальное опорное решение $\alpha_1=(x'_{11};\ \ldots;\ x'_{ij};\ \ldots;\ x'_{mn})$. (Например, методом «минимального элемента».) Если в транспортной таблице заполненных клеток оказалось меньше, чем $m+n-1$, дополнительно дописывают нули так, чтобы получился ацикличный набор из $m+n-1$ заполненных клеток;

2) вычисляют потенциалы u_i $(i=1,\ 2,\ \ldots,\ m)$; v_j $(j=1,\ 2,\ \ldots,\ n)$ опорного решения α_1. Если для всех незаполненных клеток $(i,\ j)$: $u_i+v_j\leqslant c_{ij}$, то α_1 — оптимальное решение и транспортная задача решена. В противном случае выбираем некоторую клетку $(r,\ s)$ такую, что $u_r+v_s>c_{rs}$;

3) в транспортной таблице строим цикл, одна вершина которого находится в клетке $(r,\ s)$, а все остальные вершины — в заполненных клетках (такой цикл всегда существует, и притом только один.) Каждой вершине цикла присваивают знак «+» или «—» следующим образом: вершине в клетке $(r,\ s)$ присваивают знак «+», а дальше расставляют знаки так, чтобы они от вершины к вершине чередовались.

Обозначим через ρ наименьшее из чисел $\{x'_{ij}\}$, стоящих в клетках, которым присвоен знак «—», и пусть $\rho=x'_{kl}$ (если число находится не в одной, а в нескольких клетках, выбираем любую из них.) После этого заполняем транспортную таблицу согласно следующему правилу:

а) клетки, в которые не попали вершины цикла, заполняют так же, как и раньше;

б) если клетке $(i,\ j)$ присвоен знак «+», то в эту клетку записывают число $x'_{ij}+\rho$;

в) клетку $(k,\ l)$ не заполняют, а в остальные отрицательные клетки $(i,\ j)$ записывают число $x'_{ij}-\rho$.

В результате получаем ацикличный набор из $m+n-1$ заполненных клеток транспортной таблицы, который определяет опорное решение α_2 такое, что $f(\alpha_2) \leqslant f(\alpha_1)$.

Принимая α_2 за исходное опорное решение, повторяем п. 2) и 3) и т. д. Через конечное число таких шагов будет найдено оптимальное решение транспортной задачи.

○ Рассмотрим следующую транспортную задачу:

4	3	6	4	40
1	6	2	8	30
2	4	5	7	20
30	25	15	20	b_j \ a_i

Ниже приведены все этапы решения этой задачи (начальное опорное решение построено методом «минимального элемента»):

u_i \ v_j	1	3	2	4
0	4	3 25	6	4 15
0	1 — 30	6	2 + 0	8
3	2 +	4	5 — 15	7 5

$\rho = 15$

u_i \ v_j	-1	3	0	4
0	4	3 — 25	6	4 + 15
2	1 15	6	2 15	8
3	2 15	4 +	5	7 5 —

$\rho = 5$

$u_i \backslash v_j$	1	3	2	4
0	4	3 20	6	4 20
0	1 15	6	2 15	8
1	2 15	4 5	5	7

В последней таблице записан оптимальный план перевозок, стоимость которых составляет: $20 \cdot 3 + 20 \cdot 4 + 15 \cdot 1 + 15 \cdot 2 + 15 \cdot 2 + 5 \cdot 4 = 235$ единиц.

9.15. Параметрические задачи линейного программирования

Семейство задач линейного программирования в канонической форме

$$f_t = \sum_{j=1}^{n} (\gamma'_j + \gamma''_j t) x_j \,(\min), \tag{9.44}$$

$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \; i = 1, 2, \ldots, m, \tag{9.45}$$

$$x_j \geqslant 0, \; j = 1, 2, \ldots, n \tag{9.46}$$

(где γ'_j, γ''_j — некоторые числа ($j = 1, 2, \ldots, n$); t — параметр, $-\infty < t < +\infty$) называется *параметрической задачей линейного программирования.*

Параметрическая задача линейного программирования возникла в связи с необходимостью изучить поведение оптимального решения задачи линейного программирования в зависимости от тех или иных изменений коэффициентов целевой функции.

Пусть α — некоторое опорное решение задачи (9.44)—(9.46). Обозначим через $\Delta(\alpha)$ множество всех значений параметра t, при которых α является оптимальным решением задачи (9.44)—(9.46).

Возможны лишь следующие случаи: 1) $\Delta(\alpha) = \varnothing$;

2) $\Delta(\alpha) = \{t_0\}$; 3) $\Delta(\alpha) = [t_1, t_2]$; 4) $\Delta(\alpha) =]-\infty, t_1]$;

5) $\Delta(\alpha) = [t_2, +\infty[$; 6) $\Delta(\alpha) =]-\infty, +\infty[$.

Рассмотрим некоторый базис $B_1, B_2, \ldots, B_r$ опорного решения α. Тогда оценки этого базиса для задачи (9.44)—

(9.46) имеют вид

$$\delta_j' + \delta_j'' t, \quad j = 1, 2, \ldots, n.$$

Обозначим через $T(B_1, B_2, \ldots, B_r)$ множество всех решений системы неравенств

$$\delta_j' + \delta_j'' t \leqslant 0, \quad j = 1, 2, \ldots, n.$$

Тогда:

1) $$\Delta(\alpha) \supseteq T(B_1, B_2, \ldots, B_r);$$

2) если α — невырожденное опорное решение, то

$$\Delta(\alpha) = T(B_1, B_2, \ldots, B_r);$$

3) если α — вырожденное опорное решение, то

$$\Delta(\alpha) = \bigcup T(B_1, B_2, \ldots, B_r),$$

где объединение берется по всем базисам опорного решения α.

Таким образом, всегда можно найти множество всех значений параметра t, при которых данное опорное решение является оптимальным решением параметрической задачи линейного программирования.

○ Например, $\alpha = (5; 0; 6; 0)$ — невырожденное опорное решение задачи

$$f_t = (-6 + t) x_1 + (3 + t) x_2 + (-11 + 2t) x_3 + (1 - t) x_4 \ (\min),$$
$$2x_1 + x_2 + 3x_3 = 28,$$
$$x_1 + x_2 + x_3 - x_4 = 11,$$
$$x_j \geqslant 0, \ j = 1, 2, 3, 4.$$

Симплекс-таблицу для данной задачи приведем к базису A_1, A_3 опорного решения α:

x_1	x_2	x_3	x_4	
2	1	3	0	28
1	1	1	—1	11
$6-t$	$-3-t$	$11-2t$	$t-1$	0

$\longrightarrow$

x_1	x_2	x_3	x_4	
0	—1	1	2	6
1	2	0	—3	5
$6-t$	$-3-t$	$11-2t$	$t-1$	0

x_1	x_2	x_3	x_4	
0	—1	1	2	6
1	2	0	—3	5
0	$-4-t$	0	$2t-5$	$17t-96$

Опорное решение α является оптимальным решением данной задачи тогда и только тогда, когда

$$\begin{cases} -4-t\leqslant 0 \\ 2t-5\leqslant 0 \end{cases} \Leftrightarrow \begin{matrix} t\geqslant -4, \\ t\leqslant 5/2, \end{matrix}$$

т. е. $\Delta(\alpha)=[-4,\ 5/2]$. При этом $f_t(\alpha)=17t-96$. ●

Если параметрическая задача линейного программирования (9.44)—(9.46) имеет непустое допустимое множество, то существуют опорные решения этой задачи $\alpha_1, \alpha_2, \ldots$ $\ldots, \alpha_s$ такие, что:

2) пересечение множеств $\Delta(\alpha_i)$ и $\Delta(\alpha_j)$, $i\neq j$, является либо пустым, либо состоит из единственной точки;

2) если $t\notin \bigcup\limits_{i=1}^{s} \Delta(\alpha_i)$, то целевая функция задачи (9.44)—(9.46) не ограничена снизу на допустимом множестве.

Существует специальный алгоритм, позволяющий за конечное число шагов найти эти опорные решения α_i $(i=1, 2, \ldots, s)$ и определить соответствующие множества $\Delta(\alpha_i)$.

9.16. Целочисленные задачи линейного программирования

Целочисленная задача линейного программирования — это оптимизационная задача (V, f, Ω), где целевая функция f является линейной функцией на $V=R^n$, а Ω — множеством решений некоторой системы линейных уравнений или линейных неравенств, у которых координаты с заданными номерами — целые числа.

Целочисленная задача линейного программирования отличается от общей задачи линейного программирования только дополнительным требованием о целочисленности некоторых (быть может, и всех) неизвестных. Если в задаче требуется целочисленность всех неизвестных, то ее называют *полностью целочисленной*.

Задача о загрузке корабля. В морском порту имеются предметы n видов. Предмет j-го вида имеет массу a_j и ценность c_j $(j=1, 2, \ldots, n)$. Требуется загрузить корабль данной грузоподьемности b так, чтобы ценность груза была наибольшей.

Обозначив через x_j количество предметов j-го вида $(j=1, 2, \ldots, n)$, которые необходимо погрузить на корабль, получим целочисленную задачу линейного программирования программирования:

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

$$\sum_{j=1}^{n} a_j x_j \leqslant b,$$

$$x_j \geqslant 0,\ j=1,\ 2,\ \ldots,\ n,$$

$$x_j \text{ — целое},\ j=1,\ 2,\ \ldots,\ n.$$

Задача о распределении капиталовложений между проектами. Имеется n проектов $P_1, \ldots, P_j, \ldots, P_n$, причем для каждого проекта P_j известны ожидаемый эффект γ_j от его реализации и необходимая величина капиталовложений g_j. Общий объем капиталовложений не может превышать заданной величины b. Требуется определить, какие проекты необходимо реализовать, чтобы суммарный эффект был наибольшим.

Введем неизвестные x_j $(j=1, 2, \ldots, n)$, полагая

$$x_j=\begin{cases} 1, \text{ если проект } P_j \text{ реализуется;} \\ 0, \text{ если проект } P_j \text{ не реализуется.} \end{cases}$$

Получим целочисленную задачу линейного программирования:

$$f=\sum_{j=1}^{n} \gamma_j x_j \ (\max),$$

$$\sum_{j=1}^{n} g_j x_j \leqslant b,$$

$$0 \leqslant x_j \leqslant 1,\ x_j \text{ — целое},\ j=1, 2, \ldots, n.$$

В ряде случаев оптимизационную задачу с разрывной целевой функцией удается свести к целочисленной задаче линейного программирования.

Рассмотрим, например, задачу максимизации с целевой функцией

$$f=\sum_{j=1}^{n} c_j(x_j),$$

где

$$c_j(x_j)=\begin{cases} 0, \text{ если } x_j=0, \\ c_j x_j+\mathrm{d}_j, \text{ если } x_j>0, \end{cases}$$

допустимое множество Ω которой задано системой ограничений

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \ldots, m,$$
$$x_j \geqslant 0, \quad j = 1, 2, \ldots, n.$$

Введем n новых неизвестных, полагая

$$y_j = \begin{cases} 1, & \text{если } x_j > 0, \\ 0, & \text{если } x_j = 0. \end{cases}$$

Если допустимое множество Ω ограничено, то исходная оптимизационная задача сводится к целочисленной задаче линейного программирования:

$$f = \sum_{j=1}^{n} (c_jx_j + d_jy_j) \quad (\max),$$
$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \ldots, m,$$
$$x_j \geqslant 0; \; x_j \leqslant M \cdot y_j; \; 0 \leqslant y_j \leqslant 1;$$
$$y_j \text{ — целое}, \quad j = 1, 2, \ldots, n,$$

где M — достаточно большое число.

Если допустимое множество целочисленной задачи линейного программировании является конечным множеством, то это комбинаторная оптимизационная задача. Классическим примером комбинаторной оптимизационной задачи наряду с задачами о загрузке корабля и о распределении капиталовложений между проектами является задача о коммивояжере.

Задача о коммивояжере. Имеется n городов $A_1, A_2, \ldots, A_n$ и задана матрица $C = (c_{ij})$ расстояний между этими городами. Выезжая из исходного города A_1, коммивояжер должен побывать во всех остальных городах по одному разу и вернуться в A_1. Определить, в каком порядке следует объезжать города, чтобы суммарное пройденное расстояние было наименьшим.

Если в целочисленной задаче линейного программирования отбросить требование о целочисленности неизвестных, то получим задачу линейного программирования, которая называется *ослаблением исходной целочисленной задачи*.

Целочисленное оптимальное решение ослабления является оптимальным решением и исходной целочисленной задачи.

В частности, если ослабление окажется транспортной задачей с целочисленным вектором ограничений, то оптимальное решение транспортной задачи (которое при этих условиях всегда может быть выбрано целочисленным) будет оптимальным решением исходной целочисленной задачи линейного программирования.

В большинстве же случаев ослабление целочисленной задачи линейного программирования имеет только нецелочисленное оптимальное решение, причем если округлить нецелочисленные координаты этого решения, то нельзя получить даже допустимого решения исходной задачи.

С другой стороны, для решения комбинаторной оптимизационной задачи можно попробовать перебрать все допустимые решения этой задачи и выбрать среди них такое, на котором целевая функция принимает наибольшее (наименьшее) значение. Однако встречаются задачи, у которых допустимых решений очень много и перебрать их все практически невозможно.

Таким образом, целочисленные задачи линейного программирования образуют специфический класс оптимизационных задач, для решения которых требуются специальные методы.

Известны различные методы решения целочисленных задач линейного программирования: методы отсечений, метод ветвей и границ, метод Беллмана. Эффективность того или иного метода зависит от конкретных условий целочисленной задачи линейного программирования.

9.17. Метод отсечения для целочисленных задач линейного программирования

Дана полностью целочисленная задача линейного программирования:

$$f=\sum_{j=1}^{n}\gamma_j x_j\,(\min), \tag{9.47}$$

$$\sum_{j=1}^{n}a_{ij}x_j=b_i,\quad i=1,\,2,\,\ldots,\,m, \tag{9.48}$$

$$x_j\geqslant 0,\quad j=1,\,2,\,\ldots,\,n, \tag{9.49}$$

$$x_j\text{ — целое},\quad j=1,\,2,\,\ldots,\,n. \tag{9.50}$$

Методы отсечений опираются на следующее утверждение: если ослабление задачи (9.47)—(9.50) имеет нецелочисленное оптимальное опорное решение α^0, то существует неравенство вида

$$\sum_{j=1}^{n} \lambda_j x_j \geq \Delta, \qquad (9.51)$$

которому α^0 не удовлетворяет, а любое допустимое решение исходной целочисленной задачи — удовлетворяет.

Неравенство (9.51), обладающее указанными свойствами, называют *правильным отсечением*. Правильное отсечение, например, можно построить следующим образом.

Симплекс-таблицу для ослабления задачи (9.47) — (9.50) приведем к базису α^0, все оценки которого не положительны. В полученной таблице выбираем строку с дробным свободным членом d_p:

x_1	...	x_q	...	x_n	
⋮ a'_{p1} ⋮		⋮ a'_{pq} ⋮		⋮ a'_{pn} ⋮	⋮ d_p ⋮
δ_1	...	δ_q	...	δ_n	δ_0

Обозначим через $[a'_{pq}]$ целую часть числа a'_{pq} ($q = 1, 2, \ldots, n$), т. е. ближайшее целое число, не превосходящее a'_{pq}, а через $[d_p]$ — целую часть числа d_p. Правильное отсечение в этом случае имеет вид

$$(a'_{p1} - [a'_{p1}])\, x_1 + \ldots + (a'_{pq} - [a'_{pq}])\, x_q + \ldots$$
$$\ldots + (a'_{pn} - [a'_{pn}])\, x_n \geq d_p - [d_p].$$

○ Рассмотрим симплекс-таблицу

x_1	x_2	x_3	x_4	
1/2	—5/2	1	0	3/2
—7/2	5/3	0	1	7/3
—2	—1	0	0	5

Так как $[1/2] = 0$, $[-5/2] = -3$, $[1] = 1$, $[0] = 0$, $[3/2] = 1$, то по первой строке можно построить правильное отсечение $(1/2)x_1+(1/2)x_2 \geqslant 1/2$. Аналогично, по второй строке строится правильное отсечение $(1/2)x_1 + (2/3)x_2 \geqslant 1/3$. ●

В дальнейшем будем считать, что коэффициенты при неизвестных и свободные члены системы линейных уравнений (9.48) являются целыми числами.

Для решения целочисленной задачи линейного программирования (9.47)—(9.50) методом отсечений необходимо выполнить ряд последовательных шагов. На каждом шаге решается задача линейного программирования; если эта задача оказывается неразрешимой, то и исходная целочисленная задача является неразрешимой.

Первый шаг. Находим оптимальное опорное решение $\beta_1=(x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)})$ ослабления целочисленной задачи (9.47)—(9.50). Если β_1 окажется целочисленным, то оно—искомое. В противном случае строим правильное отсечение, записываем его в виде

$$\sum_{j=1}^{n} \lambda_j^{(1)} x_j - x_{n+1} = \Delta^{(1)}, \quad x_{n+1} \geqslant 0,$$

и добавляем к системе ограничений исходной задачи.

k-й шаг ($k \geqslant 2$). Находим оптимальное опорное решение $\beta_k=(x_1^{(k)}; x_2^{(k)}; \ldots; x_n^{(k)}; x_{n+1}^{(k)}; \ldots; x_{n+k-1}^{(k)})$ ослабления задачи, построенной на предыдущем шаге. Если β_k целочисленно, то $\alpha_k=(x_1^{(k)}, x_2^{(k)}; \ldots; x_n^{(k)})$—оптимальное решение исходной задачи (9.47)—(9.50). В противном случае строим правильное отсечение $\sum\limits_{j=1}^{n+k-1} \lambda_j^{(k)} \cdot x_j - x_{n+k} = \Delta^{(k)}$, $x_{n+k} \geqslant 0$ и добавляем его к системе ограничений задачи, полученной на $(k-1)$-м шаге. После этого выполняем $(k+1)$-й шаг.

При определенных условиях всегда можно гарантировать конечность этого алгоритма, т. е. через конечное число шагов либо будет найдено оптимальное решение, либо установлено, что задача неразрешима.

9.18. Метод ветвей и границ для целочисленных задач линейного программирования

Дана оптимизационная задача минимизации (V, f, Ω). Предположим, что допустимое множество Ω разбито на конечное число непересекающихся подмножеств:

$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_k \cup \ldots \cup A_l. \tag{9.52}$$

Выберем некоторое подмножество A_k и разобьем его на конечное число непересекающихся подмножеств:

$$A_k = B_1^{(k)} \cup \ldots \cup B_m^{(k)}.$$

Тогда Ω можно представить в виде

$$\Omega = A_1 \cup \ldots \cup A_{k-1} \cup B_1^{(k)} \cup \ldots \cup B_m^{(k)} \cup A_{k+1} \cup \ldots \cup A_l. \tag{9.53}$$

Переход от разбиения (9.52) множества Ω к разбиению (9.53) того же множества называют *ветвлением* с помощью подмножества A_k. Схематически процесс ветвления можно изобразить так:

$$\begin{tikzcd} & & \Omega \arrow[dll] \arrow[d] \arrow[drr] & & \\ A_1 & \ldots & A_k \arrow[dl] \arrow[d] \arrow[dr] & \ldots & A_l \\ & B_1^{(k)} & \ldots & B_m^{(k)} & \end{tikzcd}$$

Если A — некоторое подмножество множества Ω, то *оценкой подмножества* A называется число $\xi(A)$ такое, что $f(x) \geqslant \xi(A)$ для любого $x \in A$.

Для решения задачи минимизации (V, f, Ω) методом ветвей и границ необходимо выполнить ряд последовательных шагов.

Нулевой шаг. Допустимое множество Ω некоторым способом разбиваем на конечное число непересекающихся подмножеств.

Первый шаг. Вычисляем оценки образовавшихся подмножеств и выбираем подмножество с наименьшей оценкой. Если в выбранном подмножестве удается найти элемент α^0 такой, что $f(\alpha^0)$ совпадает с оценкой этого подмножества, то α^0 — оптимальное решение задачи (V, f, Ω). В противном случае осуществляем ветвление с помощью выбранного подмножества и переходим ко второму шагу.

k-й шаг ($k \geqslant 2$). Рассматриваем разбиение множества Ω на подмножества, которое было получено на предыдущем шаге. Для каждого из этих подмножеств находим его оценку и выбираем подмножество с наименьшей оценкой. Если в выбранном подмножестве удается найти элемент α^0 такой, что $f(\alpha^0)$ совпадает с оценкой этого подмножества, то α^0 — оптимальное решение задачи (V, f, Ω).

В противном случае осуществляем ветвление с помощью выбранного подмножества и переходим к следующему шагу.

Для применения метода ветвей и границ в каждом конкретном случае необходимо:

1) задать первоначальное разбиение допустимого множества на конечное число непересекающихся подмножеств;

2) определить способ разбиения подмножеств, с помощью которых на каждом шаге осуществляются ветвления;

3) сформулировать правило вычисления оценок всех встречающихся подмножеств.

В основном метод ветвей и границ применяют для решения целочисленных задач линейного программирования, так как именно в этом случае часто удается найти приемлемое правило вычисления оценок. При этом оценки вычисляют либо с помощью решения некоторых вспомогательных задач линейного программирования, либо находят на основе комбинаторной структуры рассматриваемой задачи.

○ Рассмотрим задачу о коммивояжере с матрицей расстояний $C=(c_{ij})$ порядка n. Обозначим через Ω множество всех маршрутов, при которых коммивояжер, выезжая из города A_1, побывает во всех остальных городах по одному разу и вернется в A_1. Пусть $\Omega_{1, i_2, i_3, \ldots, i_k}$, $2 \leqslant k \leqslant n-1$, — подмножество множества Ω, состоящее из всех маршрутов, при которых коммивояжер, выезжая из A_1, последовательно посещает города $A_{i_2}, \ldots, A_{i_k}$.

Множество Ω можно разбить на непересекающиеся подмножества $\Omega_{1,2}, \Omega_{1,3}, \ldots, \Omega_{1,n}$. Для осуществления ветвления с помощью подмножества $\Omega_{1, i_2, \ldots, i_k}$, $2 \leqslant k \leqslant n-2$, это подмножество разбиваем на подмножества вида $\Omega_{1, i_2, i_3, \ldots, i_k, j}$, где $j \neq 1, i_2, i_3, \ldots, i_k$.

Например, подмножество $\Omega_{1,2}$ разбиваем на подмножества $\Omega_{1,2,3}, \Omega_{1,2,4}, \ldots, \Omega_{1,2,n}$.

Множество $\Omega_{1, i_2, i_3, \ldots, i_k}$ при $k=n-1$ состоит из единственного маршрута, и с помощью него ветвление не выполняется.

Оценку подмножества $\Omega_{1, i_2, i_3, \ldots, i_{k-1}, i_k}$, $2 \leqslant k \leqslant n-1$, вычисляем по формуле

$$\xi(\Omega_{1, i_2, i_3, \ldots, i_{k-1}, i_k}) = c_{1 i_2} + \ldots + c_{i_{k-1} i_k} + \min_{j \neq 1, i_2, \ldots, i_k} c_{i_k j} + (n-k) \min_{\substack{i \neq 1, i_2, \ldots, i_k \\ l \neq i_2, \ldots, i_k}} c_{il}.$$

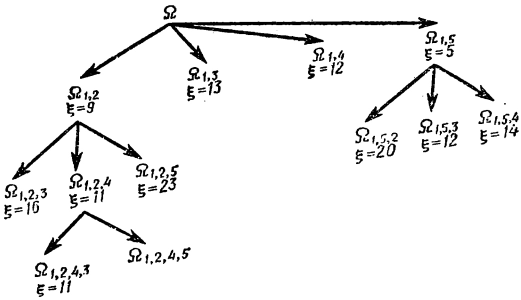
Если $k=n-1$, то оценка подмножества $\Omega_{1, i_2, \ldots, i_k}$ совпадает с длиной единственного маршрута, входящего в это множество.

Если

$$C=\begin{pmatrix} * & 4 & 9 & 6 & 1 \\ 2 & * & 9 & 2 & 10 \\ 11 & 11 & * & 8 & 1 \\ 5 & 4 & 3 & * & 8 \\ 1 & 11 & 1 & 8 & * \end{pmatrix},$$

то $\xi(\Omega_{1,2})=4+2+(5-2)\cdot 1=9$, $\xi(\Omega_{1,2,3})=4+9+1+$ $+(5-3)\cdot 1=16$, $\xi(\Omega_{1,2,4,3})=4+2+3+1+1\cdot 1=11$.

Решение задачи о коммивояжере с данной матрицей приведено на следующей схеме:



Оптимальный маршрут: $A_1 \to A_2 \to A_4 \to A_3 \to A_5 \to$ $\to A_1$. ●

9.19. Метод Беллмана для решения целочисленных задач линейного программирования

Дана целочисленная задача линейного программирования

$$f=\sum_{j=1}^{n} \gamma_j x_j \,(\max); \tag{9.54}$$

$$\sum_{j=1}^{n} g_j x_j \leqslant b; \tag{9.55}$$

$$0 \leqslant x_j \leqslant d_j, \ x_j \text{—целое}, \ j=1, 2, \ldots, n, \tag{9.56}$$

где g_j, d_j $(j=1, 2, \ldots, n)$, b—целые положительные числа.

Обозначим через Ω_z, $z=0, 1, 2, \ldots, b$, множество всех целочисленных решений системы неравенств

$$\sum_{j=1}^{n} g_j x_j \leqslant z,$$
$$0 \leqslant x_j \leqslant d_j, \quad j=1, 2, \ldots, n.$$

Множество Ω_z состоит из одного нулевого вектора при $z=0$ и совпадает с множеством допустимых решений задачи (9.54)—(9.56) при $z=b$. Кроме того, если $\alpha = (x_1, \ldots, x_k, \ldots, x_n) \in \Omega_z$, то $0 \leqslant x_k \leqslant \min(d_k, z/g_k)$.

Пусть $\varphi_k(z)$, $k=1, 2, \ldots, n$,— наибольшее значение функции $\sum_{j=1}^{k} \gamma_j x_j$ на множестве Ω_z. Функции $\varphi_1(z), \varphi_2(z), \ldots$ $\ldots, \varphi_n(z)$, определенные при $z=0, 1, 2, \ldots, b$, называются *функциями Беллмана* для задачи (9.54)—(9.56). Для этих функций имеют место следующие равенства:

$$\varphi_1(z) = \max_{\substack{0 \leqslant x_1 \leqslant \min(d_1, z/g_1), \\ x_1 - \text{целое};}} \{\gamma_1 x_1\}, \tag{9.57}$$

$$\varphi_k(z) = \max_{\substack{0 \leqslant x_k \leqslant \min(d_k, z/g_k), \\ x_k - \text{целое}.}} \{\gamma_k x_k + \varphi_{k-1}(z - g_k x_k\}, \tag{9.58}$$

Обозначим через $x_k^*(z)$, $k=1, 2, \ldots, n$, то целое значение x_k, $0 \leqslant x_k \leqslant \min(d_k, z/g_k)$, при котором достигается максимум выражения, стоящего в правой части соответствующего равенства.

Задачу (9.54)—(9.56) решают по следующей схеме:

1. По формуле (9.57) находят все значения функции $\varphi_1(z)$, $z=0, 1, 2, \ldots, b$. Одновременно определяют $x_1^*(z)$.

2. Используя рекуррентное соотношение (9.58), последовательно находят функции Беллмана $\varphi_2(z), \ldots, \varphi_{n-1}(z)$ и значение функции $\varphi_n(z)$ при $z=b$. Одновременно определяют $x_2^*(z), \ldots, x_{n-1}^*(z)$ и $x_n^*(b)$.

3. Строят оптимальное решение $\alpha^0 = (x_1^0; x_2^0; \ldots; x_{n-1}^0; x_n^0)$ задачи (9.54)—(9.56), полагая $x_n^0 = x_n^*(b)$, $x_{n-1}^0 = x_{n-1}^*(b - g_n x_n^0)$, $\ldots$, $x_1^0 = x_1^*(b - g_n x_n^0 - g_{n-1} x_{n-1}^0 - \ldots - g_2 x_2^0)$.

○ Рассмотрим следующую целочисленную задачу:

$$f = 3x_1 + 4x_2 + 5x_3 \text{ (max)},$$
$$2x_1 + 5x_2 + 4x_3 \leqslant 12,$$
$$0 \leqslant x_j \leqslant 3, \quad j=1, 2, 3.$$

В данном случае $\varphi_1(z) = \max\limits_{0 \leqslant x_1 \leqslant \min(3,\ z/2)} \{3x_1\}$. Значения этой функции и соответствующие числа $x_1^*(z)$, $z=0, 1, 2, \ldots$ $\ldots 12$, приведены в табл. 9.7.

Таблица 9.7

z	0	1	2	3	4	5	6	7	8	9	10	11	12
$\varphi_1(z)$	0	0	3	3	6	6	9	9	9	9	9	9	9
$x_1^*(z)$	0	0	1	1	2	2	3	3	3	3	3	3	3

По определению,

$$\varphi_2(z)=\max_{0\leqslant x_2\leqslant \min(3,\ z/5)}\{4x_2+\varphi_1(z-5x_2)\}.$$

В табл. 9.8 приведены числа $4x_2+\varphi_1(z-5x_2)$ при $z=$ $=0, 1, 2, \ldots, 12$ и $0\leqslant x_2\leqslant \min(3,\ z/5)$, все значения функции $\varphi_2(z)$ и числа $x_2^*(z)$.

Таблица 9.8

x_2	z												
	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	0	3	3	6	6	9	9	9	9	9	9	9
1	—	—	—	—	—	4	4	7	7	10	10	13	13
2	—	—	—	—	—	—	—	—	—	—	8	8	11
3	—	—	—	—	—	—	—	—	—	—	—	—	—
$\varphi_2(z)$	0	0	3	3	6	6	9	9	9	10	10	13	13
$x_2^*(z)$	0	0	0	0	0	0	0	0	0	1	1	1	1

Так как $\min(3,\ 12/4)=3$, то $\varphi_3(12)=\max\limits_{0\leqslant x_3\leqslant 3}\{5x_3+$ $+\varphi_2(12-4x_3)\}$, т. е. $\varphi_3(12)=\max\{0+13,\ 5+9,\ 10+6,$

$15+0\}=16$. При этом $x_3^*(12)=2$. Тогда

$$x_3^0=x_3^*(12)=2;\quad x_2^0=x_2^*(12-2\cdot 4)=x_2^*(4)=0,$$
$$x_1^0=x_1^*(12-2\cdot 4-0\cdot 5)=x_1^*(4)=2.$$

Следовательно, $\alpha^0=(2;\ 0;\ 2)$ является оптимальным решением нашей задачи, причем $f(\alpha^0)=\varphi_3(12)=16$. ●

9.20. Задачи нелинейного программирования

Оптимизационная задача $(V,\ f,\ \Omega)$ называется задачей нелинейного программирования, если целевая функция f является функцией n переменных ($V=R^n$), а допустимое множество Ω — множеством решений некоторой системы уравнений и неравенств с n неизвестными.

