Московский физико-технический институт
(Национальный исследовательский университет)

Умнов А.Е., Умнов Е.А.

ПРОПЕДЕВТИЧЕСКИЙ КУРС МАТЕМАТИКИ

МФТИ, 30апр2026г

Оглавление

Глава 1.

Введение

Предлагаемое пособие имеет своей целью помощь студентам высших учебных заведений в практическом освоении базовых понятий и методов высшей математики на начальном этапе обучения.

Этот материал может быть также полезен при изучении физики, химии, биологии, экономики, социологии и других дисциплин, использующих аппарат высшей математики в качестве инструмента исследования.

При этом предполагается, что читатель изначально владеет лишь базовыми знаниями из элементарной математики в объеме средней школы.

Важно также отметить, что данный курс является не альтернативой изучению предмета в стандартном объеме, а лишь средством предварительного ознакомления с ним. Поэтому он не претендует ни на полноту тематики материала, ни на, тем более, формальную строгость его изложения.

Данный курс, включающий как теоретические занятия, так и практические упражнения, содержит следующие разделы:

– матрицы и системы линейных уравнений,
– функции, последовательности и пределы,
– элементы аналитической геометрии,
– производные, интегралы и ряды,
– элементы теории вероятностей и математической статистики.

Подобная структура не является традиционной для учебных курсов высшей математики. Как содержание ее разделов, так и последовательность изложения материала в первую очередь обусловлены задачей формирования базы знаний, необходимой для успешного освоения основных понятий и методов высшей математики студентами самых различных специализаций.

1.1. Некоторые полезные сведения из курса элементарной математики

Обычно предполагается, что студенты изначально владеют основами элементарной математики в объеме учебных программ средней школы. Тем не менее, представляется целесообразным привести (в справочной форме) перечень некоторых базлвых сведений, используемых при изучении высшей математики.

1°. **Числа и их виды**

Как известно, основным объектом изучения в математике являются *числа,* для совокупности которых, называемой *числовым множеством,* возможно выполнение операций сравнения, сложения, умножения и т.д.

Числа разделяются на

– *натуральные,* возникающие при счете каких-либо объектов,

– *целые,* множество которых состоит из натуральных «со знаком» и числа «ноль»,

– *рациональные,* представимые в виде несокращаемого отношения двух целых чисел, и

– *иррациональные,* формой представления которых является бесконечная непериодическая десятичная дробь.

В качестве общего названия рациональных и иррациональных чисел используется термин *вещественные числа.*

2°. Формулы сокращенного умножения

Для любых чисел a и b справедливы равенства

$$\begin{aligned}
&(a+b)^2 = a^2 + 2ab + b^2,\\
&(a-b)^2 = a^2 - 2ab + b^2,\\
&a^2 - b^2 = (a+b)(a-b),\\
&(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3,\\
&(a-b)^3 = a^3 - 3a^2b + 3ab^2 - b^3,\\
&a^3 + b^3 = (a+b)(a^2 - ab + b^2),\\
&a^3 - b^3 = (a-b)(a^2 + ab + b^2).
\end{aligned}$$

Напомним также, что для любых *неотрицательных* чисел a и b имеет место соотношение $a + b \geq 2\sqrt{ab}$, причем по определению арифметического квадратного корня принимается, что

$$\sqrt{a^2} = |a| = \begin{cases} \;\;a, & \text{если } a \geq 0,\\ -a, & \text{если } a < 0. \end{cases}$$

Отметим, что $|a|$ – *абсолютную величину* числа a, принято также называть *модулем* a.

3°. Линейные уравнения

Уравнение вида

$$ax + b = 0,$$

где x – неизвестное, а $a \neq 0$ и b – фиксированные числа, называется *линейным.* Оно имеет единственное решение $x = -\dfrac{b}{a}$.

4°. Квадратные уравнения

Уравнение вида

$$ax^2 + bx + c = 0,$$

где x – неизвестное, а $a \neq 0,$ b и c – фиксированные числа, называется *квадратным.*

При $D > 0$ оно имеет *два* решения

$$x_1 = \frac{-\,b + \sqrt{D}}{2a} \quad \text{и} \quad x_2 = \frac{-\,b - \sqrt{D}}{2a},$$

где число $D = b^2 - 4ac$ называется *дискриминантом.*

При $D = 0$ квадратное уравнение имеет *одно* решение

$$x = -\frac{b}{2a}\,,$$

а при $D < 0$ оно *не имеет* решений.

Если квадратное уравнение имеет корни, то будут выполняться следующие равенства (*теорема Виета*)

$$x_1 + x_2 = -\frac{b}{a} \qquad \text{и} \qquad x_1 \cdot x_2 = \frac{c}{a}\,.$$

5°. **Степени и их свойства**

Степенью числа a *порядка* k (k – натуральное число, не меньшее, чем 2), обозначаемой a^k, называется произведение вида $\underbrace{a \cdot a \cdot a \cdot \ldots \cdot a \cdot a}_{k \text{ сомножителей}}$. В этом случае число a называют *основанием*, а число k – *показателем* степени.

Степени с натуральным показателем $k \geq 2$ обладают очевидными свойствами:

$$1)\ a^{n+m} = a^n \cdot a^m;$$
$$2)\ (a^n)^m = a^{nm}.$$

Понятие степени положительного и не равного единицы числа a можно обобщить на случай, когда ее показатель есть рациональное число вида $p = \dfrac{m}{n}$, то есть числа m и $n \neq 0$ любые целые. Для этого (*по определению*) принимают, что для любого $a > 0$ и $a \neq 1$ выполняются равенства

$$a^1 = a\ ; \qquad a^0 = 1\ ; \qquad a^{-m} = \frac{1}{a^m}\ ; \qquad a^{\frac{1}{m}} = \sqrt[m]{a}.$$

Тогда для степеней с рациональным показателем также будут справедливы свойства 1) и 2):

$$1)\ a^{p+q} = a^p \cdot a^q;$$
$$2)\ (a^p)^q = a^{pq}.$$

В курсе математического анализа показывается, что соотношения 1) и 2) справедливы и для любых вещественных чисел p и q, при любом положительном, не равном единице вещественном числе a.

Функция $y = a^x$, $a > 0$, $a \neq 1$ называется *показательной*. На рис.1.1 приведен ее графика при $a = 3$.

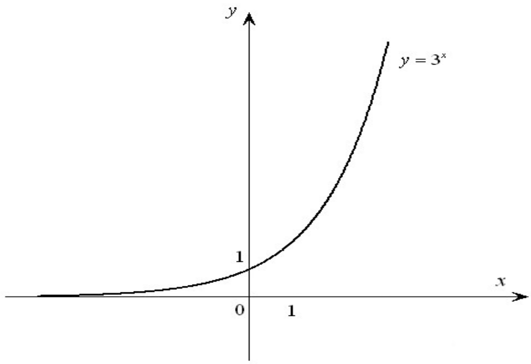


Рис. 1.1. График показательной функции

6°. **Логарифмы и их свойства**

Логарифмом положительного числа a *по основанию* b (b – положительное число и $b \neq 1$), обозначаемым $\log_b a$, называется *показатель степени,* в которую следует возвести число b, чтобы получить a. Число b принято называть *основанием* логарифма.

Отметим, что данное определение можно также представить в виде формулы

$$b^{\log_b a} = a, \quad a > 0, \quad b > 0, \quad b \neq 1,$$

которую называют *основным логарифмическим тождеством.*

Для часто используемых на практике *десятичных* логарифмов (то есть логарифмов по основанию 10) для упрощения записи применяется специальное обозначение $\log_{10} a \equiv \lg a$. По той же причине в высшей математике логарифмы по основанию e (иррациональное число $e \approx 2.72\ldots$), называемые *натуральными*, обозначают $\log_e a \equiv \ln a$.

Логарифмы для любых положительных чисел a, b, c и $c \neq 1$ обладают следующими, вытекающими из их определения, свойствами:

$$1)\ \log_c ab = \log_c a + \log_c b;$$

$$2)\ \log_c \frac{a}{b} = \log_c a - \log_c b;$$

$$3)\ \log_c a^b = b \log_c a;$$

$$4)\ \log_c a = \frac{\log_b a}{\log_b c}, \quad b \neq 1.$$

Формулу 4) можно использовать для перехода от одного основания логарифма к другому. Например, $\log_2 17 = \dfrac{\lg 17}{\lg 2}$.

Функция $y = \log_a x,\ a > 0,\ a \neq 1$ называется *логарифмической*. На рис.1.2 показан вид ее графика при $a = 3$.