Так как любое уравнение равносильно системе двух неравенств

$$\begin{cases} \Phi(x_1,\ x_2,\ \ldots,\ x_n)\leqslant 0, \\ -\Phi(x_1,\ x_2,\ \ldots,\ x_n)\leqslant 0, \end{cases}$$

то можно считать, что допустимое множество Ω задачи нелинейного программирования задается только неравенствами.

Таким образом, задача нелинейного программирования — это задача максимизации или минимизации целевой функции

$$f=f(x_1,\ x_2,\ \ldots,\ x_n) \tag{9.59}$$

на множества решений системы неравенств, которую удобно записывать в следующем виде:

$$\begin{cases} \Phi_i(x_1,\ x_2,\ \ldots,\ x_n)\leqslant 0,\quad i=1,\ 2,\ \ldots,\ m, & (9.60) \\ x_j\geqslant 0,\quad j\in J\subseteq N=\{1,\ 2,\ 3,\ \ldots,\ n\}, & (9.61) \end{cases}$$

где $\Phi_i(x_1,\ x_2,\ \ldots,\ x_n)$ — некоторые функции, определенные на всем пространстве R^n.

Задачи нелинейного программирования образуют широкий класс оптимизационных задач. В частности, к этому классу принадлежит задача максимизации или минимизации функции n переменных на всем пространстве $\mathbf{R}^n$.

Кроме того, класс задач нелинейного программирования включает в себя задачи линейного программирования.

Введем следующие обозначения:

$$(\mathbf{R}^n)^+=\{M(x_1;\ x_2;\ \ldots;\ x_n)\in\mathbf{R}^n\,|\,x_j\geqslant 0,\ j=1,\ 2,\ \ldots,\ n\},$$
$$\mathbf{R}_J^n=\{M(x_1;\ x_2;\ \ldots;\ x_n)\in\mathbf{R}^n\,|\,x_j\geqslant 0,\ j\in J\},$$

где $J \subseteq N = \{1, 2, \ldots, n\}$. Очевидно, что если $J = \varnothing$, то $\mathbf{R}_J^n = \mathbf{R}^n$; если $J = N$, то $\mathbf{R}_J^n = (\mathbf{R}^n)^+$.

Функция

$$L(M, P) = f(M) - \sum_{i=1}^{m} y_i \Phi_i(M)$$

(где $M(x_1; x_2; \ldots; x_n) \in \mathbf{R}_J^n$, $P(y_1; y_2; \ldots; y_m) \in (\mathbf{R}^m)^+$) называется *функцией Лагранжа* для задачи максимизации (9.59)—(9.61) $\Bigg($для задачи минимизации функция Лагранжа имеет вид

$$L(M, P) = -f(M) - \sum_{i=1}^{m} y_i \Phi_i(M)\Bigg).$$

Точка $M_0(x_1^0; \ldots; x_j^0; \ldots; x_n^0) \in \mathbf{R}_J^n$ является *точкой Куна—Таккера* для задачи (9.59)—(9.61), если существует точка $P_0(y_1^0; \ldots; y_i^0; \ldots; y_m^0) \in (\mathbf{R}^m)^+$ такая, что справедливы условия Куна—Таккера:

$$\frac{\partial L}{\partial x_j}(M_0, P_0) \leqslant 0, \; j \in J, \qquad \frac{\partial L}{\partial y_i}(M_0, P_0) \geqslant 0,$$

$$\frac{\partial L}{\partial x_j}(M_0, P_0) = 0, \; j \in N \setminus J, \qquad y_i^0 \frac{\partial L}{\partial y_i}(M_0, P_0) = 0, \; i = 1, 2, \ldots, m,$$

$$x_j^0 \frac{\partial L}{\partial x_j}(M_0, P_0) = 0, \; j \in J.$$

○ Найдем точки Куна—Таккера для задачи

$$f = x_1 x_2 \,(\max), \tag{9.62}$$

$$\frac{x_1^2}{4} + \frac{x_2^2}{9} \leqslant 1, \tag{9.63}$$

$$x_2 \geqslant 0. \tag{9.64}$$

Функция Лагранжа в данном случае имеет вид

$$L(x_1, x_2, y_1) = x_1 x_2 - y_1\left(\frac{x_1^2}{4} + \frac{x_2^2}{9} - 1\right).$$

Так как

$$\frac{\partial L}{\partial x_1} = x_2 - \frac{2x_1 y_1}{4}; \quad \frac{\partial L}{\partial x_2} = x_1 - \frac{2x_2 y_1}{9};$$

$$\frac{\partial L}{\partial y_1} = -\left(\frac{x_1^2}{4} + \frac{x_2^2}{9} - 1\right),$$

то для отыскания точек Куна—Таккера имеем систему

$$x_2-\frac{2x_1y_1}{4}=0, \qquad \frac{x_1^2}{4}+\frac{x_2^2}{9}-1\leqslant 0;$$
$$x_1-\frac{2x_2y_1}{9}\leqslant 0, \qquad y_1\left(\frac{x_1^2}{4}+\frac{x_2^2}{9}-1\right)=0;$$
$$x_2\left(x_1-\frac{2x_2y_1}{9}\right)=0; \quad x_2\geqslant 0,\ y_1\geqslant 0.$$

Решив ее, найдем точки $M_1(x_1;\ 0)$, $-2\leqslant x_1\leqslant 0$, и $M_2(\sqrt{2};\ 3/\sqrt{2})$, которые являются искомыми точками Куна—Таккера. ●

Необходимое условие оптимальности для задачи нелинейного программирования можно сформулировать следующим образом. Пусть M_0—оптимальное решение задачи (9.59)—(9.61), причем функции $f(M)$, $\Phi_i(M)$, $i=1, 2, \ldots, m$, дифференцируемы в этой точке. Если M_0—неособая точка допустимого множества задачи (9.59)—(9.61), то она является точкой Куна—Таккера этой задачи.

Таким образом, если допустимое множество Ω задачи (9.59)—(9.61) не имеет особых точек, а функции $f(M)$, $\Phi_i(M)$, $i=1, 2, \ldots, m$, дифференцируемы во всех точках Ω, то оптимальное решение этой задачи следует искать среди точек Куна—Таккера.

○ Оптимальное решение задачи (9.62)—(9.64) находится среди точек Куна—Таккера этой задачи: $M_1(x_1;\ 0)$, где $-2\leqslant x_1\leqslant 0$, и $M_2(\sqrt{2};\ 3\sqrt{2})$. Так как $f(M_1)=0$, а $f(M_2)=3$, то $M_2(\sqrt{2};\ 3/\sqrt{2})$—оптимальное решение задачи (9.62)—(9.64). ●

Следует отметить, что необходимое условие оптимальности малопригодно для решения большинства задач нелинейного программирования, так как лишь в редких случаях удается найти все точки Куна—Таккера задачи нелинейного программирования.

Пара точек $M_0\in \mathbf{R}_J^n$, $P_0\in(\mathbf{R}^m)^+$ называется *седловой точкой функции Лагранжа*, если

$$L(M,\ P_0)\leqslant L(M_0,\ P_0)\leqslant L(M_0,\ P)$$

для всех $M\in \mathbf{R}_J^n$ и $P\in(\mathbf{R}^m)^+$.

Достаточное условие оптимальности для задачи нелинейного программирования: если $(M_0,\ P_0)$—седловая точка функции Лагранжа для задачи (9.59)—(9.61), то M_0 является оптимальным решением этой задачи.

Это условие в общем случае не является необходимым: функция Лагранжа может не иметь седловых точек, в то время как исходная задача нелинейного программирования обладает оптимальным решением.

При решении задач нелинейного программирования в общем случае сталкиваются с большими трудностями. Точные методы решения найдены лишь для специальных классов задач нелинейного программирования, в основном когда система ограничений состоит только из линейных неравенств. Известны различные методы, позволяющие найти оптимальное решение задачи нелинейного программирования приближенно. Однако эти методы дают достаточно хорошее приближение лишь для задач выпуклого программирования.

9.21. Задачи выпуклого программирования

Задача максимизации (9.59)—(9.61) называется *задачей выпуклого программирования* в том случае, когда все функции $\Phi_i(M)$ являются выпуклыми, а целевая функция $f(M)$—вогнутой на $\mathbf{R}^n$.

Свойства задач выпуклого программирования.

1°. Любая точка условного максимума целевой функции на допустимом множестве является оптимальным решением задачи выпуклого программирования.

2°. Множество оптимальных решений задачи выпуклого программирования является выпуклым.

Говорят, что множество

$$\Omega = \{M \in \mathbf{R}^n_J \mid \Phi_i(M) \leqslant 0,\ i = 1,\ 2,\ \ldots,\ m\}$$

удовлетворяет *условию Слейтера*, если существует точка $\overline{M} \in \Omega$ такая, что для всех нелинейных функций $\Phi_i(M)$, $\Phi_i(\overline{M}) < 0$.

○ Множество $\Omega = \{M\ (x_1;\ x_2;\ x_3) \in \mathbf{R}^3 \mid \Phi_1(M) = x_1^2 + x_2^2 + x_3^2 - 4 \leqslant 0,\ \Phi_2(M) = x_1 + x_2 + x_3 - 3 \leqslant 0\}$ удовлетворяет условию Слейтера, так как если $\overline{M}\ (1,\ 1,\ 1)$, то $\Phi_1(\overline{M}) = -1 < 0$, $\Phi_2(\overline{M}) = 0$. ●

В частности, если допустимое множество задается только линейными неравенствами, то оно всегда удовлетворяет условию Слейтера.

Предположим, что допустимое множество задачи выпуклого программирования (9.59)—(9.61) удовлетворяет условию Слейтера. Тогда имеют место следующие утверждения:

1. Точка $M_0 \in \mathbf{R}^n_J$ является оптимальным решением задачи (9.59)—(9.61) тогда и только тогда, когда существует точка $P_0 \in (\mathbf{R}^m)^+$ такая, что $(M_0,\ P_0)$ будет седловой точкой функции Лагранжа для этой задачи.

На основании этого утверждения построена теория двойственности в выпуклом программировании.

2. Если функции $f(M)$ и $\Phi_i(M)$ $i=1, 2, \ldots, m$, дифференцируемы в точке $M_0 \in R_J^n$, то эта точка является оптимальным решением задачи (9.59)—(9.61) тогда и только тогда, когда она является точкой Куна—Таккера для этой задачи.

С помощью последнего утверждения построен конечный алгоритм для решения задач выпуклого квадратичного программирования. Кроме того, его можно использовать для проверки, является ли данная точка оптимальным решением задачи выпуклого программирования.

○ Рассмотрим задачу выпуклого программирования

$$f = 5x_1 + 2x_2 + 4x_3 - x_1^2 - x_2^2 - x_3^2 \, (\max), \tag{9.65}$$

$$\begin{cases} 9x_1^2 + 6x_3^2 - 15 \leqslant 0, \\ x_1 + 2x_2 + x_3 - 5 \leqslant 0, \end{cases} \tag{9.66}$$

$$x_1 \geqslant 0, \ x_3 \geqslant 0 \tag{9.67}$$

и точку $M_0(1;\ 1;\ 1)$.

Функция Лагранжа для задачи (9.65)—(9.67) имеет вид

$$L(x_1, x_2, x_3, y_1, y_2) = 5x_1 + 2x_2 + 4x_3 - x_1^2 - x_2^2 - x_3^2 - \\ - y_1(9x_1^2 + 6x_3^2 - 15) - y_2(x_1 + 2x_2 + x_3 - 5).$$

Запишем условия Куна—Таккера:

$$\begin{array}{ll} 5 - 2x_1 - 18y_1x_1 - y_2 \leqslant 0, & 9x_1^2 + 6x_3^2 - 15 \leqslant 0, \\ 2 - 2x_2 - 2y_2 = 0, & x_1 + 2x_2 + x_3 - 5 \leqslant 0, \\ 4 - 2x_3 - 12y_1x_3 - y_2 \leqslant 0, & (9x_1^2 + 6x_3^2 - 15)\, y_1 = 0, \\ (5 - 2x_1 - 18y_1x_1 - y_2)\, x_1 = 0, & (x_1 + 2x_2 + x_3 - 5)\, y_2 = 0. \\ (4 - 2x_3 - 12y_1x_3 - y_2)\, x_3 = 0, & \end{array}$$

Полагая в этой системе $x_1 = 1$, $x_2 = 1$, $x_3 = 1$, найдем $y_1 = 1/6$, $y_2 = 0$. Поэтому $M_0(1;\ 1;\ 1)$ является точкой Куна—Таккера для задачи (9.65)—(9.67). Следовательно, $M_0(1;\ 1;\ 1)$—оптимальное решение этой задачи. ●

9.22. Задачи выпуклого квадратичного программирования

Задача максимизации квадратичной функции

$$f = \sum_{j=1}^{n} \gamma_j x_j + \sum_{k=1}^{n} c_{kk} x_k^2 + 2 \sum_{1 \leqslant k < l \leqslant n} c_{kl} x_k x_l \tag{9.68}$$

на множестве решений системы

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant b_i, \quad i=1, 2, \ldots, m; \tag{9.69}$$

$$x_j \geqslant 0, \quad j=1, 2, \ldots, n, \tag{9.70}$$

называется *задачей квадратичного программирования.* Если квадратичная форма

$$\sum_{k=1}^{n} c_{kk}x_k^2 + 2\sum_{1 \leqslant k < l \leqslant n} c_{kl}x_kx_l$$

является отрицательно-определенной, то квадратичная функция (9.68) является вогнутой. В этом случае задача (9.68)—(9.70) называется *задачей выпуклого квадратичного программирования.*

Точка $M_0(x_1^0; \ldots; x_j^0; \ldots; x_n^0)$ является оптимальным решением задачи выпуклого квадратичного программирования (9.68)—(9.70) тогда и только тогда, когда существуют числа y_i^0, v_i^0, $i=1, 2, \ldots, m$, и u_j^0, $j=1, 2, \ldots, n$, такие, что:

1) $x_j^0u_j^0=0, \quad j=1, 2, \ldots, n;$ (9.71)

2) $v_i^0y_i^0=0, \quad i=1, 2, \ldots, m;$ (9.72)

3) вектор $\alpha^0=(x_1^0; \ldots; x_n^0; y_1^0; \ldots; y_m^0; u_1^0; \ldots; u_n^0; v_1^0; \ldots; v_m^0)$ является опорным решением системы ограничений

$$\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} a_{ij}y_i + u_j = 0, \quad j=1, 2, \ldots, n, \tag{9.73}$$

$$\sum_{j=1}^{n} a_{ij}x_j + v_i = b_i, \quad i=1, 2, \ldots, m,$$

$$x_j \geqslant 0, \; u_j \geqslant 0, \; v_i \geqslant 0, \quad j=1, 2, \ldots, n, \; i=1, 2, \ldots, m.$$

Два вектора условий для системы (9.73) назовем сопряженными, если они либо соответствуют неизвестным x_j, u_j при некотором j, либо неизвестным y_i, v_i при некотором i.

Таким образом, для решения задачи выпуклого квадратичного программирования (9.68)—(9.70) достаточно найти опорное решение системы (9.73), базис которого не содержит сопряженных векторов условий. Такое опорное решение можно найти методом искусственного базиса.

○ Рассмотрим задачу выпуклого квадратичного программирования

$$f = x_1 + 10x_2 - x_1^2 - x_2^2 \text{ (max)},$$
$$x_1 + x_2 \leqslant 4,$$
$$x_1 \geqslant 0, \ x_2 \geqslant 0.$$

Так как $\frac{\partial f}{\partial x_1} = 1 - 2x_1$, $\frac{\partial f}{\partial x_2} = 10 - 2x_2$, то в данном случае система (9.73) имеет вид

$$\begin{cases} 1 - 2x_1 - y_1 + u_1 = 0, \\ 10 - 2x_2 - y_1 + u_2 = 0, \\ x_1 + x_2 + v_1 = 4, \\ x_1 \geqslant 0, \ x_2 \geqslant 0, \ y_1 \geqslant 0, \\ u_1 \geqslant 0, \ u_2 \geqslant 0, \ v_1 \geqslant 0. \end{cases}$$

Составим вспомогательную задачу:

$$\begin{cases} 2x_1 + y_1 - u_1 + z_1 = 1, \\ 2x_2 + y_1 - u_2 + z_2 = 10, \\ x_1 + x_2 + v_1 + z_3 = 4, \\ x_1 \geqslant 0, \ x_2 \geqslant 0, \ y_1 \geqslant 0, \\ u_1 \geqslant 0, \ u_2 \geqslant 0, \ v_1 \geqslant 0, \\ z_1 \geqslant 0, \ z_2 \geqslant 0, \ z_3 \geqslant 0; \end{cases}$$
$$\varphi = z_1 + z_2 + z_3 \text{ (min)}.$$

Решая эту задачу симплекс-методом, необходимо следить за тем, чтобы на каждом шаге базис опорного решения не содержал сопряженных векторов условий. ●

9.23. Приближенные методы решения задач нелинейного программирования

Предположим, что задача максимизации функции

$$f = f(M) \tag{9.74}$$

на множестве

$$\Omega = \{M \in \mathbf{R}^n \mid \Phi_i(M) \leqslant 0, \quad i = 1, 2, \ldots, m\} \tag{9.75}$$

имеем оптимальное решение M_0.

Пусть ε — некоторое положительное число. Точка $M_\varepsilon \in \Omega$ называется *приближенным решением* задачи (9.74)—(9.75) с точностью ε, если она удовлетворяет неравенству

$$f(M_0)-f(M_\varepsilon)<\varepsilon.$$

Обычно приближенное решение задачи нелинейного программирования (9.74)—(9.75) ищут с помощью *релаксационного процесса*. Релаксационный процесс—это процесс построения последовательных приближений $M_1, M_2, \ldots, M_k, \ldots$ таких, что $M_k \in \Omega$ $(k=1, 2, \ldots)$ и $f(M_{k+1}) > f(M_k)$. При этом релаксационный процесс называется *сходящимся*, если

$$\lim_{k\to\infty} f(M_k)=f(M_0).$$

Сходящийся релаксационный процесс позволяет найти приближенное решение задачи нелинейного программирования с любой заданной точностью.

Чтобы задать релаксационный процесс решения задачи нелинейного программирования, достаточно указать точку $M_1 \in \Omega$ и правило перехода от точки M_k к точке M_{k+1}.

В большинстве случаев переход от точки $M_k(x_1^{(k)}; x_2^{(k)}; \ldots; x_n^{(k)})$ к точке M_{k+1} осуществляют по следующей схеме:

1. Выбирают некоторый вектор $\alpha_k=(a_1^{(k)}; a_2^{(k)}; \ldots; a_n^{(k)})$ (направление перемещения из точки M_k) и рассматривают точки вида

$$M_{k,t}(x_1^{(k)}+a_1^{(k)}t;\ x_2^{(k)}+a_2^{(k)}t;\ \ldots;\ x_n^{(k)}+a_n^{(k)}t),$$

где t—некоторый положительный параметр.

2. Подбирают значение параметра t так, чтобы

$$M_{k,t} \in \Omega \text{ и } f(M_{k,t})>f(M_k).$$

3. Если t_k—найденное значение параметра t, то полагают $M_{k+1}=M_{k,tk}$.

Известны различные методы выбора направления перемещения и значения параметра t, которые для задач выпуклого программирования гарантируют сходимость релаксационного процесса. К таким методам, например, относится метод возможных направлений (9.24).

Для решения задачи максимизации вогнутой дифференцируемой функции $f(M)$ на всем пространстве $\mathbf{R}^n$ можно использовать градиентный метод, в котором точку $M_1 \in \mathbf{R}^n$ выбирают произвольным образом, а направление перемещения из точки M_k $(k=1, 2, \ldots)$ считают равным

$$\operatorname{grad} f|_{M_k}=\left\{\frac{\partial f}{\partial x_1}(M_k);\ \frac{\partial f}{\partial x_2}(M_k);\ \ldots;\ \frac{\partial f}{\partial x_n}(M_k)\right\}.$$

При этом если $\operatorname{grad} f|_{M_k} = \theta$, то M_k — оптимальное решение задачи.

Одной из разновидностей градиентного метода является метод скорейшего подъема. В методе скорейшего подъема на k-м шаге значения параметра t подбирают так, чтобы достигалось наибольшее значение функции $f(M_{k,t})$ одной переменной t.

Точку M_ε называют *обобщенным приближенным* решением задачи (9.74)—(9.75) с точностью ε, если она удовлетворяет следующим неравенствам:

$$f(M_0) - f(M_\varepsilon) < \varepsilon,$$
$$\Phi_i(M_\varepsilon) \leqslant \varepsilon, \quad i = 1, 2, \ldots, m.$$

Заметим, что обобщенное приближенное решение задачи нелинейного программирования не обязано принадлежать допустимому множеству этой задачи. Отыскание обобщенного приближенного решения оправдано в тех случаях, когда, например, система ограничений задачи нелинейного программирования сама задана приближенно.

Для отыскания обобщенных приближенных решений используют *метод штрафных функций*.

Если функция $\xi(t)$ удовлетворяет следующим условиям:

1) $\xi(t) = 0$ при $t \leqslant 0$;
2) $\xi(t) > 0$ при $t > 0$;
3) $\xi(t_2) > \xi(t_1)$ при $t_2 > t_1 > 0$;
4) $\lim\limits_{t \to +\infty} \xi(t) = \infty$,

— то функция

$$\Lambda(M, R) = \sum_{i=1}^{m} r_i \xi(\Phi_i(M)),$$

где $R = (r_1, r_2, \ldots, r_m)$ — набор положительных параметров, называется штрафной функцией задачи (9.74)—(9.75). Основное свойство штрафной функции:

$$\Lambda(M, R) = 0 \text{ при } M \in \Omega \text{ и } \Lambda(M, R) > 0 \text{ при } M \notin \Omega.$$

Метод штрафных функций состоит в том, что вместо задачи нелинейного программирования (9.74)—(9.75) решают задачи максимизации на всем пространстве R^n функции

$$\tilde{f}_R(M) = f(M) - \Lambda(M, R)$$

при различных параметрах $r_1, r_2, \ldots, r_m$.

При благоприятных условиях метод штрафных функций позволяет найти обобщенное приближенное решение задачи нелинейного программирования с заданной точностью.

9.24. Метод возможных направлений для решения задач выпуклого программирования

Метод возможных направлений — это способ задания релаксационного процесса решения задачи выпуклого программирования.

Рассмотрим задачу максимизации функции $f(M)$ на множестве

$$\Omega = \{M \in \mathbf{R}^n \mid \Phi_i(M) \leqslant 0,\ i \in I\}.$$

Предположим, что выполнены следующие условия:

I. $I = I_л \cup I_в$, и если $i \in I_л$, то $\Phi_i(M)$ — линейная функция, если $i \in I_в$, то $\Phi_i(M)$ — выпуклая дифференцируемая функция на пространстве $\mathbf{R}^n$.

II. Множество Ω удовлетворяет условию Слейтера, т. е. существует точка $\overline{M} \in \Omega$ такая, что $\Phi_i(\overline{M}) < 0$ для всех $i \in I_в$.

III. $f(M)$ — вогнутая дифференцируемая функция на допустимом множестве Ω.

В методе возможных направлений точку $M_1 \in \Omega$ выбирают произвольным образом. Переход от точки $M_k(x_1^{(k)};\ x_2^{(k)},\ \ldots;\ x_n^{(k)})$ к точке M_{k+1} осуществляют следующим образом $(k = 1,\ 2,\ \ldots)$. Направление перемещения α_k из точки M_k ищут так, чтобы:

$$\begin{cases} \operatorname{grad} f\,|_{M_k} \cdot \alpha_k > 0, \\ \operatorname{grad} \Phi_i\,|_{M_k} \cdot \alpha_k \leqslant 0, & i \in I(M_k) \cap I_л, \\ \operatorname{grad} \Phi_i\,|_{M_k} \cdot \alpha_k < 0, & i \in I(M_k) \cap I_в, \end{cases} \tag{9.76}$$

где $I(M_k) = \{i \in I \mid \Phi_i(M_k) = 0\}$.

Заметим, что если вектора, удовлетворяющего условиям (9.76), не существует, то M_k — оптимальное решение задачи максимизации.

Пусть $\alpha_k = (a_1^{(k)};\ a_2^{(k)};\ \ldots;\ a_n^{(k)})$ удовлетворяет условиям (9.76). Тогда существует число $T > 0$ такое, что при всех t, $0 < t < T$,

$$f(M_{k,t}) > f(M_k) \text{ и } M_{k,t} \in \Omega, \tag{9.77}$$

где $M_{k,t}(x_1^{(k)} + a_1^{(k)}t;\ x_2^{(k)} + a_2^{(k)}t;\ \ldots;\ x_n^{(k)} + a_n^{(k)}t)$.

После того как определено направление перемещения α_k, необходимо подобрать значение $t > 0$ так, чтобы выполнялись условия (9.77). Это можно сделать, перебрав несколько достаточно малых значений t. Если t_k — найденное значение параметра t, то полагают

$$M_{k+1}(x_1^{(k)}+a_1^{(k)}t_k;\ x_2^{(k)}+a_2^{(k)}t_k;\ \ldots;\ x_n^{(k)}+a_n^{(k)}t_k).$$

Известны различные приемы, уточняющие выбор направления перемещения и параметра t на каждом шаге, которые обеспечивают сходимость релаксационной последовательности $M_1, M_2, \ldots, M_k, \ldots$ к оптимальному решению задачи.

○ Рассмотрим задачу максимизации функции

$$f(M)=13x_1+3x_2+17x_3-x_1^2-x_2^2-x_3^2$$

на множестве

$$\Omega=\{M(x_1;\ x_2;\ x_3)\,|\,\Phi_1(M)=2x_1^2+3x_2^2+x_3^2-24\leqslant 0,$$
$$\Phi_2(M)=x_1+2x_2+x_3-6\leqslant 0,\ \Phi_3(M)=-x_2\leqslant 0\}.$$

Эта задача удовлетворяет условиям (I)—(III). Заметим, что

$$\begin{aligned}&\operatorname{grad} f=\{13-2x_1;\ 3-2x_2;\ 17-2x_3\},\\ &\operatorname{grad}\Phi_1=\{4x_1;\ 6x_2;\ 2x_3\},\\ &\operatorname{grad}\Phi_2=\{1;\ 2;\ 1\},\ \operatorname{grad}\Phi_3=\{0;\ -1;\ 0\}.\end{aligned}$$

Рассмотрим точку $M_1(1;\ 1;\ 1)\in\Omega$. Тогда

$$f(M_1)=30;\ \Phi_1(M_1)=-18;\ \Phi_2(M_1)=-2;$$
$$\Phi_3(M_1)=-1;\ \operatorname{grad} f|_{M_1}=\{11;\ 1;\ 15\}.$$

Направление перемещения $\alpha_1=(a_1^{(1)};\ a_2^{(1)};\ a_3^{(1)})$ определим из условия $11a_1^{(1)}+a_2^{(1)}+15a_3^{(1)}>0$ ($\operatorname{grad} f|_{M_1}\cdot\alpha_1>0$). Положим $\alpha_1=(2;\ 1;\ 2)$. Тогда $M_{1,t}(1+2t;\ 1+t;\ 1+2t)$. При $t=1/3$ имеем $M_2(5/3,\ 4/3,\ 5/3)$, причем

$$f(M_2)=46^2/_3;\ \Phi_1(M_2)=-10^1/_3;\ \Phi_2(M_2)=0;$$
$$\Phi_3(M_2)=-4/3;\quad \operatorname{grad} f|_{M_2}=\{29/3;\ 1/3;\ 41/3\}.$$

Направление перемещения $\alpha_2=(a_1^{(2)};\ a_2^{(2)};\ a_3^{(2)})$ определим из следующих условий:

$$\begin{cases}{}^{29}/_3a_1^{(2)}+{}^1/_3a_2^{(2)}+{}^{41}/_3a_3^{(2)}>0\ (\operatorname{grad} f|_{M_2}\cdot\alpha_2>0),\\ a_1^{(2)}+2a_2^{(2)}+a_3^{(2)}\leqslant 0\quad(\operatorname{grad}\Phi_2|_{M_2}\cdot\alpha_2\leqslant 0).\end{cases}$$

Положим $\alpha_2=(1;\ -2;\ 1)$. Тогда $M_{2,t}({}^5/_3+t;\ {}^4/_3-2t;\ {}^5/_3+t)$. Если $t=2/3$, то $M_3(7/3,\ 0,\ 7/3)$ и

$$f(M_3)=59^1/_9;\ \Phi_1(M_3)=-7^2/_3;\ \Phi_2(M_3)=-{}^4/_3;$$
$$\Phi_3(M_3)=0;\ \operatorname{grad} f|_M=\{{}^{25}/_3;\ 3;\ {}^{37}/_3\}.$$

Направление $\alpha_3=(a_1^{(3)};\ a_2^{(3)};\ a_3^{(3)})$ определим из условий

$$\begin{cases} {}^{25}/_3 a_1^{(3)}+3a_2^{(3)}+{}^{37}/_3 a_3^{(3)}>0,\\ a_2^{(3)}\geqslant 0.\end{cases}$$

Положим $\alpha_3=(-1;\ 0;\ 1)$. В этом случае $M_{3,t}({}^7/_3-t;\ 0;\ {}^7/_3+t)$. Если $t=1/3$, то $M_4(2;\ 0;\ 4)$. Тогда

$$f(M_4)=74;\ \Phi_1(M_4)=0;\ \Phi_2(M_4)=0;$$
$$\Phi_3(M_4)=0;\ \operatorname{grad} f|_{M_4}=\{9;\ 3;\ 9\}.$$

Направление $\alpha_4=(a_1^{(4)};\ a_2^{(4)};\ a_3^{(4)})$ определим из условий:

$$\begin{cases} 9a_1^{(4)}+3a_2^{(4)}+9a_3^{(4)}>0,\\ 8a_1^{(4)}+0a_2^{(4)}+8a_3^{(4)}<0\ \ (\operatorname{grad}\Phi_1|_{M_4}\alpha_4<0),\\ a_1^{(4)}+2a_2^{(4)}+a_3^{(4)}\leqslant 0,\\ -a_2^{(4)}\leqslant 0.\end{cases}$$

Если сложить 3 и 4-е неравенства, умноженные на подходящие числа, то получим

$$9a_1^{(4)}+3a_2^{(4)}+9a_3^{(4)}\leqslant 0,$$

что противоречит 1-му неравенству. Допустимого направления не существует. Значит, $M_4(2;\ 0;\ 4)$—оптимальное решение данной задачи.

9.25. Простейшие задачи вариационного исчисления

Последовательность n функций $x_1(t),\ x_2(t),\ \ldots,\ x_n(t)$, определенных на некотором отрезке $[t_0,\ T]$, называется вектор-функцией $\bar{x}(t)$ на этом отрезке, т. е.

$$\bar{x}(t)=\{x_1(t);\ x_2(t);\ \ldots;\ x_n(t)\}. \tag{9.78}$$

Вектор-функция (9.78) называется *дифференцируемой, непрерывной* или *кусочно-непрерывной* на отрезке $[t_0,\ T]$, если все функции $x_j(t)$ соответственно дифференцируемы, непрерывны или кусочно-непрерывны на этом отрезке. Если вектор-функция $\bar{x}(t)$ дифференцируема на $[t_0,\ T]$, то

$$\frac{d\bar{x}(t)}{dt}=\left\{\frac{dx_1(t)}{dt};\ \frac{dx_2(t)}{dt};\ \ldots;\ \frac{dx_n(t)}{dt}\right\},\quad t\in[t_0,\ T].$$

Оптимизационная задача $(V,\ f,\ \Omega)$ называется *простейшей задачей вариационного исчисления в форме Лагранжа*, если:

1) V — множество вектор-функций, дифференцируемых на отрезке $[t_0, T]$ (t_0, T — фиксированные числа);

2) целевая функция f имеет вид

$$f[\bar{x}(t)] = \int_{t_0}^{T} F\left(\bar{x}, \frac{d\bar{x}}{dt}, t\right) dt, \quad \bar{x}(t) \in V; \tag{9.79}$$

3) допустимое множество Ω состоит из вектор-функций $\bar{x}(t) = \{x_1(t); x_2(t); \ldots; x_n(t)\}$, удовлетворяющих начальным условиям

$$\bar{x}(t_0) = \bar{x}^0 \tag{9.80}$$

и конечным условиям вида

$$x_i(T) = x_i^T, \; i \in I, \; I \subseteq N = \{1, 2, \ldots, n\}. \tag{9.81}$$

Если $I = N$, то конечные условия (9.81) можно записать в виде

$$\bar{x}(T) = \bar{x}^T, \text{ где } \bar{x}^T = \{x_1^T; x_2^T; \ldots; x_n^T\}.$$

Задаче вариационного исчисления в форме Лагранжа можно придать следующий экономический смысл.

Рассмотрим некоторую развивающуюся экономическую систему. Предположим, что траектория развития системы может быть описана вектор-функцией

$$\bar{x}(t) = \{x_1(t); x_2(t); \ldots; x_n(t)\},$$

где $\bar{x}(t)$ — состояние системы в момент времени t, $t_0 \leqslant t \leqslant T$.

В этом случае функционал (9.79) можно интерпретировать, например, как затраты на траекторию развития $\bar{x}(t)$ в течение времени от t_0 до T. Тогда задача минимизации (9.79)—(9.81)—это задача выбора оптимальной траектории развития экономической системы, т. е. траектории, удовлетворяющей заданным начальным и конечным условиям (9.80)—(9.81), при которой затраты будут наименьшими.

Если $\bar{x}(t) = \{x_1(t); x_2(t); \ldots; x_n(t)\}$ — дифференцируемая вектор-функция, то положим

$$\bar{u}(t) = \frac{d\bar{x}(t)}{dt},$$

т. е. $\bar{u}(t) = \{u_1(t); u_2(t); \ldots; u_n(t)\}$, где $u_j(t) = \dfrac{dx_j(t)}{dt}$, $j = 1, 2, \ldots, n$.

Задачу минимизации (9.79)—(9.81) можно записать в виде

$$f(\bar{x}) = \int_{t_0}^{T} F(\bar{x},\ \bar{u},\ t)\,\mathrm{d}t \ (\min), \tag{9.82}$$

где

$$\bar{u}(t) = \frac{\mathrm{d}\bar{x}(t)}{\mathrm{d}t}, \tag{9.83}$$

при условии, что

$$\bar{x}(t_0) = \bar{x}^0, \tag{9.84}$$

$$x_i(T) = x_i^T,\ i \in I,\ I \subseteq N = \{1,\ 2,\ \ldots,\ n\}. \tag{9.85}$$

Функция

$$H(\bar{\psi},\ \bar{x},\ \bar{u},\ t) = \sum_{j=1}^{n} \psi_j(t)\cdot u_j(t) - F(\bar{x},\ \bar{u},\ t),$$

где $\bar{\psi}(t) = \{\psi_1(t);\ \psi_2(t);\ \ldots;\ \psi_n(t)\}$ — некоторая вектор-функция на отрезке $[t_0,\ T]$, называется *функцией Гамильтона* для задачи (9.82)—(9.85).

Для задачи (9.82)—(9.85) имеет место следующее необходимое условие оптимальности. Пусть функция $F(\bar{x},\ \bar{u},\ t)$ непрерывно дифференцируема, как функция своих аргументов. Если непрерывно дифференцируемая вектор-функция $\bar{x}^*(t)$ является оптимальным решением задачи (9.82)—(9.85), то существует непрерывная вектор-функция $\bar{\psi}^*(t) = \{\psi_1^*(t);\ \psi_2^*(t);\ \ldots;\ \psi_n^*(t)\}$ такая, что пара $\bar{x}^*(t)$, $\bar{\psi}^*(t)$ удовлетворяет следующим условиям:

1) $\dfrac{\mathrm{d}\psi_j}{\mathrm{d}t} = \dfrac{\partial F}{\partial x_j},\ j = 1,\ 2,\ \ldots,\ n;$

2) $\dfrac{\partial H}{\partial u_j} = 0,\ j = 1,\ 2,\ \ldots,\ n;$

3) $\psi_i(T) = 0,\ i \in N \setminus I,$

где $H(\bar{\psi},\ \bar{x},\ \bar{u},\ t)$ — функция Гамильтона.

Приведенное необходимое условие оптимальности не является достаточным.

Для того чтобы $\bar{x}^*(t)$ было оптимальным решением задачи (9.82)—(9.85), достаточно, чтобы выполнялось неравенство

$$f[\bar{x}^*(t) + \bar{h}(t)] - f[\bar{x}^*(t)] \geqslant 0 \tag{9.86}$$

при любой дифференцируемой вектор-функции $\bar{h}(t) = \{h_1(t);\ h_2(t);\ \ldots;\ h_n(t)\}$ такой, что $h_j(t_0) = 0,\ j = 1,\ 2,\ \ldots,\ n,\ h_i(T) = 0,\ i \in I.$

○ Рассмотрим задачу минимизации

$$\int_0^1 (u_1^2 + u_2^2 + 4x_2)\,dt,$$

где $u_1 = \frac{dx_1}{dt}$, $u_2 = \frac{dx_2}{dt}$, при условиях, что

$$x_1(0) = x_2(0) = 0, \quad x_1(1) = 1.$$

В данном случае функция Гамильтона имеет вид

$$H = \psi_1 u_1 + \psi_2 u_2 - (u_1^2 + u_2^2 + 4x_2).$$

Необходимое условие оптимальности дает следующие соотношения:

1) $\frac{d\psi_1}{dt} = 0$, $\frac{d\psi_2}{dt} = 4$;

2) $\psi_1 - 2u_1 = 0$, $\psi_2 - 2u_2 = 0$;

3) $\psi_2(1) = 0$.

Из этих условий получаем: $\psi_1 = c_1$, $\psi_2 = 4t - 4$, $u_1 = \frac{c_1}{2}$, $u_2 = 2t - 2$. Так как

$$\frac{dx_1}{dt} = u_1, \quad \frac{dx_2}{dt} = u_2,$$

то $x_1 = \frac{c_1}{2}t + c_2$, $x_2 = t^2 - 2t + c_3$. Из условий $x_1(0) = x_2(0) = 0$, $x_1(1) = 0$ найдем $x_1 = t$, $x_2 = t^2 - 2t$. С помощью достаточного условия (9.86) можно проверить, что вектор-функция $\bar{x}^*(t) = \{t;\ t^2 - 2t\}$ является оптимальным решением нашей задачи. ●

9.26. Задачи оптимального управления

Оптимизационная задача (V, f, Ω) называется *задачей оптимального управления*, если:

1) V — множество пар вектор-функций $\bar{x}(t)$, $\bar{u}(t)$, где функция $\bar{x}(t) = \{x_1(t);\ x_2(t);\ \ldots;\ x_n(t)\}$ дифференцируема, а функция $\bar{u}(t) = \{u_1(t);\ u_2(t);\ \ldots;\ u_m(t)\}$ кусочно-непрерывна на отрезке $[t_0, T]$ (t_0, T фиксированы);

2) целевая функция f имеет вид

$$f[\bar{x}(t), \bar{u}(t)] = \int_{t_0}^{T} F(\bar{x}, \bar{u}, t)\,dt; \qquad (9.87)$$

3) допустимое множество $\Omega \subseteq V$ удовлетворяет следующим условиям:

а) $\frac{dx_j}{dt} = f_j[\bar{x}, \bar{u}, t]$, $j = 1, 2, \ldots, n$, (9.88)

б) при любом t, $t_0 \leqslant t \leqslant T$,

$$\bar{u}(t) \in U, \tag{9.89}$$

где U — фиксированное замкнутое подмножество пространства $\mathbf{R}^m$;

в) $\bar{x}(t_0) = \bar{x}^0$, (9.90)

$$x_i(T) = x_i^T,\ i \in I,\ I \subseteq N = \{1, 2, \ldots, n\}. \tag{9.91}$$

Например, задача вариационного исчисления (9.82) — (9.85) является задачей оптимального управления.

Экономический смысл задачи оптимального управления. Рассмотрим некоторую экономическую систему. Предположим, что в каждый момент времени t, $t_0 \leqslant t \leqslant T$, на эту систему можно оказывать управляющие воздействия

$$u_1(t),\ u_2(t),\ \ldots,\ u_m(t),$$

которые выбираются из некоторого допустимого множества U.

Вектор-функция $\bar{u}(t) = \{u_1(t); u_2(t); \ldots; u_m(t)\}$, $t_0 \leqslant t \leqslant T$, называется *управлением системой*.

Если скорость изменения состояния системы в каждый момент времени t зависит от самого состояния $\bar{x}(t) = \{x_1(t); x_2(t); \ldots; x_n(t)\}$, от управления $\bar{u}(t)$ и от момента времени t, то

$$\frac{dx_j}{dt} = f_j[\bar{x}, \bar{u}, t],\ j = 1, 2, \ldots, n.$$

При заданном управлении $\bar{u}(t)$ решение этой системы дифференциальных уравнений является некоторой траекторией развития экономической системы. Траектория развития системы должна удовлетворять некоторым начальным и конечным условиям.

Если $\bar{u}(t)$ — управление, а $\bar{x}(t)$ — определяемая им траектория развития, то пара $\{\bar{x}(t), \bar{u}(t)\}$ называется управляемым процессом. Будем считать, что затраты $f(\bar{x}, \bar{u})$ на управляемый процесс $(\bar{x}, \bar{u})$ можно вычислить по формуле

$$f(\bar{x}, \bar{u}) = \int_{t_0}^{T} F[\bar{x}, \bar{u}, t]\,dt.$$

Таким образом, задача оптимального управления (9.87)—(9.91)—это задача выбора оптимального управляемого процесса, т. е. управляемого процесса, удовлетворяющего всем приведенным условиям, при котором затраты будут наименьшими.

Функция

$$H(\overline{\psi}, \overline{x}, \overline{u}, t) = \sum_{j=1}^{n} \psi_j(t) \cdot f_j(\overline{x}, \overline{u}, t) - F(\overline{x}, \overline{u}, t),$$

где $\overline{\psi}(t) = \{\psi_1(t); \psi_2(t); \ldots; \psi_n(t)\}$ — некоторая вектор-функция на $[t_0, T]$, называется *функцией Гамильтона для задачи* (9.87)—(9.90) (если отсутствует условие 9.91).

Необходимое условие оптимальности для задачи оптимального управления (принцип максимума Понтрягина). Пусть функции $F(\overline{x}, \overline{u}, t)$, $f_j(\overline{x}, \overline{u}, t)$, $j = 1, 2, \ldots, n$ непрерывно дифференцируемы. Если $[\overline{x}^*(t), \overline{u}^*(t)]$ — оптимальное решение задачи минимизации (9.87)—(9.90), то существует непрерывная вектор-функция $\overline{\psi}^*(t) = \{\psi_1^*(t); \psi_2^*(t); \ldots; \psi_n^*(t)\}$ такая, что функции $\overline{x}^*(t)$, $\overline{u}^*(t)$, $\overline{\psi}^*(t)$ удовлетворяют следующим условиям:

1) $\dfrac{d\psi_j}{dt} = -\dfrac{\partial H}{\partial x_j}$, $j = 1, 2, \ldots, n$;

2) $\psi_i^*(T) = 0$, $i \in I$; $i = 1, 2, \ldots, n$;

3) при каждом $t \in [t_0, T]$ функция $H(\overline{x}^*(t), \overline{u}, \overline{\psi}^*(t), t)$ достигает при $\overline{u} = \overline{u}^*(t)$ максимума по всем $\overline{u} \in U$. $H(\overline{x}, \overline{u}, \overline{\psi}, t)$ — функция Гамильтона для задачи (9.87) — (9.90).

○ Рассмотрим задачу минимизации $\int\limits_0^1 (u^2 - 4x)\,dt$, где $\dfrac{dx}{dt} = u$, $0 \leqslant u \leqslant 1$, $x(0) = 0$.

Функция Гамильтона в данном случае имеет вид

$$H = \psi u - u^2 + 4x.$$

Из условий 1) и 2) находим, что $\psi = 4 - 4t$, $t \in [0, 1]$. Тогда

$$H = (4 - 4t)u - u^2 + 4x.$$

Максимум этой функции по u, $0 \leqslant u \leqslant 1$, достигается при $u = u^*(t)$, где

$$u^*(t) = \begin{cases} 1, \text{ если } 0 \leqslant t \leqslant 1/2; \\ 2 - 2t, \text{ если } 1/2 \leqslant t \leqslant 1. \end{cases}$$

По условию, $\frac{dx^*(t)}{dt} = u^*(t)$. Значит,

$$x^*(t) = \begin{cases} t + c_1, & \text{если} \quad 0 \leqslant t \leqslant 1/2; \\ 2t - t^2 - c_2, & \text{если} \quad 1/2 \leqslant t \leqslant 1. \end{cases}$$

Так как $x^*(0) = 0$, то $c_1 = 0$.

Из условия непрерывности $x^*(t)$ в точке $t = 1/2$ найдем постоянную $c_2 = -1/4$. Таким образом,

$$x^*(t) = \begin{cases} t, & \text{если} \quad 0 \leqslant t \leqslant 1/2, \\ 2t - t^2 - 1/4, & \text{если} \quad 1/2 \leqslant t \leqslant 1. \end{cases}$$

Можно проверить, что найденные $x^*(t)$ и $u^*(t)$ составляют оптимальное решение данной задачи. ●

РАЗДЕЛ X. ТЕОРИЯ ИГР

10.1. Постановка общей задачи теории игр

Теория игр — математическая теория конфликтных ситуаций. Экономические соревнования, спортивные встречи, боевые операции — примеры конфликтных ситуаций. Простейшие модели конфликтных ситуаций — это салонные и спортивные игры.

В игре могут сталкиваться интересы двух противников (игра парная или игра двух лиц), интересы n ($n>2$) противников (игра множественная или игра n лиц). Существуют игры с бесконечным множеством игроков.

Стратегией игрока называется система правил, однозначно определяющих выбор поведения игрока на каждом ходе в зависимости от ситуации, сложившейся в процессе игры. В зависимости от числа возможных стратегий игры делятся на *конечные* и *бесконечные*.

Процесс игры состоит в выборе каждым игроком i одной своей стратегии $s_i \in S_i$. В результате сложившейся ситуации s игрок i получает выигрыш $H_i(s)$.

Игры, в которых целью каждого участника является получение по возможности большего индивидуального выигрыша, называются *бескоалиционными* в отличие от *коалиционных*, в которых действия игроков направлены на максимизацию выигрышей коллективов (коалиции) без дальнейшего разделения выигрыша между участниками.

По определению бескоалиционной игрой называется система

$$\Gamma = \langle I,\ \{S_i\}_{i\in I},\ \{H_i\}_{i\in I}\rangle,$$

в которой I и S_i ($i \in I$) — множества, H_i — функции на множестве $S=\prod_{i\in I} S_i$, принимающие вещественные значения.

Бескоалиционная игра называется *игрой с постоянной суммой*, если существует такое постоянное C, что $\sum_{i\in I} H_i(s)=C$ для всех ситуаций $s \in S$.

Ситуация s в игре называется *приемлемой* для игрока i, если этот игрок, изменяя в ситуации s свою стратегию s_i на какую-либо другую s_i', не может увеличить своего выигрыша.

Ситуация s, приемлемая для всех игроков, называется *ситуацией равновесия*.

Процесс нахождения ситуации равновесия в бескоалиционной игре есть процесс решения игры.

10.2. Матричные игры

Игра называется *парной*, если в ней сталкиваются интересы двух противников. Игра называется с *нулевой суммой*, если один игрок выигрывает столько, сколько второй проигрывает в той же партии.

Каждая фиксированная стратегия, которую может выбрать игрок, называется его *чистой стратегией*.

Матричной называют парную игру с нулевой суммой при условии, что каждый игрок имеет конечное число чистых стратегий.

Пусть первый игрок имеет m чистых стратегий, а второй — n.

Парная игра с нулевой суммой задается формально системой чисел — матрицей $\|a_{ij}\|_{m\times n}$, элементы которой определяют выигрыш первого игрока (и соответственно проигрыш второго), если первый игрок выберет i-ю строку (i-ю стратегию), а второй игрок — j-й столбец (j-ю стратегию). Матрица $\|a_{ij}\|$ называется *платежной матрицей* или *матрицей игры*.

Задача первого игрока — максимизировать свой выигрыш.

Задача второго игрока — максимизировать свой выигрыш — сводится к минимизации проигрыша второго, что эквивалентно задаче минимизации выигрыша первого игрока.

10.3. Чистые стратегии

Гарантированный выигрыш первого игрока, применяющего чистую i-ю стратегию,

$$\alpha_i = \min_{1 \leqslant j \leqslant n} a_{ij}.$$

Число $\underline{v} = \max\limits_{1 \leqslant i \leqslant m} \alpha_i = \max\limits_{1 \leqslant i \leqslant m} \left\{ \min\limits_{1 \leqslant j \leqslant n} a_{ij} \right\}$ называется *нижним значением игры*, а соответствующая чистая стратегия i_0, при которой достигается $\underline{v}$, называется *максиминной* *стратегией* первого игрока. Аналогично, $\bar{v} = \min\limits_{1 \leqslant j \leqslant n} \left\{ \max\limits_{1 \leqslant i \leqslant m} a_{ij} \right\}$ называется *верхним значением* игры, а j_0 — минимаксной стратегией второго игрока.

Всегда $\underline{v} \leqslant \bar{v}$. Если $\underline{v} = \bar{v} = v$, то игра имеет седловую точку в чистых стратегиях; число v называется *значением игры* (или *ценой игры*). Игра имеет седловую точку в чистых стратегиях тогда и только тогда, когда существует элемент матрицы $a_{i_0 j_0}$, минимальный в своей строке и в то же время максимальный в столбце

$$a_{ij_0} \leqslant a_{i_0 j_0} \leqslant a_{i_0 j}. \tag{10.1}$$

Любая пара (i_0, j_0), обладающая свойством (10.1), называется *седловой точкой.*

10.4. Смешанные стратегии

Если обозначить через $x_1, x_2, \ldots, x_m$ вероятности (частоты), с которыми первый игрок выбирает соответственно первую, вторую, ..., m-ю чистую стратегию, так что $x_i \geqslant 0$, $i = 1, 2, \ldots, m$, $\sum\limits_{i=1}^{m} x_i = 1$; через $y_1, y_2, \ldots, y_n$ — вероятности, с которыми второй игрок выбирает первую, вторую, ..., n-ю свою чистую стратегию, причем $y_j \geqslant 0$, $j = 1, 2, \ldots, n$, $\sum\limits_{j=1}^{n} y_j = 1$, то наборы чисел $x = (x_1, x_2, \ldots, \ldots, x_m)$ и $y = (y_1, y_2, \ldots, y_n)$ называются *смешанными стратегиями* первого и второго игроков соответственно. Каждый игрок имеет бесчисленное множество смешанных стратегий. Множество смешанных стратегий первого игрока обозначим через s_1 и множество смешанных стратегий второго игрока — через s_2.

Задача первого игрока состоит в выборе такой стратегии $x^* \in s_1$, чтобы при отсутствии информации о выборе другого максимизировать свой выигрыш. Задача второго игрока состоит в выборе такой стратегии $y^* \in s_2$, чтобы при отсутствии информации о поведении первого игрока минимизировать выигрыш первого.

Если первый игрок применяет стратегию $x \in s_1$, а второй — стратегию $y \in s_2$, то средний выигрыш $M(x, y)$ первого игрока равен

$$M(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j.$$

Выигрыш $M(x, y)$ называют *функцией игры*.

Например, в задаче с матрицей $A=\begin{pmatrix} 2 & 3 & -1 & 4 \\ 5 & -4 & 3 & -5 \end{pmatrix}$ первый игрок имеет две чистые стратегии $x^{(1)}=(1, 0)$ и $x^{(2)}=(0, 1)$ и бесчисленное множество смешанных стратегий, таких, как $x^{(3)}=(1/2, 1/2)$, $x^{(4)}=(1/17, 16/17)$, $x^{(5)}=(12/25, 13/25)$ и т. д.; все они являются элементами множества $s_1=\{x\colon x=(x_1, x_2), x_1\geqslant 0, x_2\geqslant 0, x_1+x_2=1\}$; второй игрок имеет четыре чистые стратегии $y^{(1)}=(1, 0, 0, 0)$, $y^{(2)}=(0, 1, 0, 0)$, $y^{(3)}=(0, 0, 1, 0)$, $y^{(4)}=(0, 0, 0, 1)$ и бесчисленное множество смешанных стратегий, таких, как $y^{(5)}=(1/4, 0, 1/4, 1/2)$, $y^{(6)}=(1/3, 1/6, 1/4, 1/4)$, $y^{(7)}=$ $=(1/20, 1/5, 1/2, 1/4)$, являющихся элементами множества $s_2=\left\{y\colon y=(y_1, y_2, y_3, y_4), y_j\geqslant 0, j=1, 2, 3, 4, \sum_{j=1}^{4} y_j=1\right\}$.

Если первый игрок применяет смешанную стратегию $x^{(0)}=(1/6, 5/6)$, а второй применяет стратегию $y^{(0)}=(1/3, 0, 1/3, 1/3)$, то средний выигрыш первого игрока, определяемый функцией игры, окажется равным

$$M(x^{(0)}, y^{(0)})=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}x_i^{(0)}y_j^{(0)},$$

$$M(x^{(0)}, y^{(0)})=2\cdot(1/6)\cdot(1/3)+3\cdot(1/6)\cdot 0-1\cdot(1/6)\cdot(1/3)+$$
$$+4\cdot(1/6)\cdot(1/3)+5\cdot(5/6)\cdot(1/3)-4\cdot(5/6)\cdot 0+3\cdot(5/6)\cdot(1/3)-$$
$$-5\cdot(5/6)\cdot(1/3)=10/9.$$

Если же первый игрок применяет стратегию $x^{(0)}=(1/6, 5/6)$, а второй — стратегию $y^{(1)}=(1/3, 0, 1/2, 1/6)$, то $M(x^{(0)}, y^{(1)})=25/12$. Оптимальная стратегия первого игрока $x^*=(7/11, 4/11)$, второго игрока $y^*=(0, 4/11, 7/11, 0)$; $5/11$ — цена игры.

10.5. Седловая точка в смешанных стратегиях

Пара смешанных стратегий (x^*, y^*) называется *седловой точкой* функции $M(x, y)$, если

$$M(x, y^*)\leqslant M(x^*, y^*)\leqslant M(x^*, y). \qquad (10.2)$$

Каждая матричная игра с нулевой суммой имеет решение в смешанных стратегиях, т. е. существуют такие смешанные стратегии x^* и y^* первого и второго игроков соответственно, что выполняется условие (10.2). Гарантированный выигрыш первого игрока, применяющего смешанную стратегию x,

$$v_1(x)=\min_{y\in s_2} M(x, y),$$

$$s_2=\left\{y\colon y=(y_1, y_2, \ldots, y_n),\ y_j\geqslant 0,\ \sum_{j=1}^{n} y_j=1\right\}.$$

Стратегия x^*, при которой гарантированный выигрыш первого игрока достигает максимального значения, называется *оптимальной стратегией* первого игрока:

$$v_1(x^*)=\max_{x\in s_1} v_1(x)=\max_{x\in s_1}\min_{y\in s_2} M(x, y),$$

$$s_1=\left\{x\colon x=(x_1, x_2, \ldots, x_m),\ x_i\geqslant 0,\ \sum_{i=1}^{m} x_i=1\right\}.$$

Гарантированный проигрыш второго игрока

$$u_2(y)=\max_{x\in s_1} M(x, y);$$

y^* — оптимальная стратегия второго игрока, если

$$u_2(y^*)=\min_{y\in s_2} u_2(y)=\min_{y\in s_2}\max_{x\in s_1} M(x, y).$$

Гарантированный выигрыш первого игрока, применяющего свою оптимальную стратегию, равен гарантированному проигрышу второго игроку, применяющего свою оптимальную стратегию: $v_1(x^*)=u_2(y^*)=v^*$; v^* — цена игры.

10.6. Сведение задачи теории игр к задаче линейного программирования

Задача максимизации гарантированного выигрыша первого игрока и задача минимизации гарантированного проигрыша второго игрока сводятся к паре взаимно двойственных задач линейного программирования:

Задача первого игрока	*Задача второго игрока*
$F=v(\max)$,	$\Phi=u(\min)$,
$\sum_{i=1}^{m} a_{ij}x_i\geqslant v,\ j=1, 2, \ldots, n,$	$\sum_{j=1}^{n} a_{ij}y_j\leqslant u,\ i=1, 2, \ldots, m,$
$\sum_{i=1}^{m} x_i=1,$	$\sum_{j=1}^{n} y_j=1,$
$x_i\geqslant 0,\ i=1, 2, \ldots, m.$	$y_j\geqslant 0,\ j=1, 2, \ldots, n.$

Процесс решения задач упрощается, если перейти к переменным $\xi_i=\frac{x_i}{v}$ $(i=1, 2, \ldots, m)$, $\eta_j=\frac{y_j}{u}$ $(j=1, 2, \ldots, n)$. Это возможно, если $a_{ij}\geqslant 0$. Имеем:

$$f=\sum_{i=1}^{m}\xi_i\,(\min), \qquad \varphi=\sum_{j=1}^{n}\eta_j\,(\max),$$

$$\sum_{i=1}^{m}a_{ij}\xi_i\geqslant 1,\ j=1,2,\ldots,n, \qquad \sum_{j=1}^{n}a_{ij}\eta_j\leqslant 1,\ i=1,2,\ldots,m,$$

$$\xi_i\geqslant 0,\ i=1,\ 2,\ \ldots,\ m. \qquad \eta_j\geqslant 0,\ j=1,\ 2,\ \ldots,\ n.$$

Оптимальные стратегии игроков не изменятся, если матрицу игры $A=\|a_{ij}\|$ заменить на $\tilde{A}=\|a_{ij}+c\|$. Цена игры при этом увеличивается на c.

10.7. Методы решения задач теории игр

Методы решения задач теории игр во многом зависят от условий задачи и от матрицы A выигрышей первого игрока.

1) Если матрица A имеет седловую точку, то решение игры сводится к нахождению седловой точки матрицы A. Оптимальные стратегии игроков определяются при этом координатами (i, j) седловой точки матрицы A, а цена игры — элементом a_{ij} в седловой точке.

○ **Пример 1.** Найти оптимальные стратегии игроков и цену игры:

$$A=\begin{pmatrix} 2 & 10 & 3 & 14 & 5\\ 8 & 9 & 5 & 6 & 7\\ 10 & 8 & 4 & 8 & 12\end{pmatrix}.$$

Минимизируя элементы первой строки, получим, что $\alpha_1=\min\{2,\ 10,\ 3,\ 14,\ 5\}=2$; аналогично, $\alpha_2=\min\{8,\ 9,\ 5,\ 6,\ 7\}=5$, $\alpha_3=4$. Максимизируя элементы по столбцам, получим: $\beta_1=10$, $\beta_2=10$, $\beta_3=5$, $\beta_4=14$, $\beta_5=12$. Нижняя цена игры $\underline{v}$ определяется путем максимизации α_i:

$$\underline{v}=\max_i \alpha_i=\max\{2,\ 5,\ 4\}=5.$$

Верхняя цена игры определяется минимизацией β_j:

$$\overline{v}=\min_j \beta_j=5.$$

Имеем $\underline{v}=\overline{v}$, так что $v=5$. Игра, определяемая матрицей $\tilde{A}$ (см. пример 1), имеет седловую точку (2, 3). Задача разрешима в чистых стратегиях. Оптимальный способ игры найден. Придерживаясь чистой второй стратегии, первый игрок обеспечивает себе выигрыш, не меньший 5, второй игрок, применяя чистую третью стратегию, проигрывает не более 5. Обе стратегии $i=2$ и $j=3$ являются оптимальными для первого и второго игроков. Цена игры $v=5$. ●

2) Если матрица A имеет размер $m\times 2$ или $2\times n$, то решение задачи может быть получено графически.

Иногда процесс решения задачи можно упростить, если вычеркнуть из матрицы доминируемые (неполезные) стратегии. Вычеркнутым стратегиям соответствуют нулевые компоненты в смешанных стратегиях. Стратегия i_0 для первого игрока является *доминируемой*, если существует такая стратегия i_1 для первого игрока, что

$$a_{i_1 j}\geqslant a_{i_0 j},\quad j=1,\ 2,\ \ldots,\ n,$$

или если существуют такие числа $\mu_i\geqslant 0$, $(i\neq i_0)$, $\sum\limits_{\substack{i\\ i\neq i_0}}\mu_i=1$, что

$$\sum_{\substack{i\\ i\neq i_0}}\mu_i a_{ij}\geqslant a_{i_0 j},\quad j=1,\ 2,\ \ldots,\ n.$$

Стратегия j_0 для второго игрока является *доминирующей*, если существует такая стратегия j_1 второго игрока, что $a_{ij_1}\leqslant a_{ij_0}$, $i=1,\ 2,\ \ldots,\ m$, или если существуют такие числа $\nu_j\geqslant 0$ $(j\neq j_0)$, $\sum\limits_{\substack{j\\ j\neq j_0}}\nu_j=1$, что

$$\sum_{\substack{j\\ j\neq j_0}}\nu_j a_{ij}\leqslant a_{ij_0},\quad i=1,\ 2,\ \ldots,\ m.$$

○ **Пример 2.** Найти оптимальные стратегии игроков и цену игры

$$A=\begin{pmatrix}3&7&2&1&5&8\\4&9&3&6&2&1\\2&3&1&4&9&20\end{pmatrix}.$$

В чистых стратегиях решения игры нет, так как

$$\underline{v}=\max_i\min_j a_{ij}=\max\{1,\ 1,\ 1\}=1,$$

$$\overline{v}=\min_j\max_i a_{ij}=\min\{4,\ 9,\ 3,\ 6,\ 9,\ 20\}=3,\ \underline{v}<\overline{v}.$$

Упростим матрицу A, заметив, что

$$a_{i2} \geqslant a_{i1}, \quad i=1,\ 2,\ 3; \quad a_{i1} \geqslant a_{i3}, \quad i=1,\ 2,\ 3;$$
$$a_{1j} \leqslant \frac{1}{2} a_{2j} + \frac{1}{2} a_{3j}, \ j=3,\ 4,\ 5,\ 6; \ a_{i4} \geqslant a_{i3}, \ i=2,\ 3.$$

Задачу с матрицей $\overline{A}$ размера 2×3 решим геометрически:

$$\overline{A} = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 9 & 20 \end{pmatrix}.$$

В плоскости переменных (x, v) построим $v_j(x)$ — ожидаемый средний выигрыш первого (рис. 10.1) игрока, применяющего первую стратегию с вероятностью x при условии, что второй игрок отвечает чистой стратегией j ($j=1,\ 2,\ 3$):

$$\begin{aligned} v_1(x) &= 3x + \ \ 1(1-x) = \quad 2x+1, \\ v_2(x) &= 2x + \ \ 9(1-x) = -\ 7x+9, \quad 0 \leqslant x \leqslant 1. \\ v_3(x) &= \ \ x + 20(1-x) = -19x+20, \end{aligned}$$

В точке A (рис. 10.1) пересечения прямых $v_1(x)$ и $v_2(x)$ гарантированный выигрыш первого игрока (изображенный жирной линией) достигает наибольшего значения

$$2x+1 = -7x+9,$$
$$9x = 8, \quad x = 8/9;$$
$$\overline{x}^* = (8/9,\ 1/9); \ \overline{y}^* = (7/9,\ 2/9,\ 0); \ \overline{v}^* = 25/9.$$

Исходная задача с матрицей A имеет следующее решение: $v^* = 25/9$, $x^* = (0,\ 8/9,\ 1/9)$, $y^* = (0,\ 0,\ 7/9,\ 0,\ 2/9,\ 0)$. ●

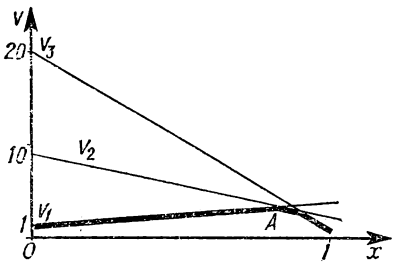


Рис. 10.1

3) Если задача теории игр не имеет решения в чистых стратегиях и не может быть решена графически, то для получения точного решения игры используют методы линейного программирования.

Целесообразно задачу второго игрока решать симплекс-методом. В последней симплекс-таблице, содержащей оптимальное решение задачи второго игрока, можно найти и оптимальное решение двойственной задачи — задачи первого игрока.

○ **Пример 3.** Получить решение задачи теории игр, если матрица выигрышей первого игрока известна:

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

Очевидно, что матрица A не имеет седловой точки и что графический способ решения неприемлем.

Построим пару взаимно двойственных задач линейного программирования, эквивалентную игре с матрицей $\tilde{A}=$ $=A+2$. Имеем:

$$\begin{array}{ll}
f=\eta_1+\eta_2+\eta_3+\eta_4\,(\max), & \varphi=\xi_1+\xi_2+\xi_3\,(\min),\\
2\eta_1+9\eta_3+6\eta_4\leqslant 1, & 2\xi_1+\xi_2+4\xi_3\geqslant 1,\\
\eta_1+3\eta_2+6\eta_3\leqslant 1, & 3\xi_2+2\xi_3\geqslant 1,\\
4\eta_1+2\eta_2+\eta_3+3\eta_4\leqslant 1, & 9\xi_1+6\xi_2+\xi_3\geqslant 1,\\
\eta_j\geqslant 0,\quad j=1,\,2,\,3,\,4; & \xi_i\geqslant 0,\quad i=1,\,2,\,3.
\end{array}$$

Сведем задачу второго игрока к канонической задаче минимизации

$$\begin{array}{l}
f'=-\eta_1-\eta_2-\eta_3-\eta_4\quad(\min)\\
2\eta_1+9\eta_3+6\eta_4+\eta_5=1,\\
\eta_1+3\eta_2+6\eta_3+\eta_6=1,\\
4\eta_1+2\eta_2+\eta_3+3\eta_4+\eta_7=1,\\
\eta_j\geqslant 0,\quad j=1,\,2,\,\ldots,\,7
\end{array}$$

и составим симплекс-таблицу, которая приведена к базису A_5, A_6, A_7 опорного решения $\alpha=(0,\,0,\,0,\,0,\,1,\,1,\,1)$:

η_1	η_2	η_3	η_4	η_5	η_6	η_7	
2	0	9	6	1	0	0	1
1	$\boxed{3}$	6	0	0	1	0	1
4	2	1	3	0	0	1	1
1	1	1	1	0	0	0	0

η_1	η_2	η_3	η_4	η_5	η_6	η_7	
2	0	9	6	1	0	0	1
$1/3$	1	2	0	0	$1/3$	0	$1/3$
$10/3$	0	-3	$\boxed{3}$	0	$-2/3$	1	$1/3$
$2/3$	0	-1	1	0	$-1/3$	0	$-1/3$

Выбираем положительную оценку $\delta_2=1$ и находим наименьшее отношение $\min\{1/3,\ 1/2\}=1/3$. С ведущим элементом $a_{22}=3$ выполняем жорданово преобразование таблицы. Получаем

η_1	η_2	η_3	η_4	η_5	η_6	η_7	
$-14/3$	0	15	0	1	$4/3$	-2	$1/3$
$1/3$	1	2	0	0	$1/3$	0	$1/3$
$10/3$	0	-1	1	0	$-2/9$	$1/3$	$1/9$
$-4/9$	0	0	0	0	$-1/9$	$-1/3$	$-4/9$

Все оценки базиса A_2, A_4, A_5 не положительны. Следовательно, $\alpha_3=(0,\ 1/3,\ 0,\ 1/9,\ 1/3,\ 0,\ 0)$ — оптимальное опорное решение канонической задачи минимизации. Поэтому $\beta=(0,\ 1/3,\ 0,\ 1/9)$ является оптимальным решением задачи линейного программирования и $f(\beta)=4/9$.

Компоненты оптимального решения двойственной задачи определяются с помощью оценочной строки последней симплекс-таблицы, а именно $\gamma=(-\delta_5,\ -\delta_6,\ -\delta_7)$, так что $\gamma=(0,\ 1/9,\ 1/3)$ — оптимальное решение двойственной задачи. Переходя к первоначальным переменным $y_j=\eta_j u$, $x_i=\xi_i v$ задачи, получим $\tilde{u}^*=9/4$; $y^*=(0,\ 3/4,\ 0,\ 1/4)$; $x^*=(0,\ 1/4,\ 3/4)$. Наконец, $v^*=u^*=9/4-2=1/4$, $x^*=(0,\ 1/4,\ 3/4)$, $y^*=(0/4,\ 0,\ 1/4)$. ●

10.8. Метод Брауна приближенного решения задач теории игр

Метод Брауна представляет собой модель практического взаимного обучения игроков, при котором каждый игрок, анализируя способ поведения противника, старается ответить наилучшим способом.

Первый игрок выбирает одну из своих стратегий. Второй игрок отвечает стратегией, которая минимизирует выигрыш первого игрока. Каждый игрок отвечает на очередной ход противника той своей стратегией, которая является оптимальной относительно всех предыдущих ходов противника. Рассмотрим решение задачи (табл. 10.1) с матрицей

$$A=\begin{pmatrix}2&0&9&6\\1&3&6&0\\4&2&1&3\end{pmatrix}.$$

Первый игрок пусть выбирает $i_0=3$. Второй игрок, просматривая третью строку, выбирает $a_{13}=1$ — наименьший элемент строки — и отмечает его звездочкой (*). Справа от матрицы выписываем элементы третьего столбца. Первый игрок, зная выбор второго и просматривая выделенный столбец, выбирает в нем максимальный элемент (отмеченный звездочкой) и назначает $i_1=1$. Второй строкой под матрицей выписываем суммарный выигрыш первого игрока за два хода.

Минимизируя выигрыш первого игрока за два хода, второй игрок выбирает $j_2=2$ и т. д.

Частота выбора строк и столбцов определяет компоненты приближенного решения, элементы со звездочкой определяют суммарный выигрыш за n ходов.

Таблица 10.1

2	0	9	6	9*	9	9	9	15	21*	23*	23*	23*	23
1	3	6	0	6	9*	12*	15*	15*	15	16	19	22	25
4	2	1	3	1	3	5	7	10	13	17	19	21	23
4	2	1*	3										
6	2*	10	9										
7	5*	16	9										
8	8*	22	9*										
9	11	28	9*										
10	14	34	9*										
12*	14	43	15										
14	14*	52	21										
16	14*	61	27										
18	17*	70	33										

Вычисления в итерационном методе Брауна удобно располагать в приведенной таблице. В левом верхнем углу записана матрица игры — в данном случае матрица размера 3×4. Строки ниже матрицы содержат сложенный за n партий возможный выигрыш первого игрока, где n — количество партий и номер строки. Столбцы правее матрицы представляют собой сложенный за n партий возможный проигрыш второго игрока, причем n — количество партий и номер столбца.

Пусть после n партий первый игрок обнаружил, что его противник выбрал j-ю чистую стратегию (по количеству «звездочек» в j-м столбце под матрицей) s_j раз ($j = 1, 2, \ldots, n$). На этом основании первый игрок допускает, что его противник придерживается смешанной стратегии $y_n = \left(\frac{s_1}{n}, \frac{s_2}{n}, \ldots, \frac{s_n}{n}\right)$ и выбирает в последующей партии чистую стратегию, дающую максимальный средний выигрыш против y_n. Номер чистой стратегии — это номер максимальной компоненты n-го столбца.

Аналогично, второй игрок после n партий предполагает, что смешанная стратегия соперника $x_n = \left(\frac{t_1}{n}, \frac{t_2}{n}, \ldots \right.$ $\left.\ldots, \frac{t_m}{n}\right)$, где t_i — частота появления i-й стратегии в предыдущих партиях. В $(n+1)$-й партии второй игрок выбирает чистую стратегию, обеспечивающую ему минимальный средний проигрыш. Номер этой чистой стратегии — это номер столбца минимального элемента n-й строки.

Степень приближения к решению зависит от выбора начальной строки и от числа ходов, поэтому:

при $n=10$

$$\frac{17}{10} \leqslant \bar{v} \leqslant \frac{25}{10},$$

$\bar{x}=(5/10,\ 5/10,\ 0)$,

$\bar{y}=(1/10,\ 6/10,\ 1/10,\ 2/10)$.

при $n=14$

$$\frac{22}{14} \leqslant \bar{v} \leqslant \frac{35}{14},$$

$\bar{x}=(5/14,\ 8/14,\ 1/14)$,

$\bar{y}=(3/14,\ 8/14,\ 1/14,\ 2/14)$.

10.9. Экономические задачи, приводимые к матричным играм

Задача 1. Пусть сельскохозяйственная бригада может посеять одну из трех культур: A_1, A_2 или A_3. Урожаи этих культур во многом зависят от погоды. Требуется установить, какую из культур сеять, чтобы обеспечить наибольший доход, если известны цена a_i 1 ц культуры A_i и средняя урожайность каждой культуры в зависимости от погоды (будет ли лето засушливым, нормальным или дождливым, достоверный прогноз погоды отсутствует).

У предприятия (первый игрок) имеется три стратегии — сеять культуру A_1, A_2 или A_3. Будем считать, что противник — природа (второй игрок) — также имеет три стратегии: погода засушливая, нормальная или дождливая. В качестве выигрыша первого игрока возьмем функцию доходов предприятия от реализации своей продукции. Получим матрицу выигрышей первого игрока:

$$A=\begin{pmatrix} a_1h_{11} & a_1h_{12} & a_1h_{13} \\ a_2h_{21} & a_2h_{22} & a_2h_{23} \\ a_3h_{31} & a_3h_{32} & a_3h_{33} \end{pmatrix},$$

где h_{ij} — урожайность i-й культуры при погодных условиях j ($j=1,\ 2,\ 3$).

Задача 2. На базе торговой организации имеется n типов одного из товаров ассортиментного минимума. В магазин должен быть завезен только один из n типов данного товара. Требуется выбрать тот тип товара, который целесообразно завезти в магазин. Если товар типа j ($j = 1,\ 2,\ \ldots,\ n$) будет пользоваться спросом, то магазин от его реализации получит прибыль p_j. Если же этот товар не будет пользоваться спросом, то убытки от его хранения принесут магазину убыток q_j. В условиях неопределенного покупательского опроса конфликт товароснабжения формализуется игрой. Пусть первый игрок — магазин, второй игрок — покупательский спрос. Каждая из сторон имеет по n стратегий. Завоз i-го товара — i-я стратегия первого игрока, спрос на j-й товар — j-я стратегия второго игрока. Матрица выигрышей первого игрока имеет вид

$$A = \begin{pmatrix} p_1 & -q_1 & \dots & -q_1 \\ -q_2 & p_2 & \dots & -q_2 \\ \vdots & \vdots & & \vdots \\ \dots & \dots & \dots & \dots \\ -q_n & -q_n & \dots & p_n \end{pmatrix}.$$

РАЗДЕЛ XI. ГРАФЫ И СЕТИ

11.1. Основные понятия теории графов

Граф G — это совокупность двух конечных множеств: множества точек, которые называются *вершинами*, и множества пар вершин, которые называются *ребрами*.

На рис. 11.1 изображен граф, имеющий пять вершин и шесть ребер.

Если рассматривается множество упорядоченных пар точек, т. е. на каждом ребре задается направление, то граф G называется *ориентированным*. В противном случае G называется *неориентированным графом*.

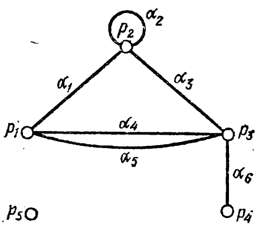


Рис. 11.1

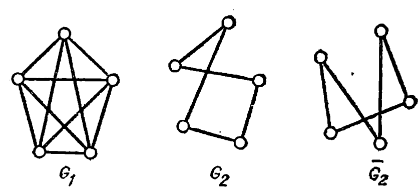


Рис. 11.2

Ребра, имеющие одинаковые концевые вершины, называются *параллельными*.

Ребро, концевые вершины которого совпадают, называется *петлей*. На рис. 11.1 α_4 и α_5 — параллельные ребра, а α_2 — петля.

Вершина и ребро называются *инцидентными* друг другу, если вершина является для этого ребра концевой точкой. На рис. 11.1 вершина p_3 и ребро α_3 инцидентны друг другу.

Две вершины, являющиеся концевыми для некоторого ребра, называются *смежными вершинами*. Два ребра, инцидентные одной и той же вершине, называются *смежными ребрами*. На рис. 11.1 p_1, p_2 — смежные вершины, а α_1, α_4 — смежные ребра.

Степенью вершины называется число ребер, инцидентных ей. Вершина степени 1 называется *висячей*. Вершина степени 0 называется *изолированной*. На рис. 11.1 степень вершины p_1 равна трем, p_4—висячая вершина, p_5—изолированная.

Теорема 11.1. *В графе G сумма степеней всех его вершин—число четное, равное удвоенному числу ребер графа:*

$$\sum_{i=1}^{n} \text{степ. } p_i = 2m,$$

где n—число вершин графа, а m—число его ребер.

Теорема 11.2. *Число нечетных вершин любого графа, т. е. вершин, имеющих нечетную степень, четно.*

Граф G называется *полным*, если любые две его различные вершины соединены ребром и он не содержит параллельных ребер.

Дополнением графа G называется граф $\bar{G}$ с теми же вершинами, что и граф G, и содержащий только те ребра, которые нужно добавить к графу G, чтобы получился полный граф. На рис. 11.2 изображены следующие графы: G_1—полный граф с пятью вершинами, G_2—некоторый граф, имеющий пять вершин, $\bar{G}_2$—дополнение графа G_2.

Путем в графе называется такая последовательность ребер, ведущая от некоторой начальной вершины p_1 в конечную вершину p_n, в которой каждые два соседних ребра имеют общую вершину и никакое ребро не встречается более одного раза. Например, в графе, изображенном на рис. 11.1, последовательность ребер $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$ образует путь, ведущий от вершины p_1 к вершине p_4.

Циклом называется путь, начальная и конечная вершины которого совпадают. На рис. 11.1 ребра $(\alpha_1, \alpha_3, \alpha_4)$ образуют цикл.

Цикл графа G называется *простым*, если он не проходит ни через одну вершину G более одного раза.

Длиной пути или цикла называется число ребер этого пути или цикла.

11.2. Связные графы

Граф G называется *связным*, если для любых двух его вершин существует путь, их соединяющий. В противном случае граф G называется *несвязным*.

Любой несвязный граф является совокупностью связных графов. Эти графы обладают тем свойством, что никакая вершина одного из них не связана путем ни с какой вершиной другого. Каждый из этих графов называется *компонентой графа* G. На рис. 11.3 изображен несвязный граф G с компонентами G_1, G_2, G_3. Каждая компонента является связным графом.

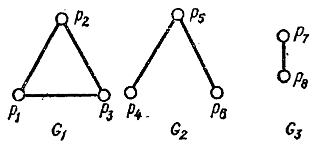


Рис. 11.3

Теорема 11.3. *Для того чтобы граф G представлял собой простой цикл, необходимо и достаточно, чтобы каждая его вершина имела степень 2.*

Ребро α называется *мостом графа* G, если граф, получившийся из G после удаления ребра α (такой граф обозначается $G \setminus \alpha$), содержит больше компонент, чем граф G.

Теорема 11.4. *Ребро α графа G является мостом тогда и только тогда, когда α не принадлежит ни одному циклу.*

11.3. Подграфы

Рассмотрим граф $G = (\mathscr{P}, A)$ с множеством вершин $\mathscr{P}$ и множеством ребер A. Граф $G' = (\mathscr{P}', A')$ называется подграфом графа G, если $\mathscr{P}'$ и A' являются подмножествами $\mathscr{P}$ и A, причем ребро содержится в A' только в том случае, если его концевые вершины содержатся в $\mathscr{P}'$. Пусть $\mathscr{P}'$ — некоторое подмножество множества вершин графа $G = (\mathscr{P}, A)$ и пусть A' — множество всех ребер графа G, концевые вершины которых входят в $\mathscr{P}'$. Тогда граф $G' = (\mathscr{P}', A')$ называется *вершинно-порожденным* подграфом графа G.

Обозначим через A' некоторое подмножество множества ребер графа $G = (\mathscr{P}, A)$ и пусть $\mathscr{P}'$ есть множество всех вершин графа G, инцидентных ребрам из A'. Тогда граф $G' = (\mathscr{P}', A')$ называется *реберно-порожденным* подграфом графа G.

На рис. 11.4 изображены вершинно-порожденный подграф G_1 графа G, представленного на рис. 11.1 (множество вершин p_1, p_3, p_4), и реберно-порожденный подграф G_2 того же графа G (множество ребер α_1, α_3, α_4, α_6).

11.4. Операции над графами

Объединением графов $G_1 = (\mathscr{P}_1, A_1)$ и $G_2 = (\mathscr{P}_2, A_2)$ называется граф $G = G_1 \cup G_2$, множество вершин которого есть объединение множеств вершин графов G_1 и G_2 ($\mathscr{P} = \mathscr{P}_1 \cup \mathscr{P}_2$), а множество ребер является объединением множеств ребер этих графов ($A = A_1 \cup A_2$).

Пересечением графов G_1 и G_2 называется граф $G = G_1 \cap G_2$, множество вершин которого есть пересечение $\mathcal{P}_1 \cap \mathcal{P}_2$, а множество ребер — пересечение $A_1 \cap A_2$.

Кольцевой суммой двух графов называется граф $G_1 \oplus G_2$, порожденный на множестве ребер $(A_1 \cup A_2) \setminus (A_1 \cap A_2)$, т. е. на множестве ребер, присутствующих либо в G_1, либо в G_2, но не принадлежащих их пересечению $G_1 \cap G_2$.

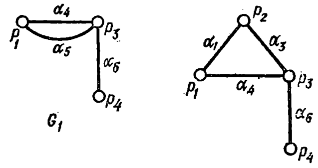


Рис. 11.4

Очевидно, что все эти три операции коммутативны.

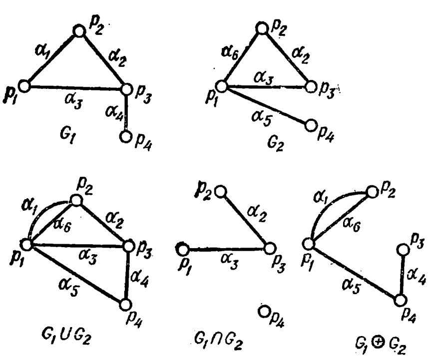


Рис. 11.5

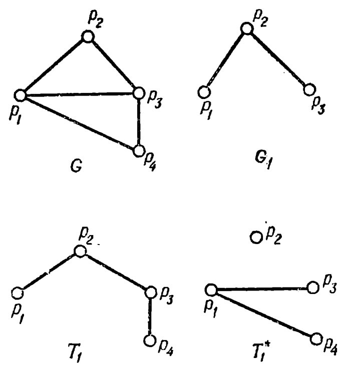


Рис. 11.6

На рис. 11.5 изображены графы G_1, G_2, $G_1 \cup G_2$, $G_1 \cap G_2$, $G_1 \oplus G_2$.

11.5. Деревья

Связный граф, не содержащий циклов, называется *деревом*.

Деревом некоторого *графа* G называется его связный подграф без циклов. Дерево графа G, содержащее все его вершины, называется *остовом графа* G или его *покрывающим деревом*.

Кодеревом T^* остова T графа G называется такой подграф G, который содержит все его вершины и только те ребра, которые не входят в T.

На рис. 11.6 представлены граф G, его дерево G_1, остов T_1 и кодерево T_1^*.

Теорема 11.5. *Граф G с n вершинами является деревом тогда и только тогда, когда G-связный граф и число его ребер равно $(n-1)$.*

Ребра остова T называются *ветвями графа* G, а ребра кодерева — *T^*-связями.*

Теорема 11.6. *Граф G является деревом тогда и только тогда, когда G не содержит циклов и при соединении ребром произвольных двух его несмежных вершин получается граф, имеющий ровно один цикл.*

11.6. Лес. Разрезы

Граф, не содержащий циклов и состоящий из k компонент, называется k-*деревом*; k-*дерево* графа G, содержащее все его вершины, называется остовным.

Подграф G, содержащий все его вершины и только те ребра, которые не входят в остовное k-дерево T графа G, называется k-*кодеревом* T^*.

Если граф G содержит k компонент, то его остовное k-дерево T называется лесом, а k-кодерево T^* в этом случае называется *КО-лесом*.

На рис. 11.7 изображены остовное 2-дерево T и 2-кодерево T^* графа G, представленного на рис. 11.6.

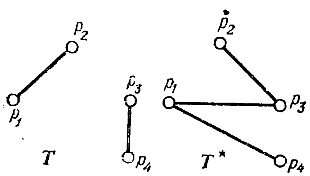


Рис. 11.7

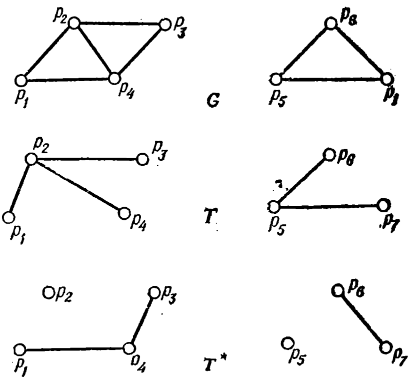


Рис. 11.8

На рис. 11.8 представлен граф G, содержащий две компоненты, его лес T и $КО$-лес T^*.

Рангом графа G, имеющего n вершин и состоящего из k компонент, называется число $r(G) = n - k$.

Цикломатическим числом графа называется число $\mu(G) = m - n + k$, где m — число ребер графа G.

Очевидно, что ранг и цикломатическое число связаны следующим соотношением:

$$r(G) + \mu(G) = m.$$

Теорема 11.7. *Ранг $r(G)$ графа G равен числу ребер леса, а цикломатическое число $\mu(G)$ равно числу ребер КО-леса.*

Ранг и цикломатическое число являются числовыми характеристиками графа, определяющими размерность подпространств циклов и разрезов.

Пусть есть некоторый связный граф G, множество вершин которого разбито на два непустых непересекающихся подмножества $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2$. Тогда множество всех ребер G, имеющих одну концевую вершину в $\mathcal{P}_1$, а другую — в $\mathcal{P}_2$, называется *разрезом графа* G.

11.7. Эйлеровы и гамильтоновы графы

Эйлеровым путем (*циклом*) графа называется путь (цикл), содержащий все ребра графа ровно один раз.

Граф, обладающий эйлеровым циклом, называется *эйлеровым графом*.

На рис. 11.9 граф G не является эйлеровым, так как вершина p_3 инцидентна только одному ребру. Если путь

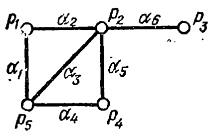


Рис. 11.9

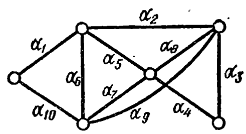


Рис. 11.10

приведет в вершину p_3, то не будет ребра, по которому можно было бы выйти из p_3.

Теорема 11.8. *Граф G является эйлеровым тогда и только тогда, когда G-связный и все его вершины имеют четную степень.*

Граф G, изображенный на рис. 11.10, является эйлеровым. Последовательность ребер (α_1, α_2, α_3, α_4, α_5, α_6, α_7, α_8, α_9, α_{10}) образует эйлеров цикл.

Теорема 11.9. *Граф G обладает эйлеровым путем с концами p_1, p_2 тогда и только тогда, когда G-связный и p_1, p_2 — единственные его вершины нечетной степени.*

На рис. 11.9 изображен граф G, обладающий эйлеровым путем (α_1, α_2, α_3, α_4, α_5, α_6) с концевыми вершинами p_5 и p_3.

Гамильтоновым путем (*циклом*) графа G называется путь (цикл), проходящий через каждую вершину G в точности по одному разу.

Граф, обладающий гамильтоновым циклом, называется *гамильтоновым*.

Критерий существования гамильтонова цикла в произвольном графе G еще не найден. Достаточным условием существования гамильтонова цикла является полнота графа G.

Граф, изображенный на рис. 11.9, не является гамильтоновым, а граф, представленный на рис. 11.11, содержит гамильтонов цикл (α_1, α_2, α_4, α_5, α_6).

11.8. Ориентированные графы

В ориентированных графах на ребрах задано направление, т. е. у каждого ребра фиксируются начало и конец. Такие направленные ребра называются *дугами*.

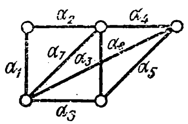


Рис. 11.11

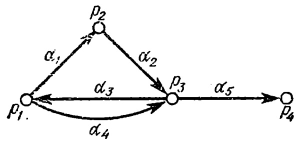


Рис. 11.12

Цепью в ориентированном графе называется такая последовательность дуг, ведущих от вершины p_1 к вершине p_n, в которой каждые две соседних дуги имеют общую вершину и никакая дуга не встречается более одного раза. Если направление цепи совпадает с направлением всех принадлежащих ей дуг, то цепь называется *путем*.

В ориентированных графах циклом называется путь, начало и конец которого совпадают. На рис. 11.12 дуги (α_1, α_4, α_5) образуют цепь, а дуги (α_1, α_2, α_5)—путь. Последовательность дуг (α_1, α_2, α_3) составляет цикл, а последовательность (α_1, α_2, α_4) не является циклом.

Цепь, путь, цикл в произвольном графе G называются *простыми*, если они не проходят ни через одну свою вершину более одного раза.

Длиной цепи, пути, цикла называется число содержащихся в них дуг.

Сетью называется граф, каждой дуге которого поставлено в соответствие некоторое число (или несколько чисел).

Многие практические задачи могут быть решены с помощью теории графов. Рассмотрим некоторые примеры.

Задача размещения. В некотором районе имеется n населенных пунктов, соединенных друг с другом сетью шоссейных дорог. Нужно построить станцию технического обслуживания автомобилей вблизи шоссейной дороги. Расположение станции должно быть удобно для жителей всех населенных пунктов. Например, можно потребовать, чтобы сумма расстояний от станции до населенных пунктов была минимальной.

Легко представить эту задачу с помощью графа. Населенные пункты представляются вершинами графа, а шоссейные дороги — дугами, которые их соединяют. Станция технического обслуживания должна быть расположена на одной из дуг графа.

Задача почтальона. Почтальон ежедневно забирает письма на почте, разносит их адресатам, располагающимся вдоль всего маршрута его следования, и возвращается обратно на почту. Путь, пройденный почтальоном, должен быть как можно короче.

Улицы представляются дугами некоторого графа. Задача почтальона заключается в том, чтобы найти маршрут обхода всех дуг графа минимальной длины.

Задача строительства дорог. Имеется n городов, которые нужно соединить между собой новыми дорогами (не обязательно каждые два города должны быть связаны друг с другом непосредственно). Известна стоимость прокладки дороги между любыми двумя городами. Требуется составить проект строительства дорог минимальной стоимости.

Очевидно, что каждый город представляется вершиной графа, а дорога — дугой, соединяющей вершины. Каждой дуге ставится в соответствие число, равное стоимости строительства соответствующей дороги. Требуется построить некоторое «покрытие» графа минимальной стоимости.

11.9. Матрицы графов

Граф может быть задан разными способами: рисунком, перечислением вершин и ребер (или дуг) и т. д. Одним из самых удобных способов является задание графа с помощью матрицы. Пусть некоторый граф G имеет n вершин $p_1, p_2, \ldots, p_n$ и m ребер $\alpha_1, \alpha_2, \ldots, \alpha_m$. Построим матрицу, имеющую n строк и m столбцов. Каждая строка матрицы будет соответствовать вершине, а столбец — ребру графа. В столбце α_j все элементы, кроме двух, будут равны нулю. Для ориентированного графа в строке, соответствующей начальной вершине дуги α_j, ставят число $+1$, а в строке, соответствующей конечной вершине, — число -1. Для неориентированного графа в строчках матрицы, соответствующих концевым вершинам ребра α_j, ставят 1, а в остальных строчках — 0. Для графов, изображенных на рис. 11.12 и 11.9, матрицы имеют соответственно следующий вид:

	α_1	α_2	α_3	α_4	α_5
p_1	1	0	-1	1	0
p_2	-1	1	0	0	0
p_3	0	-1	1	-1	1
p_4	0	0	0	0	-1

	α_1	α_2	α_3	α_4	α_5	α_6
p_1	1	1	0	0	0	0
p_2	0	1	1	0	1	1
p_3	0	0	0	0	0	1
p_4	0	0	0	1	1	0
p_5	1	0	1	1	0	0

Построенные матрицы называются *матрицами инциденций* графа.

Матричное представление графа является наиболее удобной формой задания графа при вычислениях на машине.

11.10. Максимальные потоки в сети

Поток на графе — это совокупность однородных объектов, пересылаемых из одной вершины в другую по его дугам. Если вершины p и q соединены дугой $\alpha = (p, q)$, то поток из p в q обозначается $f(p, q)$. Таким образом, поток — это некоторая функция, заданная на дугах графа.

Пусть имеется ориентированный граф, на дугах которого определена функция $C(p, q)$. В потоковых задачах $C(p, q)$ обычно означает пропускную способность дуги (p, q) или стоимость перевозки единицы потока по этой дуге.

Дивергенцией потока f в вершине p называется разность выходящих и входящих потоков:

$$\mathrm{div}_f(p) = \sum_{(p,\, q) \in A(p)} f(p, q) - \sum_{(r,\, p) \in B(p)} f(r, p),$$

где $A(p)$ — множество дуг, выходящих из вершины p, а $B(p)$ — множество входящих в нее дуг.

Вершины, в которых $\operatorname{div}_f(p) > 0$, называются *источниками потока* f, а вершины, в которых $\operatorname{div}_f(p) < 0$,— его *стоками*.

Сложим дивергенцию потока f во всех вершинах графа. Очевидно, что

$$D = \sum_p \operatorname{div}_f(p) = 0.$$

Пусть задана сеть, в которой выделены две вершины s и t. Рассмотрим такие потоки $f(\alpha)$, что:

1) $0 \leqslant f(\alpha) \leqslant C(\alpha)$ на всех дугах α;

2) $\operatorname{div}_f(p) = 0$ при всех p, кроме, быть может, $p = s$ и $p = t$.

Вершины s и t называются *полюсами сети*, а остальные вершины называются *внутренними*.

Из предыдущего следует, что $\operatorname{div}_f(s) = -\operatorname{div}_f(t) = M$. Если $M = 0$, то поток называется *циркуляцией*. В противном случае ($M \neq 0$) вершина s является источником потока, а вершина t—стоком. Число $M > 0$ называется мощностью потока f.

Задача состоит в том, чтобы найти поток f максимальной мощности, удовлетворяющий условиям 1) и 2).

Если рассматриваемая сеть G содержит параллельные дуги, то нужно рассмотреть новую сеть G' с теми же вершинами, что и G, в которой параллельные дуги α и β склеены. Пропускная способность полученной при этом дуги γ есть $C(\gamma) = C(\alpha) + C(\beta)$. После получения решения f для сети G' оно разбивается на сумму $f(\gamma) = f_1 + f_2$, где $0 \leqslant f_1 \leqslant C(\alpha)$, $0 \leqslant f_2 \leqslant C(\beta)$.

Нахождение увеличивающей цепи

Пусть заданы сеть G и какой-то поток f на этой сети. Дуги α, в которых $f(\alpha) \leqslant C(\alpha)$, называются *увеличивающими*, так как поток через эти дуги можно увеличить на величину $\Delta f(\alpha) = C(\alpha) - f(\alpha)$. Множество увеличивающих дуг обозначается A^+.

Дуги α, в которых $f(\alpha) > 0$, называются *уменьшающими*, так как поток через эти дуги можно уменьшить на величину $f(\alpha)$. Множество уменьшающих дуг обозначают A^-.

Заметим, что дуга α может принадлежать одновременно множествам A^+ и A^-.

Пример. Рассмотрим сеть, изображенную на рис. 11.13, и некоторый поток f. Возьмем цепь (s, p, r, m, n, t), соединяющую вершины s и t. Вдоль этой цепи можно увеличить поток на минимальную из величин

$$\Delta f(s,\ p)=9-7=2,\ \Delta f(p,\ r)=4-2=2,$$
$$f(m,\ r)=1,\ f(n,\ m)=1,\ \Delta f(n,\ t)=4-3=1.$$

Очевидно, что мощность потока возрастет на 1 (рис. 11.14).

В общем случае вдоль некоторой цепи, соединяющей вершины s и t, можно увеличить поток, если ее прямые дуги — увеличивающие, а все обратные — уменьшающие.

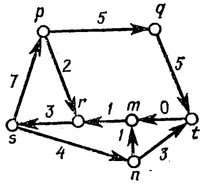


Рис. 11.13

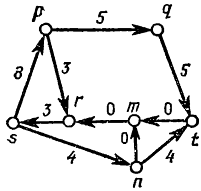


Рис. 11.14

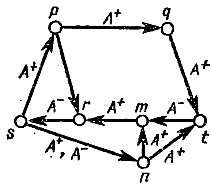


Рис. 11.15

Максимальный дополнительный поток δ, который можно переслать вдоль такой цепи, равен минимуму из всех $\Delta f(\alpha)$ для прямых дуг α и всех $f(\alpha)$ для обратных дуг α.

Алгоритм нахождения увеличивающей цепи

1. Находят множество дуг A^+ и A^-.

2. Отмечают вершину s (источник).

3. Далее отмечают некоторые дуги и вершины сети G. Именно: если вершина p отмечена, а вершина q — нет и дуга $\alpha=(p,\ q)\in A^+$, то отмечают дугу α и вершину q. Если вершина p отмечена, а вершина q не отмечена и дуга $\alpha=(q,\ p)\in A^-$, то отмечают дугу α и вершину q.

В результате получают некоторое подмножество отмеченных вершин, соединенных отмеченными дугами. При этом, отмечая вершину q, нужно одновременно запомнить вершину p, отмеченную ранее, откуда в q привела отмеченная дуга α (прямая или обратная).

Замечание. Приведенный алгоритм не указывает, в каком порядке нужно просматривать дуги. Например, можно сначала просмотреть все дуги, инцидентные s, отметить некоторые вершины в соответствии с правилами; затем просмотреть все дуги, инцидентные отмеченным вершинам, и т. д. В другом варианте всякий раз просматривают дуги, инцидентные последней отмеченной вершине, и т. д. Этот алгоритм позволяет получить увеличивающую цепь из s в t.

○ **Пример.** Применим алгоритм построения увеличивающей цепи к графу, изображенному на рис. 11.15.

Для любой отмеченной вершины будем просматривать все инцидентные ей дуги. Отмечаем источник s. Дуги (s, p), (s, n) отмечаем как выходящие увеличивающие, а дугу (r, s)—как входящую уменьшающую (рис. 11.16).

В скобках указаны вершины, соединенные с новыми вершинами отмеченными дугами. Далее просматриваем дуги, инцидентные p (рис. 11.17). Дугу (p, r) не рассматриваем, поскольку вершина r уже отмечена. Далее,

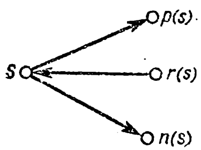


Рис. 11.16

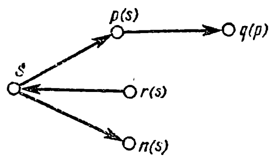


Рис. 11.17

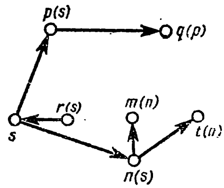


Рис. 11.18

просматривая дуги, инцидентные n, отмечаем дугу (n, t) и вершину t (рис. 11.18).

На этом просмотр заканчиваем, так как отмечен сток t. Идя в обратном направлении, получаем увеличивающую цепь (s, n, t). ●

Алгоритм отыскания максимального потока в сети

Данный алгоритм, использующий увеличивающие цепи, называется *алгоритмом Форда.* Он состоит из следующих шагов.

1. Выбирают некоторый поток $f(\alpha)$ из s в t (например, $f(\alpha) \equiv 0$).
2. Находят классы увеличивающих (A^+) и уменьшающих (A^-) дуг.
3. Применяют алгоритм поиска увеличивающей цепи из s в t. Если такой цепи нет, то поток f—максимальный. Если цепь C найдена, то перейти к шагу 4.
4. Находят $\min\limits_{\alpha \in A^+ \cap C} (\Delta f(\alpha))$, $\min\limits_{\alpha \in A^- \cap C} f(\alpha)$.
5. Пусть $\delta > 0$—наименьшее из этих чисел.
6. Увеличивают поток вдоль цепи C на δ и переходят к шагу 2.

Очевидно, что построенный таким образом поток удовлетворяет следующим условиям:

$$f(\alpha) \leqslant c(\alpha),$$
$$\operatorname{div}_f(p) = 0 \text{ при } p \neq s,\ p \neq t.$$

Этот поток является максимальным. Если начальные данные целочисленны, то выполнение алгоритма Форда завершается за конечное число шагов.

11.11. Задача о кратчайшем пути между двумя вершинами графа

Пусть каждой дуге (x, y) графа G поставлено в соответствие число $l(x, y) \geqslant 0$, которое называется длиной этой дуги (если вершины x и y не соединены дугой, то полагают $l(x, y) = \infty$). Длина пути определяется как сумма длин отдельных дуг, составляющих этот путь. Требуется найти кратчайший путь между двумя заданными вершинами s и t графа G.

Алгоритм поиска кратчайшего пути

В ходе выполнения алгоритма окрашивают вершины и дуги графа и вычисляют величины $d(x)$, равные кратчайшему пути из вершины s в вершину x, включающему только окрашенные вершины.

1. Полагают $d(s) = 0$, $d(x) = \infty$ для любого $x \neq s$. Окрашивают вершину s и полагают $y = s$.

2. Для каждой неокрашенной вершины x пересчитывают величину $d(x)$ по формуле $d(x) = \min\{d(x), d(y) + l(y, x)\}$.

Если $d(x) = \infty$ для всех вершин, то вычисления заканчивают. В графе G отсутствуют дуги из вершины s в неокрашенные вершины.

В противном случае окрасить вершину x, для которой величина $d(x)$ минимальна, и дугу, ведущую в вершину x. Полагают $y = x$.

3. Если $y = t$, то кратчайший путь найден. В противном случае переходят к шагу 2.

С помощью описанного алгоритма можно определить кратчайший путь из s во все вершины исходного графа. Для этого процедуру окрашивания нужно продолжить до тех пор, пока все вершины графа не будут окрашены. При этом для графа G будет построено покрывающее дерево с корнем в вершине s (если такое дерево существует). В вершину s не ведет ни одна дуга и существует ориентированный путь из s в любую другую вершину графа.

○ **Пример.** Найти кратчайший путь между вершинами s и t для графа, изображенного на рис. 11.19.

Окрасим вершину s, положим $d(s) = 0$, $d(x) = \infty$ для любого $x \neq s$ и $y = s$. Легко найти, что $d(a) = 6$, $d(b) = 9$, $d(c) = 5$, $d(d) = d(t) = \infty$.

Окрашиваем вершину c и дугу (s, c), так как величина $d(c)$ минимальна. Полагаем $y = c$ и, поскольку вершина t не окрашена, снова пересчитываем величину $d(x)$. Имеем: $d(a) = 6$, $d(b) = 9$, $d(d) = 10$, $d(t) = \infty$.

Окрашиваем вершину a и дугу (s, a). Полагаем $y = a$ и пересчитываем величины $d(x)$. Получаем $d(b) = 9$, $d(d) = 10$, $d(t) = \infty$.

Окрашиваем вершину b и дугу (a, b). Полагаем $y = b$, находим величины $d(x)$: $d(d) = 10$, $d(t) = 13$.

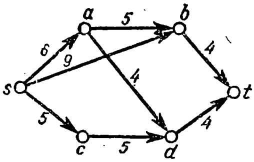


Рис. 11.19

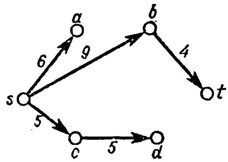


Рис. 11.20

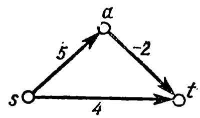


Рис. 11.21

Окрашиваем вершину d и дугу (a, d) [либо дугу (c, d)]. Полагаем $y = d$ и находим величину $d(t) = 13$.

Окрашиваем вершину t. Кратчайшим является путь (s, b, t). Покрывающее дерево кратчайших путей изображено на рис. 11.20. ●

Обобщим алгоритм на тот случай, когда некоторые дуги имеют отрицательные длины.

1. При выполнении шага 2 алгоритма пересчет величин $d(x)$ производят для всех вершин.

2. Если для некоторой окрашенной вершины величина $d(x)$ уменьшается, то с нее и с инцидентной ей дуги окраску снимают.

3. Процесс вычислений заканчивают, когда все вершины графа G окрашены и ни одно из чисел $d(x)$ не меняется при пересчете.

○ **Пример.** На рис. 11.21 изображен граф, имеющий дугу отрицательной длины. Найти кратчайший путь между вершинами s и t.

Окрашиваем вершину s. Полагаем $d(s) = 0$, $d(a) = d(t) = \infty$, $y = s$, пересчитываем величины $d(x)$: $d(a) = \min\{\infty, 0 + 5\} = 5$; $d(t) = \min\{\infty, 0 + 4\} = 4$.

Окрашиваем вершину t, дугу (s, t) и полагаем $y = t$. Пересчитываем числа $d(x)$. Поскольку из вершины t не исходит на одна дуга, эти числа не меняются.

Окрашиваем вершину a и дугу (s, a). Полагаем $y = a$

и пересчитываем $d(x)$: $d(t)=\min\{4,\ 5+(-2)\}=3$. Величина $d(t)$ уменьшилась, поэтому с вершины t и дуги $(s,\ t)$ окраску снимаем.

В исходном графе неокрашенной является только вершина t. Она окрашивается вместе с дугой $(a,\ t)$, поскольку $y=a$.

Полагая $y=t$, пересчитываем величины $d(x)$, но они не меняются. Кратчайшим является путь $(s,\ a,\ t)$. ●

Описанный алгоритм можно применять к задаче поиска кратчайших путей между каждой парой вершин графа. Однако эта процедура требует больших вычислений и обычно для решения применяют специальный алгоритм.

11.12. Алгоритм построения деревьев

Пусть имеется некоторый граф G и каждому его ребру $(x,\ y)$ поставлено в соответствие число $m(x,\ y)$, которое называется его *весом*.

Вес дерева определяется как сумма весов составляющих его ребер. Для графа G необходимо построить покрывающее дерево минимального веса.

Алгоритм построения минимального покрывающего дерева

Просматривают ребра графа G в порядке возрастания их весов. Если ребро включается в покрывающее дерево, то его окрашивают в голубой цвет, в противном случае — в оранжевый цвет. Ребра, включенные в дерево, образуют граф, состоящий из нескольких компонент. Если концевые вершины просматриваемого ребра принадлежат одной и той же компоненте, то ребро образует цикл с ребрами, ранее включенными в дерево. Такое ребро не включают в дерево. В противном случае ребро включают в дерево. Вершины одной связной компоненты составляют букет.

1. Берут любое ребро, не являющееся петлей. Окрашивают его в голубой цвет, а его концевые вершины включают в первый букет.

2. Выбирают неокрашенное ребро.

Если его концевые вершины принадлежат одному и тому же букету, то окрашивают ребро в оранжевый цвет.

Если ни одна из его концевых вершин не принадлежит ни одному из букетов, то включают их в новый букет и окрашивают ребро в голубой цвет.

Если концевые вершины принадлежат разным букетам, то объединяют эти букеты в один и окрашивают ребро в голубой цвет.

Если один конец ребра принадлежит некоторому букету, а второй не входит ни в один букет, то нужно включить второй конец в тот же букет и окрасить ребро в голубой цвет.

3. Заканчивают процедуру, если все вершины графа вошли в один букет. В противном случае переходят к шагу 2.

Число шагов при выполнении алгоритма конечно, так как оно не превышает числа ребер графа. Если голубые ребра не образуют покрывающего дерева, то у исходного графа его нет.

○ **Пример.** В новом районе имеется шесть жилых массивов. Нужно соединить их между собой дорогами, стоимость прокладки которых была бы минимальна. В следующей таблице приведены стоимости постройки дорог между каждой парой жилых массивов:

	a	b	c	d	e	f
a	—	12	1	13	11	10
b	12	—	14	3	5	12
c	1	14	—	16	17	15
d	13	3	16	—	5	7
e	11	5	17	5	—	6
f	10	12	15	7	6	—

Решение задачи запишем в виде следующей таблицы:

Ребро	Цвет	Букет 1	Букет 2
(a, c)	Голубой	a, c	
(b, d)	»	a, c	b, d
(d, e)	»	a, c	b, d, e
(b, e)	Оранжевый	a, c	b, d, e
(e, f)	Голубой	a, c	b, d, e, f
(d, f)	Оранжевый	a, c	b, d, e, f
(a, f)	Голубой	a, b, c, d, e, f	

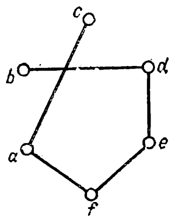


Рис. 11.22

Построено покрывающее дерево, вес которого равен 25 (рис. 11.22). ●

11.13. Задачи сетевого планирования

Сетевое планирование применяют для организации и составления календарных планов реализации больших комплексов работ. Это, например, научно-исследовательские работы с участием нескольких институтов, разработка автоматизированной системы бухгалтерского учета, строительство большого объекта и т. д. Управление всеми этими работами можно осуществлять с помощью метода критического пути. Использование этого метода позволяет сравнительно просто выяснить, когда необходимо начинать и заканчивать выполнение отдельных операций, как задержка хода выполнения некоторой операции влияет на время завершения всего проекта.

Для использования метода критического пути нужно прежде всего разбить крупный проект на отдельные операции и составить перечень операции. Некоторые из них могут выполняться одновременно, другие — только в определенном порядке. Например, при строительстве дома нельзя возводить стены раньше, чем сделан фундамент. Необходимо выяснить очередность выполнения всех операций списка. Для этого составляют список операций, непосредственно предшествующих каждой операции. После этого нужно запланировать время, необходимое для выполнения каждой операции. Полученные данные обычно помещаются в таблицу. В табл. 11.1 приведены данные для проекта, состоящего из шести работ. Для каждой из них задана продолжительность и указаны непосредственно предшествующие ей операции.

При построении графа каждую операцию изображают в виде ориентированной дуги. Связи между операциями также представляются в виде дуги. Дугу-связь проводят из конца дуги, соответствующей предшествующей операции, в начало следующей операции. Чтобы отличить операции от связей, операции изображают сплошными линиями, а связи — пунктирами. Вершины графа называют событиями. Временем наступления события считают время, когда завершено выполнение всех операций, входящих в соответствующую вершину.

Граф, представляющий взаимосвязь отдельных работ проекта, называется сетевым графиком. На рис. 11.23

построен сетевой график для комплекса операций, задаваемых табл. 11.1.

Таблица 11.1

Операция	Предшествующие операции	Время
a_1	—	10
a_2	—	5
a_3	—	15
a_4	a_1a_2	18
a_5	a_2a_3	19
a_6	a_4a_5	18

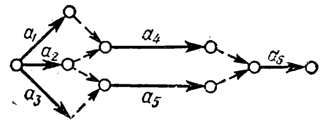


Рис. 11.23

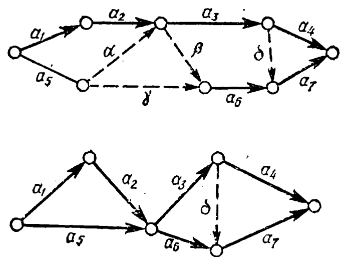


Рис. 11.24

Большое количество дуг в сети усложняет решение задачи. Поэтому прежде всего нужно упростить полученную сеть. Для этого можно выбросить некоторые дуги-связи. Начало и конец выбрасываемой дуги объединяют в одну вершину. При этом нужно проверять, не нарушится ли порядок выполнения операций после выбрасывания дуги. Проверку проводят по таблице, задающей проект.

○ **Пример.** На рис. 11.24 изображены сеть G и упрощенная сеть G_1. При упрощении выброшены дуги α, β, γ. Последовательность выполнения работ при этом не изменилась. Дугу δ выбросить нельзя, так как после этого дуги — работы a_4 и a_7 — будет неразличимы.●

Сетевой график не может содержать циклов. Если предположить, что имеется некоторый цикл $(a_1, a_2, \ldots, a_{n-1}, a_n, a_1)$, то операция a_1 может быть выполнена только после завершения операции a_n, операция a_n только после завершения операции $a_{n-1}, \ldots$, операция a_2—только после завершения операции a_1. В этом случае проект никогда не может быть выполнен.

Сеть может содержать несколько начальных вершин (таких вершин, в которые не входит ни одна дуга). В этом случае можно добавить еще одну вершину и провести из нее дуги во все начальные вершины. Тогда сеть будет иметь одну начальную вершину. Аналогично вводят конечную вершину (вершина, из которой не выходит ни одна дуга).

После построения сетевого графика нумеруют его вершины. Нумерацию, при которой номер начала любой дуги меньше номера ее конца, называют правильной.

Алгоритм получения правильной нумерации вершин

1. Нумеруют все начальные вершины

2. Вычеркивают все дуги, выходящие из начальных вершин. При этом получают новые начальные вершины. Переходят к шагу 1.

Процесс повторяют до тех пор, пока все вершины не будут перенумерованы. Конечная вершина получает при этом наибольший номер.

Временные параметры сетевого графика

Предположим, что выполнение работы начато в момент времени $t=0$. Пусть t_{ij} — заданная продолжительность работы (p_i, p_j). Величины t_{ij} записывают на соответствующих дугах сетевого графика и считают их длинами.

Ранним сроком начала работы называют наименьшее допустимое время, когда работа может быть начата.

Если из вершины p_i выходит несколько работ, то ранние сроки начал этих работ совпадают и называются ранним сроком наступления события p_i. Ранний срок начала работы (p_i, p_j) обозначают $t_{ij}^{\text{рн}}$, а ранний срок наступления события p_i — $T_i^{\text{р}}$. Для удобства величины $T_i^{\text{р}}$ записывают в верхней трети каждой вершины (рис. 11.25).

Если работа начата в ранний срок начала, то время ее окончания называется *ранним сроком окончания работы.* Ранний срок окончания работы (p_i, p_j) обозначается $t_{ij}^{\text{ро}}$.

Для вычисления ранних сроков наступления событий используют алгоритм Форда. Считают, что нумерация вершин является правильной.

Алгоритм расчета ранних сроков начал и окончаний работ

1. Полагают $T_1^{\text{р}}=0$.

2. Для $i=2, 3, \ldots, n$ вычисляют

$$T_i^{\text{р}} = \max_{k:\ (p_k,\ p_i) \in B(p_i)} (T_k^{\text{р}} + t_{ki}).$$

Номер k той вершины, при движении из которой получено значение $T_i^{р}$, заносят в левую треть вершины p_i (рис. 11.25).

После нахождения величин $T_i^{р}$ можно подсчитать ранние сроки начал и окончаний работ:

$$t_{ij}^{рн} = T_i^{р}, \quad t_{ij}^{ро} = t_{ij}^{рн} + t_{ij}.$$

○ **Пример.** Найти ранние сроки начал и окончаний работ для сети, изображенной на рис. 11.26.

Полагаем $T_1^{р} = 0$. После этого рассматриваем вершины в порядке их номеров; $T_2^{р} = T_1^{р} + t_{12} = 0 + 2 = 2$. В левую треть вершины p_2 ставим номер вершины p_1; $T_3^{р} =$

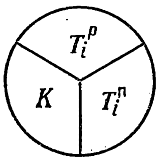


Рис. 11.25

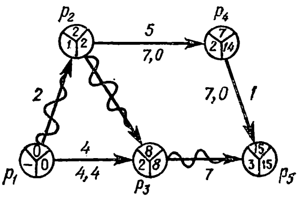


Рис. 11.26

$= \max(T_2^{р} + t_{23}, T_1^{р} + t_{13}) = \max(8, 4) = 8$. В левую треть вершины p_3 записываем номер вершины p_2 (так как при движении из p_2 получено значение $T_3^{р}$); $T_4^{р} = T_2^{р} + t_{24} = 2 + 5 = 7$, $T_5^{р} = \max(T_4^{р} + t_{45}, T_3^{р} + t_{35}) = \max(7 + 1{,}8 + 7) = 15$. После этого находим ранние сроки начал и окончаний работ: $t_{12}^{рн} = 0$, $t_{13}^{рн} = 0$, $t_{23}^{рн} = 2$, $t_{24}^{рн} = 2$, $t_{35}^{рн} = 8$, $t_{45}^{рн} = 7$, $t_{12}^{ро} = 2$, $t_{13}^{ро} = 4$, $t_{23}^{ро} = 8$, $t_{24}^{ро} = 7$, $t_{35}^{ро} = 15$, $t_{45}^{ро} = 8$. ●

Критическое время и критический путь. Ранний срок наступления конечного события называется критическим временем и обозначается $T_{кр}$. Весь проект не может быть завершен раньше момента времени $T_{кр}$, т. е. критическое время — это минимальный срок окончания всего комплекса работ. На сетевом графике $T_{кр}$ — это длина пути наибольшей длины из начальной вершины в конечную.

Всякий путь длины равной $T_{кр}$ из начальной вершины в конечную называется *критическим путем*.

Алгоритм построения критического пути

Начинают построение с конечной вершины. В ее левой трети стоит номер той вершины, при движении из которой определялся ранний срок наступления события. Критический путь идет из конечной вершины в вершину с этим номером; затем в вершину, номер которой стоит в левой трети полученной при движении вершины, и так до начальной вершины.

Если в какой-то вершине стоят два номера, то критический путь распадается на два. Таким образом, критических путей может быть несколько.

Для сети, изображенной на рис. 11.26, критический путь выделен волнистой линией.

Всякий некритический путь короче критического. Поэтому при выполнении работ, лежащих на этом пути, можно допустить задержку времени, которая не превышает разности между критическим временем и длиной пути. Такая задержка не влияет на срок выполнения всего проекта. Любая задержка выполнения работ, лежащих на критическом пути, вызывает такую же задержку выполнения всей работы.

Поздние сроки начал и окончаний работ. Задают время T выполнения всего комплекса работ. Очевидно, что должно выполняться неравенство $T \geqslant T_{\text{кр}}$. Обычно берут $T = T_{\text{кр}}$.

Поздним сроком окончания работы называется наибольшее допустимое время окончания работы без нарушения срока завершения всего проекта. Поздний срок окончания работы (p_i, p_j) обозначается $t_{ij}^{\text{по}}$. Можно определить поздний срок начала работы (p_i, p_j) $t_{ij}^{\text{пн}}$ по формуле

$$t_{ij}^{\text{пн}} = t_{ij}^{\text{по}} - t_{ij}.$$

Поздним сроком $T_j^{\text{п}}$ наступления событич p_j называется наиболее поздний срок окончания всех работ, входящих в соответствующую вершину.

Алгоритм вычисления поздних сроков наступления событий

1. Полагают $T_n^{\text{п}} = T$.

2. Для $j = n-1, n-2, \ldots, 2, 1$ вычисляют $T_j^{\text{п}} = \min\limits_{k:\ (p_j,\ p_k) \in A(j)} (T_k^{\text{п}} - t_{jk})$.

Таким образом, для конечной вершины поздний срок наступления событий совпадает со временем выполнения всего проекта. Затем просматривают все вершины в порядке убывания их номеров. Для каждой вершины рассматривают множество всех выходящих работ. Из поздних сроков наступления их концов вычитают продолжительность этих работ. Минимальная из этих разностей и равна $T_j^{\text{п}}$. Величину $T_j^{\text{п}}$ записывают для удобства вычислений в правой трети вершины p_j (рис. 11.25).

○ **Пример.** Положим для сети, изображенной на рис. 11.26, время окончания всего комплекса работ $T = T_{кр} = 15$ и поставим это значение в правую треть вершины p_5. Перейдем к событию p_4: $T_4^п = T_5^п - t_{45} = 15 - 1 = 14$. Аналогично находим $T_3^п = 15 - 7 = 8$.

Из вершины p_2 выходят две работы, поэтому $T_2^п = \min(T_3^п - t_{23},\ \ T_4^п - t_{24}) = \min(8-6,\ \ 14-5) = 2$. Аналогично получаем $T_1^п = 0$. ●

Из алгоритма вычисления поздних сроков следует, что увеличение наиболее позднего срока окончания проекта T на t единиц ведет к увеличению поздних сроков наступления всех событий также на t единиц.

После определения $T_j^п$ можно вычислить поздние сроки начала и окончаний всех работ проекта: $t_{ij}^{по} = T_j^п$, $t_{ij}^{пн} = t_{ij}^{по} - t_{ij}$.

Резервы времени. Рассмотрим некоторую работу $(p_i,\ p_j)$. Найдем время, которое можно выделить для выполнения этой работы без задержки срока окончания всего проекта. Работа $(p_i,\ p_j)$ не может быть начата раньше срока $T_i^р$ и должна быть закончена не позднее времени $T_j^п$. Для выполнения этой работы нужно затратить не более $T_j^п - T_i^р$ единиц времени. По плану эту работу можно сделать за t_{ij} единиц времени.

Максимально допустимое время, на которое можно увеличить продолжительность выполнения работы $(p_i,\ p_j)$ или отложить начало так, что это не вызовет задержки выполнения всего проекта, называется полным резервом времени.

Полный резерв времени работы $(p_i,\ p_j)$ обозначают R_{ij}, он равен

$$R_{ij} = T_j^п - T_i^р - t_{ij}.$$

Если полный резерв времени некоторой работы равен нулю, то задержка ее выполнения вызовет такую же по времени задержку выполнения всего проекта.

Если на некоторой работе использовать ее полный резерв, то путь, проходящий через эту работу, станет критическим. Полный резерв времени любой работы на этом пути станет равным нулю.

Найдем время, которое можно дополнительно выделить для выполнения работы $(p_i,\ p_j)$ без введения дополнительных ограничений на время выполнения последующих работ. Для этого выполнение работы должно быть закончено к моменту времени $T_j^р$. Таким образом можно выделить $T_j^{\text{р}} - T_i^{\text{р}}$ единиц времени на выполнение работы (p_i, p_j).

Величина $r_{ij} = T_j^{\text{р}} - T_i^{\text{р}} - t_{ij}$ называется *свободным резервом времени работы* (p_i, p_j). Если использовать свободный резерв на некоторой операции, то последующие работы могут быть по-прежнему начаты в свои ранние сроки.

Резервы времени удобно рассчитывать по сетевому графику, так как величины $T_i^{\text{р}}$, $T_i^{\text{п}}$ записаны в его вершинах. Полученные значения резервов записывают около соответствующих дуг сетевого графика. Сначала ставят полный резерв, а затем свободный.

Для сети, изображенной на рис. 11.26, имеем:

$$R_{24} = T_4^{\text{п}} - T_2^{\text{р}} - t_{24} = 14 - 2 - 5 = 7,$$
$$r_{24} = T_4^{\text{р}} - T_2^{\text{р}} - t_{24} = 7 - 2 - 5 = 0,$$
$$R_{45} = T_5^{\text{п}} - T_4^{\text{р}} - t_{45} = 15 - 7 - 1 = 7,$$
$$r_{45} = T_5^{\text{р}} - T_4^{\text{р}} - t_{45} = 15 - 7 - 1 = 7,$$
$$R_{13} = T_3^{\text{п}} - T_1^{\text{р}} - t_{13} = 8 - 0 - 4 = 4,$$
$$r_{13} = T_3^{\text{р}} - T_1^{\text{р}} - t_{13} = 8 - 0 - 4 = 4.$$

Если поздние сроки найдены при $T = T_{\text{кр}}$, то для любой вершины, лежащей на критическом пути, $T_j^{\text{р}} = T_j^{\text{п}}$, $T_j^{\text{р}} = T_i^{\text{р}} + t_{ij}$. Следовательно, $R_{ij} = r_{ij} = 0$.

РАЗДЕЛ XII. ИНТЕРПОЛЯЦИЯ

12.1. Задача интерполяции

Задачей интерполирования является построение такой функции, которая для данных значений аргумента принимала бы данные значения. Пусть для значений аргумента $x_0, x_1, x_2, \ldots, x_n$, которые называются *узлами интерполяции*, вычислены значения некоторой функции $f(x)$:

$$f(x_0)=y_0,$$
$$f(x_1)=y_1, \ldots,$$
$$f(x_n)=y_n.$$

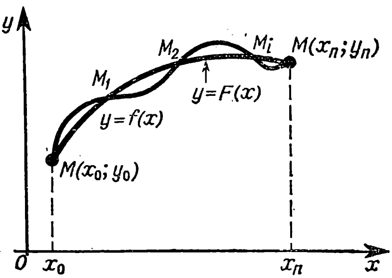


Рис. 12.1

Требуется построить функцию $F(x)$ (интерполирующую функцию), принимающую в узлах интерполяции те же значения, что и $f(x)$, т. е. такую, что

$$F(x_0)=y_0, \ F(x_1)=y_1, \ \ldots, \ F(x_n)=y_n.$$

Существуют различные варианты записи интерполяционных многочленов (см. пп. 12.3; 12.4; 12.5). Геометрически интерполирование означает, что нужно найти кривую $y=F(x)$ (рис. 12.1), проходящую через заданную систему точек $M(x_i;\ y_i)$.

12.2. Конечные разности

Пусть функция $y=f(x)$ задана в точках $x_i=x_0+kh$ (h — постоянная, k — целое). Тогда

$$\Delta y_i=\Delta f_i=f(x_{i+1})-f(x_i)$$

называются *конечными разностями первого порядка.*

Разности первых разностей образуют конечные разности второго порядка и обозначаются

$$\Delta^2 y_i = \Delta^2 f_i = \Delta f_{i+1} - \Delta f_i = f(x_{i+2}) - 2f(x_{i+1}) + f(x_i).$$

Так же определяются и разности более высоких порядков:

$$\Delta^n y_i = \Delta^n f_i = \Delta^n f_{i+1} - \Delta^{n-1} f_i.$$

Индексы при Δy берутся те же, что и индексы у вычитаемых, т. е. вторых членов разностей.

Конечные разности различных порядков располагают в форме таблиц двух видов: горизонтальной таблицы разностей (табл. 12.1) или диагональной таблицы разностей (табл. 12.2).

Таблица 12.1

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
x_0	y_0	Δy_0	$\Delta^2 y_0$	$\Delta^3 y_0$	$\Delta^4 y_0$
x_1	y_1	Δy_1	$\Delta^2 y_1$	$\Delta^3 y_1$	
x_2	y_2	Δy_2	$\Delta^2 y_2$		
x_3	y_3	Δy_3			
x_4	y_4				

Таблица 12.2

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
x_0	y_0				
		Δy_0			
x_1	y_1		$\Delta^2 y_0$		
		Δy_1		$\Delta^3 y_0$	
x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$
		Δy_2		$\Delta^3 y_1$	
x_3	y_3		$\Delta^2 y_2$		
		Δy_3			
x_4	y_4				

В табл. 12.1 и 12.2 всякое число (кроме находящихся в первых двух столбцах) является разностью двух чисел предыдущего столбца.

Вычисления разностей следует прекращать, если все числа некоторого столбца оказываются почти равными между собой («разности постоянны»). Разности k-го порядка постоянны для многочлена k-й степени. Поэтому приблизительное их постоянство показывает, что данная функция может быть с достаточной точностью изображена многочленом k-й степени.

12.3. Интерполяционная формула Лагранжа

Пусть поставлена задача о построении функции такой, что $L(x_k) = f(x_k)$, $k = 0, 1, 2, \ldots, n$. Часто такие функции строят в виде обыкновенных многочленов

$$L(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n.$$

Чтобы найти коэффициенты a_i интерполирующего многочлена возможно низшей степени, принимающей в точках $x_0, x_1, \ldots, x_n$ заданные значения, нужно решить систему уравнений относительно a_i:

$$\begin{cases} a_0+a_1x_0+a_2x_0^2+\ldots+a_nx_0^n=y_0, \\ a_0+a_1x_1+a_2x_1^2+\ldots+a_nx_1^n=y_1, \\ a_0+a_1x_2+a_2x_2^2+\ldots+a_nx_2^n=y_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_0+a_1x_n+a_2x_n^2+\ldots+a_nx_n^n=y_n, \end{cases} \tag{12.1}$$

где $y_i=f(x_i)$ $(i=0,\ 1,\ 2,\ \ldots,\ n)$.

Многочлен, коэффициенты которого определяются из системы (12.1), называется *интерполяционным многочленом Лагранжа* $L_n(x)$ для функции $f(x)$ и может быть записан в виде

$$L_n(x)=\sum_{i=0}^{n}\frac{(x-x_0)(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)}{(x_i-x_0)(x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)}y_i.$$

Интерполяционная формула Лагранжа при $n=2$ имеет вид

$$L_2(x)=y_0\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}+y_1\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}+$$
$$+y_2\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)};$$

при $n=3$ — вид

$$L_3(x)=y_0\frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}+y_1\frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}+$$
$$+y_2\frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}+y_3\frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}.$$

○ **Пример.** Построить многочлен наименьшей степени, принимающий в данных точках $x_0=1$, $x_1=3$, $x_2=6$ значения функции $y_0=10$, $y_1=16$, $y_2=4$.

Подставляя данные значения в интерполяционную формулу $L_2(x)$, имеем

$$L_2(x)=10\frac{(x-3)(x-6)}{(1-3)(1-6)}+16\frac{(x-1)(x-6)}{(3-1)(3-6)}+4\frac{(x-1)(x-3)}{(6-1)(6-3)},$$
$$L_2(x)=2{,}8-8{,}6x-1{,}4x^2. \ ●$$

Заметим, что если функция $f(x)$ задана аналитически и имеет в рассматриваемом интервале достаточное число непрерывных производных, то погрешность, получающаяся от замены $f(x)$ интерполяционным многочленом Лагранжа, равна

$$f(x)-L_n(x)=\frac{(x-x_0)(x-x_1)\dots(x-x_n)}{(n+1)!}f^{(n+1)}(\xi),$$

где ξ — некоторое промежуточное значение между наибольшим и наименьшим из чисел x, x_0, x_1, ..., x_n.

12.4. Интерполяционные формулы Ньютона

В том случае, когда интерполяционные узлы находятся на равном расстоянии, опираясь на понятие конечных разностей, интерполяционный многочлен можно найти по следующей формуле:

$$P_n(x)=y_0+q\Delta y_0+\frac{q(q-1)}{2!}\Delta^2 y_0+\dots$$
$$\dots+\frac{q(q-1)\dots(q-n+1)}{n!}\Delta^n y_0, \tag{12.2}$$

где $(x-x_0)/h=q$, $h=x_{i+1}-x_i$ — шаг заданной таблицы значений $f(x)$. Формула (12.2) называется *интерполяционной формулой Ньютона*, ее используют для интерполирования при значениях аргумента, расположенных в начале таблицы.

При $n=1$ имеем формулу линейного интерполирования

$$y\approx y_0+\frac{y_1-y_0}{h}(x-x_0),$$

а при $n=2$ имеем формулу квадратичного интерполирования

$$y\approx y_0+\frac{y_1-y_0}{h}(x-x_0)+\frac{y_2-2y_1+y_0}{2h^2}(x-x_0)(x-x_1).$$

Погрешность приближения функции интерполяционным многочленом Ньютона степени n (12.2) вычисляют по формуле

$$R_n(x)=f(x)-P_n(x)\approx\frac{\Delta^{n+1}y_0}{(n+1)!}q(q-1)\dots(q-n). \tag{12.3}$$

В частности, для линейной интерполяции ($n=1$) формула погрешности (12.3) имеет вид

$$R_1(x)\approx\frac{\Delta^2 y_0}{2!}q(q-1).$$

Для интерполирования в конце таблицы удобно использовать вторую формулу Ньютона, а именно:

$$P_n(x)=y_n+q\,\Delta y_{n-1}+\frac{q(q+1)}{2!}\Delta^2 y_{n-2}+$$
$$+\frac{q(q+1)(q+2)}{3!}\Delta^3 y_{n-3}+\ldots+\frac{q(q+1)(q+2)\ldots(q+n-1)}{n!}\Delta^n y_0. \tag{12.4}$$

При применении формулы Ньютона (12.2) удобнее использовать горизонтальную таблицу разностей (см. табл. 12.1), так как тогда необходимые значения разностей функции находятся в соответствующей горизонтальной строке таблицы.

Для формулы (12.4) составляют диагональную таблицу разностей (см. табл. 12.2).

Если из таблицы разностей будет обнаружено, что k-е разности функции для равноотстоящих значений аргумента постоянны, то интерполяционную формулу Ньютона (12.2) можно использовать в качестве эмпирической формулы, а вычисление разностей прекратить.

○ **Пример 1.** Построить эмпирическую формулу для функции y, заданной таблицей

x_i	0	1	2	3	4
y_i	13	20	6	31	35

Составим следующую таблицу разностей для заданных x_i и y_i:

x_i	y_i	Δy	$\Delta^2 y$
0	13	7	—1
1	20	6	—1
2	6	5	—1
3	31	4	
4	35		

Из этой таблицы следует, что вторые разности $\Delta^2 y$ постоянны. Используя интерполяционную формулу Ньютона (12.2) и учитывая, что в данном примере $h=1$, $q=x$, имеем

$$y=13+7x-\frac{x(x-1)}{2!} \text{ или } 13+7{,}5x-0{,}5x^2=y.$$

Пример 2. Найти значения нфукции y при $x=1{,}05$, и $x=1{,}25$, если известно, что

x_i	1	1,1	1,2	1,3
y_i	0,84147	0,89121	0,93204	0,96356

Составим таблицу конечных разностей. Это горизонтальная таблица конечных разностей. Табличные разности записываем целыми числами в единицах последнего знака без нулей впереди.

Таблица 12.3

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
1	0,84147	4974	—901	—20
1,1	0,99121	4073	—921	
1,2	0,93204	3152		
1,3	0,96356			

Так как $x=1{,}05$ находится между 1 и 1,1, т. е. в начале табл. 12.3, то воспользуемся первой формулой Ньютона (12.2).

В данном случае $h=0{,}1$, $x_0=1$, $x=1{,}05$, $q=(x-x_0)/h=$ $=0{,}5$. Далее имеем

$$y=0{,}84147+0{,}5\cdot 0{,}04974+\frac{0{,}5\,(0{,}5-1)}{2}(-0{,}00901)+$$
$$+\frac{0{,}5\,(0{,}5-1)\,(0{,}5-2)}{6}(-0{,}00020)=0{,}86746,$$
$$y\,(1{,}05)=0{,}86746.$$

Определим теперь $y\,(1{,}25)$. Так как $x=1{,}25$ заключено между 1,2 и 1,3, т. е. находится в конце таблицы, то для вычислений пользуемся второй формулой Ньютона (12.4). Таблицу конечных разностей в этом случае удобнее записывать в виде таблицы разностей (табл. 12.4).

Таблица 12.4

x	y	Δy	$\Delta^2 y$	$\Delta^3 y$
1	0,84147			
		4974		
1,1	0,89121		—901	
		4073		—20
1,2	0,93204		—921	
		3152		
1,3	0,96356			

В данном случае $h=0,1$, $x_0=1,3$, $x=1,25$, $q=(x-x_0)/h=0,5$. Далее имеем

$$y(1,25)=0,96356-0,5\cdot 0,03152+$$
$$+\frac{-0,5(-0,5+1)}{2}(-0,00921)+$$
$$+\frac{-0,5(-0,5+1)(-0,5+2)}{6}(-0,0020)=0,94678,$$

$y(1,25)=0,94678$.

12.5. Интерполяционные формулы Стирлинга и Бесселя

При интерполировании значений функции, находящихся в середине таблицы, для значений x, близких к x_k, используют формулу Стирлинга при $|q|\leqslant 0,25$ и формулу Бесселя при $0,25\leqslant q\leqslant 0,75$.

В этом случае исходное значение функции обозначают через y_0 и считают индексы вниз и вверх от нуля, как показано в табл. 12.5.

Таблица 12.5

Индекс	Аргумент	Функция						
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
—4	x_{-4}	y_{-4}	⋮	⋮	⋮	⋮	⋮	⋮
			Δy_{-4}					
—3	x_{-3}	y_{-3}		$\Delta^2 y_{-4}$				
			Δy_{-3}		$\Delta^3 y_{-4}$			
—2	x_{-2}	y_{-2}		$\Delta^2 y_{-3}$		$\Delta^4 y_{-4}$		
			Δy_{-2}		$\Delta^3 y_{-3}$		$\Delta^5 y_{-4}$	
—1	x_{-1}	y_{-1}		$\Delta^2 y_{-2}$		$\Delta^4 y_{-3}$		$\Delta^6 y_{-4}$
			Δy_{-1}		$\Delta^3 y_{-2}$		$\Delta^5 y_{-3}$	
0	x_0	y_0		$\Delta^2 y_{-1}$		$\Delta^4 y_{-2}$		$\Delta^6 y_{-3}$
			Δy_0		$\Delta^3 y_{-1}$		$\Delta^5 y_{-2}$	
1	x_1	y_1		$\Delta^2 y_0$		$\Delta^4 y_{-1}$		$\Delta^6 y_{-2}$
			Δy_1		$\Delta^3 y_0$		$\Delta^5 y_{-1}$	
2	x_2	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$		
			Δy_2		$\Delta^3 y_1$			
3	x_3	y_3		$\Delta^2 y_2$				
			Δy_3					
4	x_4	y_4						

Формула Стирлинга имеет вид

$$P(x)=y_0+q\frac{\Delta y_{-1}+\Delta y_0}{2}+\frac{q^2}{2}\Delta^2 y_{-1}+\frac{q(q^2-1)}{3!}\frac{\Delta^3 y_{-2}+\Delta^3 y_{-1}}{2}+$$
$$+\frac{q^2(q^2-1)}{4!}\Delta^4 y_{-2}+\frac{q(q^2-1)(q^2-2^2)}{5!}\frac{\Delta^5 y_{-3}+\Delta^5 y_{-2}}{2}+$$
$$+\frac{q^2(q^2-1)(q^2-2^2)}{6!}\Delta^6 y_{-3}+\dots$$
$$\dots+\frac{q^2(q^2-1)(q^2-2^2)\dots(q^2-(n-1)^2)}{(2n)!}\Delta^{2n} y_{-n},$$

где $q=(x-x_0)/h$.

Формула Бесселя имеет вид

$$P(x)=\frac{y_0+y_1}{2}+\left(q-\frac{1}{2}\right)\Delta y_0+\frac{q(q-1)}{2}\frac{\Delta^2 y_{-1}+\Delta^2 y_0}{2}+$$
$$+\frac{(q-1/2)(q(q-1))}{3!}\Delta^3 y_{-1}+\frac{q(q-1)(q+1)(q-2)}{4!}\cdot\frac{\Delta^4 y_{-2}+\Delta^4 y_{-1}}{2}+$$
$$+\frac{(q-1/2)q(q-1)(q+1)(q-2)}{5!}\Delta^5 y_{-2}+$$
$$+\frac{q(q^2-1)(q^2-2^2)(q-3)}{6!}\cdot\frac{\Delta^6 y_{-3}+\Delta^6 y_{-2}}{2}+\dots$$
$$\dots+\frac{q(q^2-1)(q^2-2^2)\dots(q-n)(q+n-1)}{2n!}\cdot\frac{\Delta^{2n} y_{-n}+\Delta^{2n} y_{-n+1}}{2}+$$
$$+\frac{(q-1/2)q(q^2-1)(q^2-2^2)\dots(q-n)(q+n-1)}{(2n+1)!}\Delta^{2n+1} y_{-n},$$

где $q=(x-x_0)/h$.

При $q=1/2$ формула Бесселя значительно упрощается и называется *формулой интерполирования на середину:*

$$P(x)=\frac{y_0+y_1}{2}-\frac{1}{8}\frac{\Delta^2 y_{-1}+\Delta^2 y_0}{2}+\frac{3}{128}\frac{\Delta^4 y_{-2}+\Delta^4 y_{-1}}{2}-$$
$$-\frac{5}{1024}\cdot\frac{\Delta^6 y_{-3}+\Delta^6 y_{-2}}{2}+\dots$$
$$\dots+(-1)^n\frac{(1\cdot 3\cdot 5\dots(2n-1))^2}{2^{2n}(2n)!}\frac{\Delta^{2n} y_{-n}+\Delta^{2n} y_{-n+1}}{2}.$$

При $n=1$ имеем формулу квадратичной интерполяции по Бесселю

$$P(x)=y_0+q\,\Delta y_0-\frac{q(1-q)}{4}(\Delta y_1-\Delta y_{-1}).$$

12.6. Интерполирование сплайнами

Для повышения точности аппроксимации функции интерполяционными многочленами необходимо увеличивать число узлов интерполяции, что, в свою очередь, приводит к увеличению степени этих многочленов. Разбиение же заданного отрезка на несколько частей с построением на каждой части самостоятельного интерполяционного многочлена неудобно тем, что «на стыках» первая производная двух соседних интерполяционных многочленов может терпеть разрыв. Поэтому во многих задачах удобнее использовать сплайны.

Разобьем заданный отрезок $[a, b]$ на n частей точками $a=x_0<x_1<\ldots<x_n=b$. *Сплайном* $S_m(x)$ называется функция, определенная на отрезке $[a, b]$ и имеющая на нем непрерывную производную $(m-1)$-го порядка, которая на каждом частичном отрезке $[x_i, x_{i+1}]$ совпадает с некоторым многочленом степени не выше m; при этом хотя бы на одном из частичных отрезков степень многочлена точно равна m.

Сплайн, принимающий в узлах x_i те же значения f_i, что и некоторая функция $f(x)$, называется *интерполяционным*.

На практике широко применяют сплайны третьей степени (кубические сплайны $S_3(x)$). Для построения итерполяционного кубического сплайна разобьем отрезок $[a, b]$ на n равных частичных отрезков длины $h=(b-a)/n$. В этом случае кубический сплайн на отрезке $[x_i, x_{i+1}]$, $i=0, 1, \ldots, m-1$, запишется в следующем виде:

$$S_3(x)=\frac{(x_{i+1}-x)^2(2(x-x_i)+h)}{h^3}f_i+\frac{(x-x_i)^2(2(x_{i+1}-x)+h)}{h^3}f_{i+1}+$$
$$+\frac{(x_{i+1}-x)^2(x-x_i)}{h^2}m_i+\frac{(x-x_i)^2(x-x_{i+1})}{h^2}m_{i+1}, \quad (12.5)$$

где m_i, m_{i+1} — некоторые числа. При этом $S_3'(x_i)=m_i$, $S_3'(x_{i+1})=m_{i+1}$.

Кубический сплайн (12.5) на каждом из промежутков $[x_i, x_{i+1}]$ непрерывен вместе со своей первой производной всюду на $[a, b]$. Выберем величины m_i так, чтобы была непрерывна и вторая производная. Условие непрерывности второй производной в точках x_i, $i=1, 2, \ldots, n-1$, принимает вид

$$m_{i-1}+4m_i+m_{i+1}=\frac{3}{h}(f_{i+1}-f_{i-1}), \quad i=1, 2, \ldots, n-1. \quad (12.6)$$

Выражение (12.6) — система линейных алгебраических уравнений относительно m_i. Для однозначного определения m_i добавляют еще два условия. Эти условия задают в виде ограничений на значения сплайна и его производных на концах промежутка $[a, b]$ и называют *краевыми*.

Существует несколько различных видов краевых условий, из которых наиболее употребительными являются следующие:

$$\text{I. } S_3'(a)=f'(a),\quad S_3'(b)=f'(b). \qquad \text{II. } S_3''(a)=f''(a),\quad S_3''(b)=f''(b). \qquad \text{III. } S_3^{(r)}(a)=S_3^{(r)}(b),\quad r=1,\ 2.$$

Условия типа III называются *периодическими*. Их применяют в том случае, если интерполируемая функция $f(x)$ периодическая с периодом $b—a$.

В случае краевых условий типа I система уравнений для определения m_i имеет вид

$$\begin{cases} m_0=f_0', \\ m_n=f_n', \\ m_{i-1}+4m_i+m_{i+1}=(3/h)(f_{i+1}-f_{i-1}), \quad i=1,\ 2,\ \ldots,\ n-1. \end{cases}$$

Для краевых условий типа II система уравнений для определения m_i такова:

$$\begin{cases} 2m_0+m_1=3/h\,(f_1-f_0)-\dfrac{h}{2}f_0'', \\ 2m_n+m_{n-1}=3/h\,(f_n-f_{n-1})+\dfrac{h}{2}f_n'', \\ m_{i-1}+4m_i+m_{i+1}=(3/h)(f_{i+1}-f_{i-1}), \quad i=1,\ 2,\ \ldots,\ n-1. \end{cases}$$

Если $f(x)$ — периодическая функция, то, полагая $f_0=f_n$, $f_{n+1}=f_1$, $m_0=m_n$, $m_1=m_{n+1}$, можно записать следующую систему для определения m_i:

$$\begin{cases} 4m_1+m_2+m_n=(3/h)(f_2-f_0), \\ m_{i-1}+4m_i+m_{i-1}=(3/h)(f_{i+1}-f_{i-1}), \quad i=2,\ \ldots,\ n-1, \\ m_1+m_{n-1}+4m_n=(3/h)(f_1-f_{n-1}). \end{cases} \tag{12.7}$$

Матрицы систем во всех трех случаях не вырождены, поэтому системы имеют, и притом единственное, решение. Решив соответствующую заданным краевым условиям систему уравнений, находят m_i. Затем по формуле (12.5) строят сплайн на каждом частичном отрезка $[x_i,\ x_{i+1}]$.

○ **Пример.** Сплайн — интерполяция функции $f(x)=\sin x$, $n=4$. Функция $f(x)=\sin x$ — периодическая.

Воспользуемся системой уравнений (12.7):

i	x_i	f_i
0	0	0
1	$\pi/2$	1
2	π	0
3	$(3/2)\,\pi$	-1
4	2π	0

$$h=\frac{b-a}{n}=\frac{\pi}{2};$$

$$\begin{cases} 4m_1+m_2+m_4=0, \\ m_1+4m_2+m_3=-12/\pi, \\ m_2+4m_3+m_4=0, \\ m_1+m_3+4m_4=12/\pi. \end{cases}$$

Решив систему, имеем: $m_1=m_3=0$, $m_2=-3/\pi$, $m_4=3/\pi$: Подставляя m_i в формулу (12.5), получим сплайн-функцию.

$$\begin{array}{ll} -\dfrac{4}{\pi^3}x^3+\dfrac{3}{\pi}x, & 0\leqslant x\leqslant\dfrac{\pi}{2}, \\ \dfrac{4}{\pi^3}(x-\pi)^3-\dfrac{3}{\pi}(x-\pi), & \dfrac{\pi}{2}\leqslant x\leqslant\pi, \\ \dfrac{4}{\pi^3}(x-\pi)^3-\dfrac{3}{\pi}(x-\pi), & \pi\leqslant x\leqslant\dfrac{3}{2}\pi, \\ -\dfrac{4}{\pi^3}(x-2\pi)^3+\dfrac{3}{\pi}(x-2\pi), & \dfrac{3}{2}\pi\leqslant x\leqslant 2\pi. \end{array}$$

РАЗДЕЛ XIII. ЭЛЕМЕНТАРНЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ВЕРОЯТНОСТЕЙ И ТЕОРИИ СЛУЧАЙНЫХ ФУНКЦИЙ

13.1. Случайные события

Теория вероятностей занимается изучением закономерностей случайных событий и случайных величин при массовом ох появлении.

Под *случайным событием* в теории вероятностей понимается событие, которое в результате опыта (испытания) может произойти или не произойти.

События называются *несовместными*, если они не могут наблюдаться в одном и том же испытании одновременно.

Суммой событий A_1, A_2, ..., A_n называется событие, состоящее в появлении хотя бы одного из этих событий.

Произведением событий A_1, A_2, ..., A_n называется событие, состоящее в одновременном появлении всех этих событий.

13.2. Вероятность события

Количественной мерой возможности появления события является *вероятность*. Наиболее широкое распространение имеют два определения вероятности события: классическое и статистическое.

Классическое определение вероятности события связано с понятием благоприятствующего случая.

Случай называется *благоприятствующим* данному событию, если появление его влечет за собой появление этого события.

При классическом определении за вероятность события принимается отношение числа благоприятствующих случаев к общему числу равновозможных случаев

$$P(A)=\frac{m}{n},$$

где $P(A)$ — вероятность события A, m — число благоприятствующих событию A случаев, n — общее число возможных случаев.

Статистическое определение вероятности связано с понятием частоты события. *Частота события* вычисляется по формуле

$$P^*(A)=\frac{m_i}{n_i},$$

где m_i — число появления события A в серии из n_i опытов.

С увеличением числа опытов частота $P^*(A)$ во многих случаях стабилизируется около некоторой постоянной величины.

При статистическом определении вероятности за вероятность события A принимают то число, относительно которого стремится стабилизироваться частота $P^*(A)$ при увеличении числа опытов. Из определения вероятностей события следует, что

$$0 \leqslant P(A) \leqslant 1.$$

13.3. Теоремы сложения и умножения вероятностей

В задачах, использующих вероятностные количественные характеристики, приходится по вероятностям одних событий оценивать вероятности других событий. Для этого используют различные соотношения, в основе которых лежат теоремы теории вероятностей.

Теорема сложения вероятностей. *Вероятность суммы несовместных событий $A_1, A_2, \ldots, A_n$ равна сумме вероятностей этих событий*:

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

Если в единичном опыте обязательно должно произойти одно из событий $A_1, A_2, \ldots, A_n$, то такая группа событий называется *полной группой событий*. Сумма вероятностей несовместных событий, образующих полную группу, равна единице:

$$\sum_{i=1}^{n} P(A_i)=1.$$

В ряде случаев вероятности появления одних событий зависят от того, произошло другое событие или нет.

Вероятность события A, вычисленная при условии, что имело место другое событие B, называется *условной вероятностью* события A и обозначается $P(A/B)$.

Теорема умножения вероятностей. *Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную при условии, что первое событие имело место*:

$$P(AB) = P(A)\,P(B/A).$$

Если же появление одного из событий не меняет вероятности появления другого, то события называются *независимыми*. Вероятность произведения независимых событий равна произведению вероятностей каждого события:

$$P(AB) = P(A)\cdot P(B).$$

13.4. Биномиальное распределение и распределение Пуассона вероятностей события

Распределение вероятностей события A может быть задано формулой биномиального распределения

$$P_n(m) = C_n^m P^m q^{n-m},$$

где $P_n(m)$ — вероятность появления ровно m раз данного события в серии из n опытов, C_n^m — число сочетаний из n элементов по m, P — вероятность появления события A в одном опыте, $q = 1 - P$.

При большом числе опытов вычисления по этой формуле становятся громоздкими. Поэтому на практике обычно используют пуассоновское приближение к биномиальному распределению, точность которого увеличивается при увеличении числа опытов и уменьшении вероятности P. Оно задается формулой

$$P_n(m) = \frac{\lambda^m}{m!}\,\mathrm{e}^{-\lambda},$$

где λ — среднее значение числа появлений рассматриваемого события в серии опытов, представляющее собой произведение nP.

Вероятность появления m раз события A на заданном интервале времени t находят по формуле Пуассона, которая в этом случае принимает вид

$$P_m(t) = \frac{(\lambda t)^m}{m!}\,\mathrm{e}^{-\lambda t},$$

где λ — интенсивность события, т. е. среднее число событий в единицу времени.

13.5. Случайные величины

Случайной называется величина, которая в результате опыта может принимать различные заранее не известные значения.

Случайные величины можно разделить на два основных вида: дискретные и непрерывные.

Дискретной случайной величиной называется такая величина, которая может принимать любое значение из конечного или бесконечного счетного множества значений, т. е. такого множества, элементы которого могут быть занумерованы в каком-нибудь порядке и выписаны в последовательности $x_1, x_2, \ldots, x_n, \ldots$.

Непрерывной случайной величиной называется такая величина, которая может принимать любые не известные заранее значения из рассматриваемого участка или интервала.

13.6. Функция распределения и плотность распределения случайной величины

Характеристикой случайной величины служат вероятности появления различных ее значений. Для их задания используют функцию распределения вероятностей случайной величины

$$F(x) = P(X < x),$$

где $P(X < x)$ — вероятность выполнения неравенства $X < x$, рассматриваемая как функция переменной x.

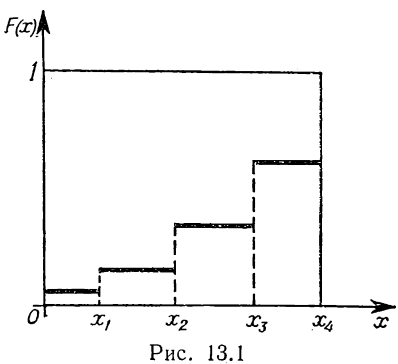


Рис. 13.1

Если X — дискретная случайная величина, возможные значения которой пронумерованы в порядке их возрастания $x_1, x_2, \ldots, x_n$, и вероятности этих значений равны соответственно $P_1, P_2, \ldots, P_n$, $P(X = x_i) = P_i$ $(i = 1, 2, \ldots, n)$, то функция распределения определяется формулой

$$F(x) = \sum_{x_i < x} P_i.$$

Эта функция изменяется ступенчато при значениях $x_1, x_2, \ldots, x_n$ (рис. 13.1).

Для непрерывной случайной величины ее функция распределения непрерывна.

Производная от функции распределения непрерывной случайной величины называется *плотностью распределения вероятности* $f(x) = F'(x)$.

В свою очередь, функция распределения через плотность распределения выражается формулой

$$F(x) = \int_{-\infty}^{\infty} f(x_1)\, dx_1.$$

13.7. Математическое ожидание случайной величины

Математическим ожиданием дискретной случайной величины называется сумма произведений всех ее значений на вероятности этих значений:

$$M[X] = M_x = \sum_{i=1}^{n} x_i P_i.$$

Для непрерывной случайной величины математическое ожидание выражается интегралом

$$M[X] = M_x = \int_{-\infty}^{\infty} x f(x)\, dx,$$

где $f(x)$ — функция плотности распределения вероятностей.

13.8. Дисперсия случайной величины

Дисперсией случайной величины X называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания

$$D[X] = D_x = M[(X - M_x)^2].$$

Для дискретной случайной величины

$$D_x = \sum_{i=1}^{n} (x_i - M_x)^2 P_i.$$

Для непрерывной случайной величины

$$D_x = \int_{-\infty}^{\infty} (x - M_x)^2 f(x)\, dx.$$

Положительный квадратный корень из дисперсии называется *средним квадратическим отклонением* случайной величины: $\sigma_x = \sqrt{D_x}$.

13.9. Векторные случайные величины

Во многих задачах приходится рассматривать совместно несколько случайных величин $X_1, X_2, \ldots, X_n$. Совокупность таких величин называют *векторной* или *многомерной случайной величиной*.

Для векторных случайных величин, так же как и для одной случайной величины, вводится понятие функции распределения вероятностей. Например, *функцией распределения вероятностей* двумерного случайного вектора с составляющими X, Y называется вероятность совместного выполнения неравенств $X < x$, $Y < y$, рассматриваемая как функция двух переменных

$$F(x, y) = P\begin{pmatrix} X < x \\ Y < y \end{pmatrix}.$$

Дискретным векторным случайным величинам соответствуют вероятности их совместного появления в опыте. Для непрерывных векторных случайных величин вводится понятие плотности вероятности случайного вектора. Так, например, *плотностью вероятности случайного вектора* (X, Y) называется предел отношения вероятности попадания его конца в бесконечно малую область к площади этой области при стягивании ее в точку:

$$f(x, y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P\begin{pmatrix} x \leqslant X \leqslant x + \Delta x \\ y \leqslant Y \leqslant y + \Delta y \end{pmatrix}}{\Delta x\, \Delta y}.$$

13.10. Корреляционный момент связи двух случайных величин. Коэффициент корреляции

Корреляционным моментом связи двух случайных величин X, Y называется математическое ожидание произведения отклонений этих величин от их математических ожиданий:

$$K_{xy} = M\,[(X - M_x)(Y - M_y)].$$

Для дискретной двумерной случайной величины корреляционный момент определяется формулой

$$K_{xy} = \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - M_x)(y_j - M_y)\, P_{ij},$$

где P_{ij} — вероятность появления величины (x_i, y_j).

Для непрерывной двумерной случайной величины корреляционный момент связи выражается интегралом

$$K_{xy}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-M_x)(y-M_y)f(x,y)\,dx\,dy.$$

Коэффициентом корреляции двух случайных величин X, Y называется число, определяемое отношением корреляционного момента к произведению средних квадратических отклонений этих величин:

$$r_{xy}=\frac{K_{xy}}{\sigma_x\sigma_y}.$$

Коэффициент корреляции изменяется от нуля до единицы. Для некоррелированных случайных величин он равен нулю.

13.11. Примеры законов распределения случайных величин

Закон равномерного распределения описывает поведение плотности вероятности случайной величины X следующим образом:

$$f(x)=\begin{cases} a\ \forall x\in[x_1;\ x_2] \\ 0\ \forall x\in\,]-\infty,\ x_1[\,\cup\,]x_2,\ +\infty[,\end{cases}$$

где a — постоянная величина.

Нормальный закон распределения, описывающий изменение плотности вероятности случайной величины, имеет вид

$$f(x)=\frac{1}{\sigma_x\sqrt{2\pi}}\,e^{-\frac{(x-M_x)^2}{2\sigma_x^2}}.$$

Закон распределения, который описывает поведение плотности вероятности и функции распределения в виде

$$f(x)=ke^{-kx},\quad F(x)=1-e^{-kx},\quad x>0,$$

называется *показательным*.

Закон распределения Рэлея для плотности и функции распределения имеет вид

$$f(x)=2h^2xe^{-h^2x^2},\quad F(x)=1-e^{-h^2x^2},\quad x>0.$$

13.12. Случайные функции. Законы распределения

Случайной называется функция, значения которой при каждом данном значении аргумента являются случайными величинами. Наиболее часто аргументом случайной функции является время.

Каждое конкретное представление случайной функции называется ее *реализацией*.

Характеристиками случайной функции являются ее законы распределения.

Одномерным законом распределения случайной функции $X(t)$ называется закон распределения ее значений при разных значениях (в сечениях) аргумента t. Обычно он задается одномерной плотностью вероятности $f(x, t)$.

Двумерным законом распределения случайной функции $X(t)$ называется закон совместного распределения ее значений при разных значениях двух независимых аргументов: t_1 и t_2. Он задается двумерной плотностью вероятности $f(x_1, x_2, t_1, t_2)$.

Закон совместного распределения значений функции в n произвольных сечениях $f(x_1, x_2, \ldots, x_n, t_1, t_2, \ldots, t_n)$ называется ее *n-мерным законом распределения*.

13.13. Математическое ожидание случайной функции

Математическим ожиданием случайной функции $X(t)$ называется такая неслучайная функция, значение которой при каждом данном значении аргумента t равно математическому ожиданию случайной функции при этом значении аргумента.

Математическое ожидание случайной функции представляет собой функцию, около которой группируются и относительно которой колеблются все возможные реализации случайной функции.

Математическое ожидание выражается через одномерную плотность вероятности следующим образом:

$$M_x(t) = M[X(t)] = \int_{-\infty}^{\infty} xf(x, t)\,\mathrm{d}x.$$

13.14. Корреляционная функция случайной функции

Корреляционной функцией случайной функции $X(t)$ называется такая неслучайная функция $K_x(t_1, t_2)$, которая для каждой пары сечений аргумента (t_1, t_2) равна соответствующему корреляционному моменту связи.

Корреляционная функция выражается через двумерную плотность вероятности $f(x_1, x_2, t_1, t_2)$ с помощью формулы

$$K_x(t_1, t_2) =$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 - M_x(t_1))(x_2 - M_x(t_2)) f(x_1, x_2, t_1, t_2)\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

При $t_1=t_2$ корреляционная функция равна дисперсии случайной функции $K_x(t_1, t_2)=D_x(t_1)$.

Нормированной корреляционной функцией случайной функции $X(t)$ называется функция

$$r_x(t_1, t_2)=\frac{K_x(t_1, t_2)}{\sigma_x(t_1)\,\sigma_x(t_2)},$$

где $\sigma_x(t)$ — среднее квадратическое отклонение случайной функции в сечении.

13.15. Каноническое разложение случайной функции

Для практического применения случайную функцию обычно представляют в виде канонического разложения.

Каноническим разложением случайной функции $X(t)$ называется представление ее в виде

$$X(t)=M_x(t)+\sum_{k=1}^{n} v_k f_k(t),$$

где v_k — центрированные*, некоррелированные случайные величины, $f_k(t)$ — неслучайные функции, называемые *координатными функциями канонического разложения.*

При каноническом разложении случайной функции ее корреляционная функция выражается в виде

$$K_x(t_1, t_2)=\sum_{k=1}^{n} D_k f_k(t_1) f_k(t_2),$$

где D_k — дисперсия случайной величины V_k.

13.16. Стационарные случайные функции

Случайная функция $X(t)$ называется *стационарной в широком смысле*, если ее математическое ожидание постоянно, а корреляционная функция зависит только от разности аргументов t_1 и t_2:

$$K_x(t_1, t_2)=K_x(\tau),$$

где $\tau=t_2-t_1$.

Случайная функция $X(t)$ называется *стационарной в узком смысле*, если ее n-мерный закон распределения при любом n зависит только от интервалов t_2-t_1, ... и совсем не зависит от положения этих интервалов в области изменения аргумента t.

В практических задачах обычно применяют понятие стационарной функции в широком смысле.

* *Центрированной случайной величиной* называется отклонение случайной величины от ее математического ожидания.

Каноническое разложение стационарной случайной функции, называемое *спектральным разложением*, имеет вид

$$X(t) = M_x + \sum_{k=0}^{n} (u_k \cos \omega_k t + v_k \sin \omega_k t),$$

где u_k, v_k—центрированные, некоррелированные случайные величины с попарно равными дисперсиями $D[u_k] = D[v_k] = D_k$.

Корреляционная функция в этом случае принимает вид

$$K_x(\tau) = \sum_{k=0}^{n} D_k \cos \omega_k \tau.$$

Стационарная случайная функция называется *эргодической*, если ее характеристики M_x и $K_x(\tau)$ могут быть определены осреднением по времени одной произвольной реализации на некотором отрезке $[0, T]$.

13.17. Марковские случайные процессы. Марковская цепь

Частным видом случайных функций являются марковские случайные процессы.

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

Состояния системы могут изменяться либо дискретно, либо непрерывно.

Случайный марковский процесс называется *процессом с дискретными состояниями*, если возможные состояния системы S_1, S_2, ..., S_n можно пронумеровать, а сам процесс состоит в том, что время от времени система S скачком (мгновенно) переходит из одного состояния в другое. Примером такого процесса является процесс, протекающий в техническом устройстве. Можно представить два состояния такой системы: S_1—система работает, S_2—система вышла из строя.

Случайный марковский процесс называется *процессом с непрерывными состояниями*, если эти состояния меняются непрерывно, постепенно. Примером такого процесса является процесс движения самолета, автомашины.

В системе с дискретными состояниями переход от состояния в состояние может происходить в определенные, фиксированные моменты времени либо в случайные моменты.

Случайный марковский процесс называется *процессом с дискретным временем*, если переходы системы из состояния в состояние возможны только в строго определенные, заранее фиксированные моменты времени t_1, t_2, ... В промежутки времени между этими моментами система S сохраняет свое состояние.

Случайный марковский процесс называется *процессом с непрерывным временем*, если переход системы из состояния в состояние возможен в любой заранее не известный случайный момент времени.

Так как для марковского процесса с дискретными состояниями и дискретным временем времена $t_1, t_2, \ldots, t_k, \ldots$ фиксированы, то процесс можно рассматривать как функцию целочисленного аргумента k ($k=1, 2, \ldots$) — номера шага. В этом случае переходы системы из состояния в состояние представляют собой последовательность (цепочку) событий или состояний $S_1^{(1)}$, $S_1^{(2)}$, $S_3^{(3)}$, $S_5^{(4)}$, $S_2^{(5)}$, ... Число в скобках обозначает номер шага, нижний индекс — номер состояния.

Случайная последовательность событий с фиксированным шагом называется *дискретной марковской цепью*, если для каждого шага вероятность перехода из любого состояния S_i в любое другое состояние S_j не зависит от того, когда и как система перешла в состояние S_i.

Если переход системы из состояния в состояние происходит в случайные моменты времени, то соответствующая цепочка состояний называется *непрерывной цепью Маркова*.

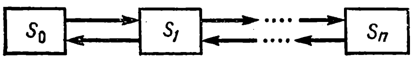


Рис. 13.2

При анализе случайных процессов с дискретными состояниями используют графы состояний. Граф состояний геометрически изображает возможные состояния системы и ее возможные переходы из состояния в состояние.

Важное место в исследовании экономических систем занимает процесс гибели и размножения.

Марковская непрерывная цепь называется *процессом гибели и размножения*, если ее граф состояний представляет собой цепочку, в которой каждое из промежуточных состояний связано прямой и обратной связью с каждым соседним состоянием (рис. 13.2).

РАЗДЕЛ XIV. МЕТОДЫ СТАТИСТИЧЕСКОГО АНАЛИЗА

14.1. Статистические испытания

Исследование свойств изучаемого объекта часто проводят с помощью метода статистических испытаний (метода Монте—Карло).

Статистическим испытанием называется специально организованное испытание натурного объекта или его модели с учетом воздействия на объект или модель случайных возмущений.

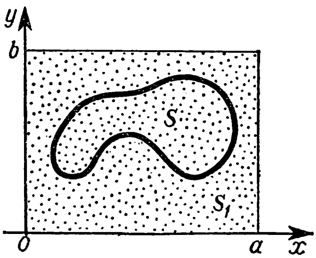


Рис. 14.1

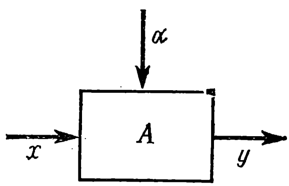


Рис. 14.2

Моделирование реальных явлений с учетом случайных воздействий иногда называют *имитацией*.

С помощью метода статистических испытаний решают задачи исследования любого реального процесса, на протекание которого влияют случайные факторы. Этот метод используют также при решении задач, для которых возможно создание искусственной вероятностной модели. Примером последнего может служить определение площади криволинейной фигуры.

○ Пусть требуется вычислить площадь фигуры S, изображенной на рис. 14.1. Ограничив искомую площадь S площадью прямоугольника S_1, находим отношения числа точек N, принадлежащих площади S, к общему числу точек N_1 в площади S_1 ($S_1 = ab$). Площадь S приближенно находим из отношения

$$\frac{S}{S_1} = \frac{N}{N_1}, \text{ откуда } S = \frac{N}{N_1} S_1.$$

При этом закон распределения точек в области S_1 должен быть равномерным. ●

Еще одним примером использования метода статистических испытаний является вычисление экстремальных значений функции. Для многоэкстремальных многопараметрических задач определение экстремумов функций с помощью последовательного перебора точек в области определения функции — сложная трудоемкая задача. Использование метода случайного выбора точек в указанной области позволяет с меньшей трудоемкостью и достаточной вероятностью находить экстремальные значения функции.

Примером использования метода статистических испытаний при исследовании случайных процессов является моделирование нестационарных систем массового обслуживания, которые не удается представить в виде стационарных линейных моделей.

При статистическом моделировании случайного процесса случайные факторы (возмущения) представляют в виде конечного набора случайных величин с известными распределениями.

Испытание модели процесса с конкретными значениями факторов называют *реализацией случайного процесса в данной модели.*

Модель статистических испытаний в общем виде можно представить следующим отношением:

$$y = A(x, \alpha),$$

где A — известный оператор преобразования, x — вектор входных неслучайных воздействий, y — вектор выходных параметров, α — вектор случайных параметров с известными законами распределения вероятностей. Условно процесс статистических испытаний модели объекта с учетом случайных воздействий изображен на рис. 14.2.

Задавая последовательность значений случайных величин в соответствии с их законами распределения и осуществляя для каждого конкретного значения этих случайных величин реализацию процесса, получают множество реализаций, которое подвергают статистической обработке для получения статистических характеристик выходных параметров объекта.

Моделирование случайной величины — основное содержание метода статистических испытаний. Исходными при моделировании случайных величин являются случайные величины, равномерно распределенные на промежутке [0, 1]. Другие виды распределений получают с помощью специальных методов преобразования такой равномерно распределенной величины.

Процесс статистических испытаний обычно осуществляют с помощью ЭВМ. Значения случайных величин для каждой реализации процесса выбирают с помощью датчиков случайных чисел, которые входят в математическое обеспечение современных ЭВМ.

14.2. Метод наименьших квадратов

Метод наименьших квадратов служит для оценки неизвестных величин по результатам измерений, содержащим случайные погрешности. Среди многих приложений метода наиболее важным является нахождение наилучшего уравнения (функциональной зависимости) определенного вида для представления опытных данных.

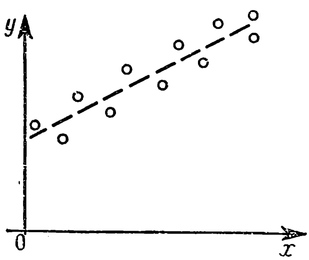


Рис. 14.3

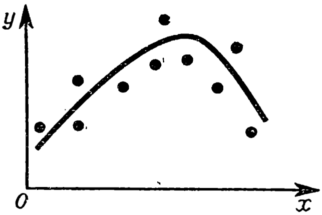


Рис. 14.4

Процесс выражения опытных данных функциональной зависимостью с помощью метода наименьших квадратов состоит из двух этапов: на первом выбирают вид искомой формулы, а на втором для данной формулы подбирают параметры. На рис. 14.3 приведены опытные данные, для которых в качестве эмпирической формулы (полученной на основании опытных данных) можно принять линейную зависимость $y = ax + b$. Для данных, приведенных на рис. 14.4, эмпирическую зависимость целесообразно принять в виде $y = ax^2 + bx + c$.

В соответствии с идеей метода наименьших квадратов необходимо минимизировать сумму

$$S=\sum_{i=1}^{n}(\bar{y}(x_i)-y_i)^2, \tag{14.1}$$

где x_i, y_i — значения опытных данных, $\bar{y}(x_i)$ — значение функции, взятое на эмпирической зависимости в точке x_i, n — число опытов.

В случае линейной эмпирической формулы сумма (14.1) принимает вид

$$S=\sum_{i=1}^{n}(ax_i+b-y_i)^2, \tag{14.2}$$

а в случае квадратической зависимости — следующий вид:

$$S=\sum_{i=1}^{n}(ax_i^2+bx_i+c-y_i)^2. \tag{14.3}$$

Минимум функции (14.2) и (14.3) имеют в тех точках, в которых частные производные от S по параметрам a, b, c обращаются в нуль. В результате дифференцирования и элементарных преобразований для определения параметров получают нормальную систему линейных уравнений. В случае линейной эмпирической зависимости составляют нормальную систему двух уравнений с двумя неизвестными a и b:

$$\begin{cases} a\sum\limits_{i=1}^{n}x_i^2+b\sum\limits_{i=1}^{n}x_i=\sum\limits_{i=1}^{n}x_iy_i, \\ a\sum\limits_{i=1}^{n}x_i+bn=\sum\limits_{i=1}^{n}y_i. \end{cases} \tag{14.4}$$

В случае квадратической зависимости нормальная система состоит из трех уравнений с тремя неизвестными:

$$\begin{cases} a\sum\limits_{i=1}^{n}x_i^4+b\sum\limits_{i=1}^{n}x_i^3+c\sum\limits_{i=1}^{n}x_i^2=\sum\limits_{i=1}^{n}x_i^2y_i, \\ a\sum\limits_{i=1}^{n}x_i^3+b\sum\limits_{i=1}^{n}x_i^2+c\sum\limits_{i=1}^{n}x_i=\sum\limits_{i=1}^{n}x_iy_j, \\ a\sum\limits_{i=1}^{n}x_i^2+b\sum\limits_{i=1}^{n}x_i+cn=\sum\limits_{i=1}^{n}y_i. \end{cases}$$

Кроме линейной и квадратической зависимости в практических задачах обработки результатов наблюдений используют и другие зависимости, например гиперболическую зависимость

$$y = a + \frac{b}{x}. \tag{14.5}$$

Для зависимости (14.5) система нормальных уравнений имеет вид

$$\begin{cases} na + b\sum_{i=1}^{n}\frac{1}{x_i} = \sum_{i=1}^{n} y_i, \\ a\sum_{i=1}^{n}\frac{1}{x_i} + b\sum_{i=1}^{n}\frac{1}{x_i^2} = \sum_{i=1}^{n}\frac{y_i}{x_i}. \end{cases}$$

○ **Пример.** Опытные данные о значениях x и y представлены в следующей таблице:

x	1	2	3	4	5	6
y	15	10	2	2	—4	—10

Анализ опытных данных показывает, что в качестве эмпирической зависимости можно использовать линейную зависимость $y = ax + b$. Найти методом наименьших квадратов значения a и b.

Коэффициенты нормальной системы уравнений находят с помощью таблицы подсчетов (табл. 14.1).

Таблица 14.1

i	x_i	y_i	x_i^2	$x_i y_i$
1	1	15	1	15
2	2	10	4	20
3	3	2	9	6
4	4	2	16	8
5	5	—4	25	—20
6	6	—10	36	—60
Σ	21	15	91	—31

Подставляя полученные в таблице данные в систему уравнений (14.4), получаем

$$\begin{cases} 91a + 21b = -31, \\ 21a + 6b = 15. \end{cases} \tag{14.6}$$

Решая систему уравнений (14.6), получаем следующие значения коэффициентов: $a \approx -4{,}76$, $b \approx 19{,}2$. Эмпирическая формула принимает вид

$$y = -4{,}76x + 19{,}2. \bullet$$

Не существует общего правила для выбора подходящего вида эмпирической формулы; можно лишь догадываться о подходящей форме уравнения по форме кривой, изображающей данные. Однако существуют способы, с помощью которых можно проверить, является ли догадка удачной или нет.

Для наиболее часто встречающихся зависимостей с двумя параметрами, а именно: I) $y = ax + b$, II) $y = ax^b$, III) $y = ab^x$, IV) $y = a + \frac{b}{x}$, V) $y = \frac{1}{ax+b}$, VI) $y = \frac{x}{ax+b}$, VII) $y = a\lg x + b$, эмпирическую формулу можно выбирать с помощью табл. 14.2.

Таблица 14.2

Номер формулы	$\bar{x}_s$	$\bar{y}_s$	Вид эмпирической формулы
I	$\frac{x_1+x_n}{2}$	$\frac{y_1+y_n}{2}$	$y = ax + b$
II	$\sqrt{x_1x_n}$	$\sqrt{y_1y_n}$	$y = ax^b$
III	$\frac{x_1+x_n}{2}$	$\sqrt{y_1y_n}$	$y = ab^x$ $y = ae^{\beta x}$, где $\beta = \ln b$
IV	$\frac{2x_1x_n}{x_1+x_n}$	$\frac{y_1+y_n}{2}$	$y = a + \frac{b}{x}$
V	$\frac{x_1+x_n}{2}$	$\frac{2y_1y_n}{y_1+y_n}$	$y = \frac{1}{ax+b}$
VI	$\frac{2x_1x_n}{x_1+x_n}$	$\frac{2y_1y_n}{y_1+y_n}$	$y = \frac{x}{ax+b}$
VII	$\sqrt{x_1x_n}$	$\frac{y_1+y_n}{2}$	$y = a\lg x + b$

Для проверки пригодности выбранной эмпирической формулы, используя исходные данные, находят значение $\bar{x}_s$ и $\bar{y}_s$. Затем сравнивают y_s, соответствующее $\bar{x}_s$ в исходных данных, со значением $\bar{y}_s$. Если $\bar{x}_s$ не находится среди исходных данных x_i, то соответствующее значение можно определить с помощью линейной интерполяции:

$$\hat{y}_s = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(\overline{x}_s - x_i),$$

где x_i и x_{i+1} — промежуточные значения, между которыми содержится $\overline{x}_s$ $(x_i < \overline{x}_s < x_{i+1})$.

Если величина $|\hat{y}_s - \overline{y}_s|$ большая, то соответствующая эмпирическая формула непригодна.

Зависимости I—VII, приведенные в табл. 14.2, монотонные и, следовательно, пригодны только в том случае, если в исходных данных $x_{i+1} - x_i > 0$, а $y_{i+1} - y_i$ обладает постоянным знаком.

○ **Пример.** Определить вид эмпирической формулы, отвечающей следующей таблице:

x	2	3	4	5	6	7	8	9	10
y	12	35	75	125	210	315	445	600	800

Таблица 14.3

Номер формул	$\overline{x}_s$	$\overline{y}_s$	$\hat{y}_s$	$\lvert \hat{y}_s - \overline{y}_s \rvert$	Вид эмпирической формулы
I	$\frac{x_1 + x_n}{2} = \frac{2+10}{2} = 6$	$\frac{y_1 + y_n}{2} = \frac{12+800}{2} = 406$	210	196	$y = ax + b$ не подходит
II	$\sqrt{x_1 x_n} = \sqrt{2 \cdot 10} = 4{,}47$	$\sqrt{y_1 y_n} = \sqrt{12 \cdot 800} = 98$	98,5	0,5	$y = ax^b$ подходит лучше других формул
III	$\frac{x_1 + x_n}{2} = 6$	$\sqrt{y_1 y_n} = 98$	210	112	$y = ab^x$ не подходит
IV	$\frac{2x_1 x_n}{x_1 + x_n} = 3{,}3$	$\frac{y_1 + y_n}{2} = 406$	47	359	$y = a + \frac{b}{x}$ не подходит
V	$\frac{x_1 + x_n}{2} = 6$	$\frac{2y_1 \cdot y_n}{y_1 + y_n} = 23{,}6$	210	186,4	$y = \frac{1}{ax + b}$ не подходит
VI	$\frac{2x_1 x_n}{x_1 + x_n} = 3{,}3$	$\frac{2y_1 y_n}{y_1 + y_n} = 23{,}6$	47	23,4	$y = \frac{x}{ax + b}$ не подходит
VII	$\sqrt{x_1 x_n} = 4{,}47$	$\frac{y_1 + y_n}{2} = 406$	98,5	307,5	$y = a \lg x + b$ не подходит

Подбор эмпирической формулы по приведенным выше критериям приведен в табл. 14.3. ●

14.3. Дисперсионный анализ

Дисперсионный анализ является одним из методов статистической обработки наблюдений и служит для оценки влияния на наблюдаемую величину различных факторных признаков, таких, от которых зависит наблюдаемая величина.

Пусть производится n измерений случайной величины y. Каждое измерение y_j ($j=1, 2, \ldots, n$) зависит от некоторого числа параметров x_{ij}, которые могут принимать или дискретные, или непрерывные значения. Эту зависимость обычно представляют в виде линейной комбинации параметров x_{ij} с коэффициентами β_i:

$$y_j = \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots + \beta_m x_{mj} + e_j, \qquad (14.7)$$

где e_j — случайная ошибка измерения.

Величины $\beta_1, \beta_2, \ldots, \beta_m$ называются *факторами*. Уравнение (14.7) называется *линейной многофакторной моделью*.

Параметры x_{ij} в дисперсионном анализе обычно принимают равными 0 или 1, что указывает на то, какие из факторов учитываются при таком анализе.

Для оценки влияния факторных признаков x_{ij} на наблюдаемую величину y_j (результативный признак) значения этой величины разбивают на несколько уровней, соответствующих определенному значению факторного признака.

Пусть, например, наблюдаются значения производительности труда на разных предприятиях. Требуется оценить влияние концентрации производства на производительность. Предприятия по концентрации производства можно разделить на следующие уровни (группы): мелкие, средние и крупные. В каждый из уровней будут входить предприятия с некоторыми конкретно наблюдаемыми значениями производительности. В этом случае наблюдаемые значения записывают с двумя индексами (y_{rj}), где r — номер уровня, j — номер измерения на каждом уровне. В данном случае $r=1, 2, 3$. В общем случае $r=1, 2, \ldots, p$, где p — число уровней.

Для однофакторного дисперсионного анализа наблюдаемые значения можно представить в виде

$$y_{rj} = \beta_r + e_{rj} \quad (r=1, 2, \ldots, p; \; j=1, 2, \ldots, q_r),$$

где β_r — среднее значение наблюдаемой величины на r-уровне.

Среднее всех наблюдаемых значений определяют по формуле

$$\beta = \frac{\sum_{r=1}^{p} \sum_{j=1}^{q_r} y_{rj}}{q_r}.$$

Затем находят групповые (уровневые) средние:

$$\beta_r = \frac{\sum_{j=1}^{q_r} y_{rj}}{q_j}.$$

Далее подсчитывают факторную дисперсию:

$$D_F = \frac{\sum_{r=1}^{p} (\beta_r - \beta)^2}{p-1} \qquad (14.8)$$

и остаточную дисперсию:

$$D_0 = \frac{\sum_{r=1}^{p} \sum_{j=1}^{q_r} (y_{rj} - \beta_r)^2}{N-p}. \qquad (14.9)$$

Наконец, составляют отношение $F = \frac{D_F}{D_0}$, которое характеризует влияние факторного признака. Чем больше влияние факторного признака на результативный, тем больше значение F.

В знаменателях выражений (14.8) и (14.9) находятся значения чисел степеней свободы* $k_1 = p-1$, $k_2 = N-p$. Для различных значений k_1 и k_2 рассчитаны распределения вероятностей $p\ (F > F_1)$, где F_1 — заданные значения отношения $\frac{D_F}{D_0}$. На основании этих распределений составлены таблицы для определения значений F_1, входом в которые являются k_1, k_2 и заданные уровни вероятности $p\ (F > F_1)$.

Например, для уровня вероятности, равного 0,05, и значений $k_1 = 2$, $k_2 = 10$ значение $F_1 = 4{,}1$. В результате расчетов с использованием выражений (14.8) и (14.9) по-

---

* *Числом степеней свободы* называется число независимых выборок, используемых при оценке соответствующей статистической характеристики.

лучено значение F_1, равное 3,2. А так как $3{,}2 < 4{,}1$, то только с вероятностью не выше чем 0,05 случайные значения величины F будут превосходить расчетное значение. Следовательно, с малой вероятностью факторный признак будет оказывать влияние на результативный признак и это влияние можно не учитывать.

14.4. Регрессионный анализ

Регрессионный анализ является методом статистической обработки наблюдений, в результате которой оказывается возможным составить уравнение регрессии и получить количественную оценку влияния факторных признаков на результативный признак.

Пусть имеется линейная многофакторная модель (14.7). Оценивая с помощью метода наименьших квадратов для уравнения (14.7) факторы $\beta_1, \beta_2, \ldots, \beta_m$, составим сумму $\beta_{10}x_{1j} + \beta_{20}x_{2j} + \ldots + \beta_{m0}x_{mj}$, где $\beta_{10}, \beta_{20}, \ldots, \beta_{m0}$ — средняя квадратическая оценка случайных факторов; $x_{1j}, x_{2j}, \ldots \ldots, x_{mj}$ — значения непрерывных переменных $x_1, x_2, \ldots, x_m$.

Уравнение

$$y_j = \beta_{10}x_{1j} + \beta_{20}x_{2j} + \ldots + \beta_{m0}x_{mj} \qquad (14.10)$$

называется *уравнением регрессии*.

Главной задачей регрессионного анализа является получение оптимальных оценок $\beta_{10}, \beta_{20}, \ldots, \beta_{m0}$, называемых коэффициентами регрессии.

Уравнения (14.7) можно записать в виде

$$y_j = x_{1j}\beta_1 + x_{2j}\beta_2 + \ldots + x_{mj}\beta_m + e_j, \quad j = 1, 2, \ldots, n, \qquad (14.11)$$

или в матричной форме

$$y = x^{\mathrm{T}}\beta + e, \qquad (14.12)$$

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \cdot \\ \cdot \\ \cdot \\ \beta_m \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{pmatrix};$$

x^{T} — матрица, транспонированная к матрице

$$x = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ x_{m1} & x_{m2} & \ldots & x_{mn} \end{pmatrix}.$$

Оценку факторов $\beta_1, \beta_2, \ldots, \beta_m$ в уравнении (14.12) на основе метода наименьших квадратов можно получить по формуле

$$\beta_0 = (x^{\mathrm{T}}x)^{-1} x^{\mathrm{T}}y,$$

где $(x^{\mathrm{T}}x)^{-1}$—матрица, обратная к матрице $x^{\mathrm{T}}x$.

14.5. Планирование эксперимента

Под *экспериментом* принимают специально подготовленное испытание объекта исследования или его модели, позволяющее получать новую информацию об объекте.

Эксперименты, связанные с организацией и управлением предприятия и отрасли народного хозяйства, называются *хозяйственными*. Эксперименты подразделяют на натурные и машинные.

Натурный хозяйственный эксперимент проводят на отдельном или нескольких предприятиях или в отраслях народного хозяйства для оценки эффективности предлагаемых нововведений с целью более широкого или длительного использования их на практике.

Машинный эксперимент использует некоторую абстрактную (математическую) модель изучаемого объекта (звена народного хозяйства), которая исследуется с помощью современных вычислительных средств.

Математический аппарат планирования эксперимента опирается на такие методы статистического анализа, как метод статистических испытаний, метод наименьших квадратов, регрессионный анализ и т. п.

Математическая постановка задачи планирования эксперимента состоит в следующем. Пусть имеется некоторый объект (предприятие, объединение и т. п.), характеризуемый определенным набором входных параметров (факторов) и выходных переменных. Входные и выходные переменные можно записать соответственно в виде векторов $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_p)$. На объект действуют возмущающие (неуправляемые) факторы $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_q)$.

В различные моменты времени объект будет находиться в различных состояниях, которые определяются статическими или динамическими зависимостями, связывающими входные и выходные переменные. Эти зависимости при проведении машинного эксперимента обычно бывают заданными. При проведении натурного эксперимента чаще всего предполагается отсутствие функциональных связей между входом и выходом, т. е. система рассматривается как «черный ящик». В этом случае требуется по известным значениям входных и выходных параметров составить функциональную зависимость между ними. Для этого выдвигают гипотезу о характере такой зависимости. В большинстве случаев предполагают, что искомая зависимость является полиномом от одной или многих переменных. Далее, составив план и проведя серию экспериментов, после соответствующей обработки результатов находят коэффициенты принятого полинома.

Рассмотренную модель записывают в виде следующего векторного уравнения регрессии:

$$y = A_0 + \sum_{i=1}^{n} A_i x_i + \sum_{i<j} A_{ij} x_i x_j + \sum_{i=1}^{n} A_{ii} x_i^2 + \ldots, \quad (14.13)$$

где векторы — коэффициенты A_0, A_i, A_{ij}, ... — определяют в результате проведения эксперимента. Тем самым оценивается влияние отдельных факторов или их совместное влияние на выход объекта.

Значения коэффициентов модели (14.13) обычно оценивают с помощью метода наименьших квадратов.

Одной из главных задач планирования эксперимента является задача составления плана проведения эксперимента, обеспечивающего наиболее достоверное определение коэффициентов модели при ограниченном числе испытаний. Одновременно должна решаться задача уменьшения трудоемкости обработки информации в методе наименьших квадратов. Эти задачи решаются путем использования оптимальных планов. В качестве такого плана часто применяют план полного факторного эксперимента ПФЭ-2. Для такого плана число опытов равно 2^n, где n — число учитываемых факторов. Исследуемые факторы c_i изменяются на двух уровнях: верхнем c_i^+ и нижнем c_i^- [2].

Литература к разделу XIV

1. *Ветров А. А., Ломовицкий Г. И.* Дисперсионный анализ в экономике. М., 1975.
2. *Вознесенский В. А.* Статистические методы планирования эксперимента в технико-экономических исследованиях. М., 1981.

РАЗДЕЛ XV. ЭЛЕМЕНТЫ ТЕОРИИ И ПРОСТЕЙШИЕ МОДЕЛИ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

15.1. Классификация систем массового обслуживания

Большинство экономических задач связано с системами массового обслуживания.

Системы, в которых, с одной стороны, возникают массовые запросы (требования) на выполнение киких-либо видов услуг, а с другой стороны, происходит удовлетворение этих запросов, называются *системами массового обслуживания*.

Система массового обслуживания включает следующие элементы: источник требований, входящий поток требований, очередь, обслуживающее устройство (обслуживающий аппарат, канал обслуживания), выходящий поток требований.

Системы массового обслуживания классифицируют по разным признакам. К таким признакам относятся условия ожидания требованием начала обслуживания. В соответствии с этим признаком системы подразделяются на следующие виды:

1) системы массового обслуживания с потерями (отказами);

2) системы массового обслуживания с ожиданием;

3) системы массового обслуживания с ограниченной длиной очереди;

4) системы массового обслуживания с ограниченным временем ожидания.

Системы массового обслуживания, у которых требования, поступающие в момент, когда все приборы обслуживания заняты, получают отказ и теряются, называются *системами с потерями или отказами*.

Системы массового обслуживания, у которых возможно появление как угодно длинной очереди требований к обслуживающему устройству, называются *системами с ожиданием*.

Системы массового обслуживания, допускающие очередь, но с ограниченным числом мест в ней, называются *системами с ограниченной длиной очереди*.

Системы массового обслуживания, допускающие очередь, но с ограниченным сроком пребывания каждого требования в ней, называются *системами с ограниченным временем ожидания*.

По числу каналов или приборов системы делятся на *одноканальные* и *многоканальные*.

По месту нахождения источника требований системы массового обслуживания делятся на *разомкнутые*, когда источник находится вне системы, и *замкнутые*, когда источник находится в самой системе. К последнему виду относится, например, станочный участок, в котором станки являются источником неисправностей, а следовательно, и требований на их обслуживание.

Одной из форм классификации систем массового обслуживания является кодовая (символьная) *классификация Д. Кендалла*. При этой классификации характеристику системы записывают в виде трех, четырех или пяти символов, например $A|B|S$, где A — тип распределения входящего потока требований, B — тип распределения времени обслуживания, S — число каналов обслуживания. Для экспоненциального распределения принимают символ M, для любого (произвольного) распределения — символ G. Запись $M|M|3$ означает, что входящий поток требований пуассоновский (простейший), время обслуживания распределено по экспоненциальному закону, в системе имеется три канала обслуживания.

Четвертый символ указывает допустимую длину очереди, а пятый — порядок отбора (приоритета) требований.

15.2. Показатели эффективности систем массового обслуживания

Показатели эффективности делятся на показатели, характеризующие качество и условия работы обслуживающей системы, и показатели, отражающие экономические особенности системы.

Показатели первой группы обычно формируют на основе полученных из расчетов значений вероятностей состояний системы. Показатели второй группы рассчитывают на основе показателей первой группы.

Среди показателей первой группы можно выделить следующие.

1) Вероятность того, что поступающее в систему требование откажется присоединяться к очереди и теряется ($P_{отк}$). Этот показатель для системы массового обслуживания с отказами равен вероятности того, что в системе находится столько требований, сколько она содержит приборов (каналов) обслуживания:

$$P_{отк} = P_m,$$

где m — число каналов обслуживания.

Для системы с ограниченной длиной очереди $P_{отк}$ равно вероятности того, что в системе находится $m+l$ требований:

$$P_{отк} = P_{m+l},$$

где l — допустимая длина очереди.

Противоположным показателем является вероятность обслуживания требования

$$P_{обсл} = 1 - P_{отк}.$$

2) Среднее количество требований, ожидающих начала обслуживания,

$$M_{ож} = \sum_{n=m+1}^{m+l} (n-m) P_n,$$

где P_n — вероятность того, что в системе находятся n требований.

При условии простейшего потока требований и экспоненциального закона распределения времени обслуживания формулы для $M_{ож}$ принимают следующий вид:

система с ограниченной длиной очереди

$$M_{ож} = \frac{P_0 \rho^m}{m!} \sum_{n=1}^{l} n \left(\frac{\rho}{m}\right)^n,$$

где $\rho = \frac{\lambda}{\mu}$, λ — интенсивность входящего потока требований (среднее число требований, поступающих в единицу времени), μ — интенсивность обслуживания (среднее число обслуженных требований в единицу времени);

система с ожиданием

$$M_{ож} = \frac{P_0 \rho^{m+1}}{m \cdot m!} \frac{1}{\left(1 - \frac{\rho}{m}\right)^2}.$$

3) Относительная (q) и абсолютная (A) пропускные способности системы. Эти величины находят соответственно по формулам

$$q = 1 - P_{\text{отк}}, \quad A = \lambda q.$$

4) Среднее число занятых обслуживанием приборов в случае экспоненциального характера потока требований и времени обслуживания

$$m_з = \rho q.$$

Для системы массового обслуживания с отказами $m_з$ можно найти по формуле

$$m_з = \sum_{n=1}^{m} n \cdot P_n.$$

5) Общее количество требований, находящихся в системе (M). Эту величину определяют следующим образом:

с и с т е м а м а с с о в о г о о б с л у ж и в а н и я с о т к а з а м и

$$M = m_з,$$

с и с т е м а м а с с о в о г о о б с л у ж и в а н и я с о г р а н и ч е н н о й д л и н о й о ч е р е д и и о ж и д а н и е м

$$M = m_з + M_{\text{ож}}.$$

6) Среднее время ожидания требованиям начала обслуживания ($T_{\text{ож}}$). Если известна функция распределения вероятности времени ожидания требованием начала обслуживания

$$F(t) = P(T_{\text{ож}} < t),$$

то среднее время ожидания находится как математическое ожидание случайной величины $T_{\text{ож}}$:

$$T_{\text{ож}} = M[T_{\text{ож}}] = \int_0^\infty t \, dF;$$

$T_{\text{ож}}$ при показательном законе распределения требований во входящем потоке можно определить по формуле

$$T_{\text{ож}} = \frac{M_{\text{ож}}}{\lambda}.$$

Показатели, характеризующие экономические особенности, формируют обычно в соответствии с конкретным видом системы и ее назначением. Одним из общих экономических показателей является *экономическая эффективность*

$$E = P_{обсл}\lambda c T - G_п,$$

где c — средний экономический эффект, полученный при обслуживании одного требования, T — рассматриваемый интервал времени, $G_п$ — величина потерь в системе.

Величину потерь можно определить по следующим формулам:

система с отказами

$$G_п = (q_к m_з + q_у P_{отк}\lambda + q_{пк} m_{св})\, T,$$

где $q_к$ — стоимость эксплуатации одного прибора в единицу времени, $q_у$ — стоимость убытков в результате ухода требований из системы в единицу времени, $q_{пк}$ — стоимость единицы времени простоя прибора системы, $m_{св} = m - m_з$;

система с ожиданием

$$G_п = (q_{ож} M_{ож} + q_{пк} m_{св} + q_к m_з)\, T,$$

где $q_{ож}$ — стоимость потерь, связанных с простоем требований в очереди в единицу времени.

15.3. Уравнения Колмогорова для вероятностей состояний

Системы, представляемые в виде непрерывной цепи Маркова, обычно исследуют с помощью уравнений Колмогорова для вероятностей состояний.

Плотностью вероятности перехода λ_{ij} из состояния S_i в состояние S_j называется предел отношения вероятности этого перехода за время Δt к длине промежутка Δt, когда последний стремится к нулю:

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

где $P_{ij}(\Delta t)$ — вероятность того, что система, находившаяся в момент t в состоянии S_i, за время Δt перейдет в состояние S_j.

Марковская непрерывная цепь называется *однородной*, если плотности вероятностей λ_{ij} не зависят от времени t, в противном случае она называется *неоднородной*.

Для однородных марковских непрерывных цепей, характеризующих процессы гибели и размножения, уравнения Колмогорова имеют вид

$$\frac{dP_0}{dt} = -\lambda_{01}P_0(t) + \lambda_{10}P_1(t),$$

$$\dots\dots\dots\dots\dots$$

$$\frac{dP_i}{dt} = \lambda_{i-1,\,i}P_{i-1}(t) - (\lambda_{i,\,i-1} + \lambda_{i,\,i+1})P_i(t) + \lambda_{i+1,\,i}P_{i+1}(t) \quad (i=1,\ 2,\ \dots,\ n), \tag{15.1}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots$$

где $P_i(t)$ — вероятность состояния S_i, когда в системе находится i требований в момент времени t, $n+1$ — общее число возможных состояний S_0, S_1, ..., S_n.

При гипотезе о стационарном режиме работы системы (вероятности состояний не зависят от времени) уравнения Колмогорова (15.1) принимают вид

$$-\lambda_{01}P_0 + \lambda_{10}P_1 = 0,$$

$$\dots\dots\dots\dots$$

$$\lambda_{i-1,\,i}P_{i-1} - (\lambda_{i,\,i+1} + \lambda_{i,\,i+1})P_i + \lambda_{i+1,\,i}P_{i+1} = 0 \quad (i=1,\ 2,\ \dots,\ n). \tag{15.2}$$

$$\dots\dots\dots\dots$$

В большинстве практических задач оказывается допустимой гипотеза о стационарном режиме работы системы. Поэтому могут быть использованы уравнения Колмогорова вида (15.2).

Математические модели систем массового обслуживания, приводимые ниже, соответствуют уравнениям Колмогорова для стационарного режима работы системы (15.2) при условиях простейшего потока входящих требований и экспоненциального закона распределения времени обслуживания.

15.4. Системы массового обслуживания с отказами

Граф состояний многоканальной системы массового обслуживания с отказами при принятых допущениях (см.

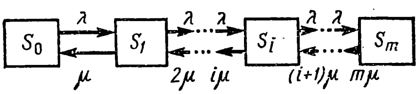


Рис. 15.1

п. 15.3) имеет вид, изображенный на рис. 15.1. Здесь λ — интенсивность входящего потока требований, μ — производительность одного канала (прибора) обслуживания, S_0, S_1, ..., S_m — состояния системы (индекс указывает число требований в системе), m — общее число каналов.

Вероятности состояний определяются по формулам

$$P_i = \frac{\rho^i}{i!} P_0 \quad (i = 1, 2, \ldots, m),$$

где $\rho = \frac{\lambda}{\mu}$, а вероятность P_0 находится из выражения

$$P_0 = \left[\sum_{i=0}^{m} \frac{\rho^i}{i!}\right]^{-1}.$$

○ **Пример.** В вычислительный центр коллективного пользования с тремя ЭВМ поступают заказы от предприятий на вычислительные работы. Если работают все три ЭВМ, то вновь поступающий заказ не принимается и предприятие вынуждено обратиться в другой вычислительный центр. Пусть среднее время работы с одним заказом составляет 3 ч. Интенсивность потока заявок 0,251/ч. Найти вероятность отказа и среднее число занятых ЭВМ.

Имеем: $m = 3$, $\lambda = 0{,}25\,(1/\text{ч})$, $T_{\text{обс}} = 3$ (ч). Находим:

$$\rho = \frac{\lambda}{m} = \lambda \cdot \overline{T}_{\text{обсл}} = 3 \cdot 0{,}25 = 0{,}75,$$

$$P_0 = \left[\sum_{i=0}^{3} \frac{\rho^i}{i!}\right]^{-1} = \left[1 + 0{,}75 + \frac{0{,}75^2}{2!} + \frac{0{,}75^3}{3!}\right]^{-1} = [2{,}1]^{-1},$$

$$P_{\text{отк}} = \frac{\rho^m}{m!} P_0 = \frac{0{,}75^3}{3!} \cdot \frac{1}{2{,}1} = 0{,}033,$$

$$m_3 = \sum_{n=1}^{m} n P_n = P_0 \sum_{n=1}^{m} \frac{\rho^n}{(n-1)!} =$$

$$= \frac{1}{2{,}1}\left[0{,}75 + 0{,}75^2 + \frac{0{,}75^3}{2}\right] \approx 0{,}72.$$

Таким образом, $P_{\text{отк}} = 0{,}033$; $m_з = 0{,}72$ (ЭВМ). ●

15.5. Системы массового обслуживания с ограниченной длиной очереди

Граф состояний многоканальной системы массового обслуживания, имеющей m каналов с ограниченной очередью, число мест в которой ограничено величиной l, при принятых допущениях (см. п. 15.3) имеет вид, изображенный на рис. 15.2.

Вероятности состояний $S_1, S_2, \ldots, S_m$ находят по формуле

$$P_i=\frac{\rho^i}{i!}P_0 \quad (i=1, 2 \ldots, m).$$

Вероятности состояний $S_{m+1}, S_{m+2}, \ldots, S_{m+l}$ определяют с помощью формул

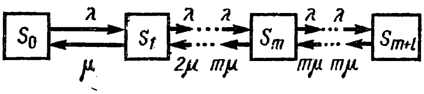


Рис. 15.2

$$P_i=\frac{\rho^i}{m^{i-m}m!}P_0 \quad (i=m+1, \ldots, m+l).$$

Вероятность P_0 подсчитывают по формуле

$$P_0=\left[\sum_{i=0}^{m}\frac{\rho^i}{i!}+\sum_{i=m+1}^{m+l}\frac{\rho^i}{m^{i-m}m!}\right]^{-1}.$$

В большинстве практических задач отношение $\frac{\rho}{m}<1$. Формула для P_0 используется в виде

$$P_0=\left[\sum_{i=0}^{m}\frac{\rho^i}{i!}+\frac{\rho^{m+1}}{m\cdot m!}\frac{1-\left(\frac{\rho}{m}\right)^l}{1-\frac{\rho}{m}}\right]^{-1}.$$

○ **Пример.** На автозаправочной станции установлены три колонки для выдачи бензина. Около станции находится площадка на три машины для их ожидания в очереди. На станцию прибывает в среднем две машины в минуту. Среднее время заправки одной машины 1 мин. Требуется определить вероятность отказа и среднюю длину очереди.

Имеем: $m=3$, $l=3$, $\lambda=2$ (1/мин), $\overline{T}_{\text{обс}}=1$ (мин), $\mu=1/\overline{T}_{\text{обс}}=1$ (1/мин). Далее находим: $\rho=\lambda/\mu=2/1=2$, $\rho/m=2/3$,

$$P_0=\left[\sum_{i=0}^{m}\frac{\rho^i}{i!}+\frac{\rho^{m+1}}{m\cdot m!}\cdot\frac{1-(\rho/m)^l}{1-\rho/m}\right]^{-1}=$$

$$=\left[1+2+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^4}{3\cdot 3!}\cdot\frac{1-(2/3)^3}{1-2/3}\right]^{-1}\approx 0{,}122,$$

$$P_{\text{отк}}=P_{m+l}=\frac{\rho^{m+l}}{m^l\cdot m!}P_0=\left(\frac{\rho}{m}\right)^l\frac{\rho^m}{m!}P_0=$$

$$=(2/3)^3\,\frac{2^3}{3!}\cdot 0{,}122=0{,}048,$$

$$M_{\text{ож}}=\frac{P_0\rho^m}{m!}\sum_{n=1}^{l}n\left(\frac{\rho}{m}\right)^n=\frac{0{,}122\cdot 2^3}{3!}\left[\frac{2}{3}+2\left(\frac{2}{3}\right)^2+\right.$$

$$\left.+3\cdot\left(\frac{2}{3}\right)^3\right]=0{,}35.$$

Таким образом, $P_{\text{отк}}=0{,}048$, $M_{\text{ож}}=0{,}35$ (машины). ●

15.6. Системы массового обслуживания с ожиданием

Граф состояний системы массового обслуживания с ожиданием аналогичен графу состояний системы с ограниченной длиной очереди при условии, что граница очереди

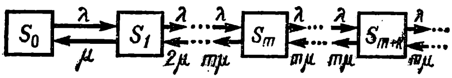


Рис. 15.3

отодвигается в бесконечность. Такой граф состояний изображен на рис. 15.3.

Вероятности состояний находят по формулам

$$P_i=\frac{\rho^i}{i!}P_0\quad(i=1,\ 2,\ \ldots,\ m),$$

$$P_i=\frac{\rho^i}{m!\,m^{i-m}}P_0\quad(i=m+1,\ \ldots,\ m+k,\ \ldots).$$

При $\rho/m<1$ для определения P_0 используют формулу

$$P_0=\left[\sum_{i=0}^{m}\frac{\rho^i}{i!}+\frac{\rho^{m+1}}{m!\,(m-\rho)}\right]^{-1}.$$

○ **Пример.** В порту имеется два причала для разгрузки грузовых судов. Интенсивность потока судов равна 0,8 (судов в сутки). Среднее время разгрузки одного судна составляет 2 сут. Предполагается, что очередь ожидающих разгрузки судов может быть неограниченной длины.

Найти среднее число занятых причалов и среднее время пребывания судна в порту.

Имеем: $m=2$, $\lambda=0{,}8\,(1/\text{сут})$, $\mu=1/\overline{T}_{\text{обс}}=0{,}5\,(1/\text{сут})$, $\rho=\lambda/\mu=0{,}8/0{,}5=1{,}6$. Находим:

$$P_0=\left[\sum_{i=0}^{m}\frac{\rho^i}{i!}+\frac{\rho^{m+1}}{m!\,(m-\rho)}\right]^{-1}=$$

$$=\left[1+\frac{1{,}6}{1!}+\frac{1{,}6^2}{2!}+\frac{1{,}6^3}{2!\,(2-1{,}6)}\right]^{-1}=0{,}11,$$

$$m_з=\rho q,\quad q=1\Rightarrow m_з=1{,}6,$$

$$M_{\text{ож}}=\frac{P_0\rho^{m+1}}{m\cdot m!}\cdot\frac{1}{(1-\rho/m)^2}=\frac{0{,}11\cdot 1{,}6^3}{2\cdot 2\cdot(1-0{,}8)^2}=2{,}8,\quad \overline{T}_{\text{ож}}=\frac{M_{\text{ож}}}{\lambda}=3{,}5.$$

Итак, $m_з=1{,}6$ (причалов), $\overline{T}_{\text{ож}}=3{,}5$ (сут). ●

15.7. Система массового обслуживания с ограниченным временем ожидания

В системах массового обслуживания с ограниченным временем ожидания время ожидания в очереди каждого требования ограничено случайной величиной $t_{\text{ож}}$, среднее значение которого $\overline{t}_{\text{ож}}$.

Величина, обратная среднему времени ожидания, означает среднее количество требований, покидающих очередь в единицу времени, вызванное появлением в очереди одного требования: $\nu=1/\overline{t}_{\text{ож}}$.

При наличии в очереди k требований интенсивность потока покидающих очередь требований составляет $k\nu$. Граф состояний такой системы изображен на рис. 15.4.

Формулы для определения вероятностей состояний такой системы имеют вид

$$P_i=\frac{\rho^i}{i!}P_0\quad (i=1,\ 2,\ \ldots,\ m),$$

$$P_i=\frac{\rho^m}{m!}\cdot\frac{\lambda^k}{\prod\limits_{j=1}^{k}(m\mu+j\nu)}\qquad (i=m+1,\ \ldots,\ m+k,\ \ldots),$$

где $\prod\limits_{j=1}^{k}(m\mu+j\nu)$ — произведение сомножителей $m\mu+j\nu$.

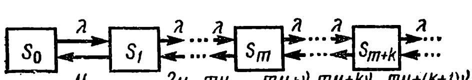


Рис. 15.4

Вероятность P_0 определяют по формуле

$$P_0=\left[\sum_{i=0}^{m}\frac{\rho^i}{i!}+\frac{\rho^m}{m!}\sum_{k=1}^{\infty}\frac{\lambda^k}{\prod\limits_{j=1}^{k}(m\mu+j\nu)}\right]^{-1}.$$

В практических задачах сумму бесконечного ряда вычислить достаточно просто, так как члены ряда быстро убывают с увеличением номера.

○ **Пример.** В пункте химчистки имеется три аппарата для чистки. Интенсивность потока посетителей $\lambda=6$ (посетителей в час). Интенсивность обслуживания посетителей одним аппаратом $\mu=3$ (посетителей в час). Среднее количество посетителей, покидающих очередь, не дождавшись обслуживания, $\nu=1$ (посетитель в час). Найти абсолютную пропускную способность пункта.

Имеем: $m=3$, $\lambda=6$, $\mu=3$, $\nu=1$. Находим: $\rho=\lambda/\mu=$ $=6/3=2$,

$$P_0=\left[1+\frac{2}{1!}+\frac{2^2}{2!}+\frac{2^3}{3!}+\frac{2^3}{3!}\times\right.$$
$$\left.\times\left(\frac{6}{3\cdot3+1}+\frac{6^2}{(3\cdot3+1)(3\cdot3+2\cdot1)}\right)\right]^{-1}=0{,}13.$$

Вероятность занятости всех приборов равна $P_{\text{зан}}=1-$ $-P_0=0{,}87$. Тогда абсолютная пропускная способность может быть получена как произведение: $A=mP_{\text{зан}}=$ $=3\cdot0{,}87=2{,}61$. Таким образом, $A=2{,}61$ (посетителя в час). ●

15.8. Замкнутые системы массового обслуживания

В замкнутых системах массового обслуживания источник требований находится внутри системы и интенсивность потока требований зависит от состояния самой системы. Чаще всего потоком требований в такой системе является поток неисправностей от некоторой группы работающих устройств. Пусть имеется m работающих устройств, которые могут выходить из строя за счет неисправностей. Имеется также n приборов (каналов) обслуживания этих требований. В качестве таких каналов могут выступать и люди. Обычно предполагают, что $n<m$.

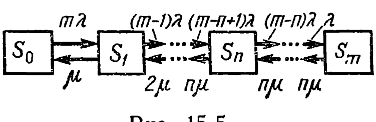


Рис. 15.5

Обозначим через S_0 состояние, при котором все устройства работают, а приборы обслуживания не заняты; S_1— состояние, при котором одно устройство вышло из строя и обслуживается одним прибором обслуживания; S_n—n устройств не работают и все приборы заняты обслуживанием; S_m—все устройства не работают, из них n обслуживаются и $m-n$ ждут обслуживания. Граф состояний такой системы изображен на рис. 15.5.

Вероятности состояний замкнутой системы определяются следующими зависимостями:

$$P_i = \frac{\prod\limits_{j=0}^{i-1}(m-j)}{i!}\rho^i P_0 \quad (i=1,\ 2,\ \ldots,\ n),$$

$$P_i = \frac{\prod\limits_{j=0}^{i-1}(m-j)}{n!n^{i-n}}\rho^i P_0 \quad (i=n+1,\ n+2,\ \ldots,\ m),$$

где П—знак произведения сомножителей $m-j$,

$$P_0 = \left[1 + \sum_{i=1}^{n}\frac{\prod\limits_{j=0}^{i-1}(m-j)}{i!}\rho^i + \sum_{i=n+1}^{m}\frac{\prod\limits_{j=0}^{i-1}(m-j)}{n!n^{i-n}}\rho^i\right]^{-1}.$$

○ **Пример.** Рабочий обслуживает группу из трех станков. Каждый станок останавливается в среднем два раза в час. Процесс наладки занимает в среднем 10 мин. Определить абсолютную пропускную способность наладки рабочим станков.

Имеем: $n=1$, $m=3$, $\lambda=2$, $T_{обс}=1/6$, $\mu=6$. Находим: $\rho=\lambda/\mu=1/3$,

$$P_0 = \left[1 + m\rho + \frac{m(m-1)}{1!1^1}\rho^2 + \frac{m(m-1)(m-2)}{1!1^2}\rho^3\right]^{-1} =$$

$$= \left[1 + 3\cdot\frac{1}{3} + 3\cdot2\cdot\left(\frac{1}{3}\right)^2 + 3\cdot2\cdot1\cdot\left(\frac{1}{3}\right)^3\right]^{-1} = 0{,}346.$$

Определяем вероятность того, что рабочий будет занят обслуживанием:

$$P_з = 1 - P_0 = 1 - 0{,}346 = 0{,}654.$$

Если рабочий занят обслуживанием, то он обслуживает 6 станков в час. Следовательно, абсолютная способность находится как произведение:

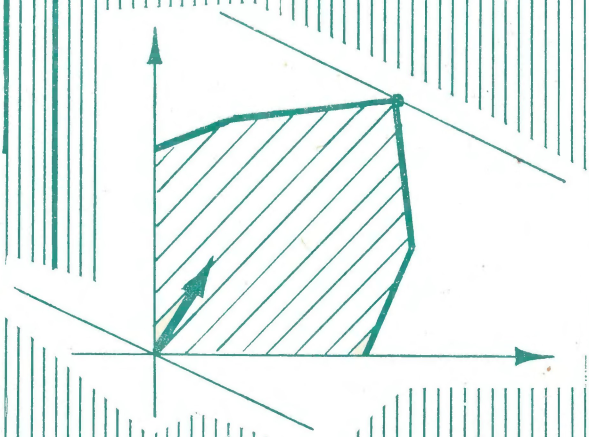
$$A = \mu P_з = 6\cdot 0{,}654 = 4.$$

Таким образом, $A=4$ (станка в час). ●

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

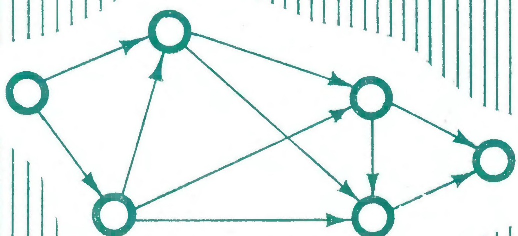
$$\mathrm{grad}\, F(\bar{x}) = \left\{ \frac{\partial F}{\partial x_1} \right.$$



$$x_k - \sum_{i=1}^{n} a_{ik} x_i = y_n \qquad k = 1, 2, \ldots, n$$

$$\frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n} \right\}$$

$$F = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$



$$1(X) = \int_{-\infty}^{\infty} x f(x) dx$$

1 руб.