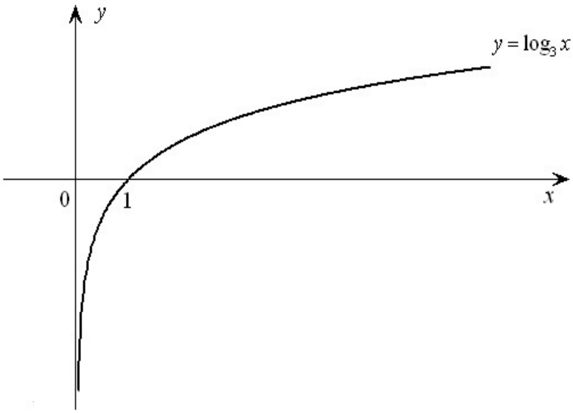


Рис. 1.2. График логарифмической функции

7°. Тригонометрические функции, тождества и уравнения

Напомним, что углы можно измерять как в *градусной*, так и в *радианной* мере.

За один *градус* принимают величину центрального угла, опирающегося на дугу в окружности, длина которой равна $\frac{1}{360}$ длины окружности. За один *радиан* принимают величину центрального угла в окружности, опирающегося на дугу, длина которой равна радиусу этой окружности. Поскольку длина окружности $2\pi r$, то угол в 360^o имеет величину 2π радиан.

Определение основных тригонометрических функций удобно давать при помощи так называемого *тригонометрического круга* показанного на рис.1.3.

> *Синусом угла* α называется отношение y-координаты точки A к длине вектора $\overrightarrow{r} = \overrightarrow{OA}$.
> *Косинусом угла* α называется отношение x-координаты точки A к длине вектора $\overrightarrow{r} = \overrightarrow{OA}$.
> *Тангенсом угла* α называется отношение y-координаты точки A к ее x-координате.

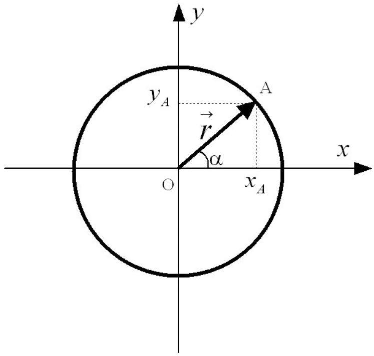


Рис. 1.3. Определение основных тригонометрических функций

Заметим, что в силу этих определений

1) синус и косинус имеют значения (не превосходящие по модулю 1) для *любых* углов α. В то время как тангенс может принимать любые значения от $-\infty$ до $+\infty$, но при этом он не существует для углов равных $\frac{\pi}{2} + \pi n$, где n любое целое число, то есть для углов $\pm 90^o$, $\pm 270^o$, $\pm 450^o, \ldots$.

2) тригонометрические функции обладают свойством *периодичности*, то есть их значения повторяются при изменении аргумента α на одно и то же минимально возможное, положительное число, называемое *периодом.*
Период синуса и косинуса равен 2π , а тангенса — π .

Тригонометрические функции вещественного аргумента x (обычно измеряемого в радианной мере) принято обозначать как $y = \sin x$, $y = \cos x$ и $y = \operatorname{tg} x$. Их графики приведены на рисунках 1.4 – 1.6.

Иногда используются также функции *котангенс, секенс* и *косекенс* $\operatorname{ctg} x = \dfrac{1}{\operatorname{tg} x}$, $\sec x = \dfrac{1}{\cos x}$, $\operatorname{cosec} x = \dfrac{1}{\sin x}$.

Для тригонометрических функций справедливы равенства, позволяющие выражать одни из них через другие. Приведем наиболее часто используемые на практике соотношения.

1) Основные тригонометрические тождества

$$\sin^2 x + \cos^2 x = 1 \ , \quad \operatorname{tg} x = \frac{\sin x}{\cos x} \ , \quad \frac{1}{\cos^2 x} = 1 + \operatorname{tg}^2 x \ .$$

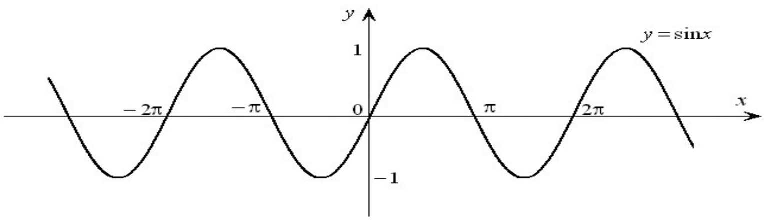


Рис. 1.4. График функции *синус*

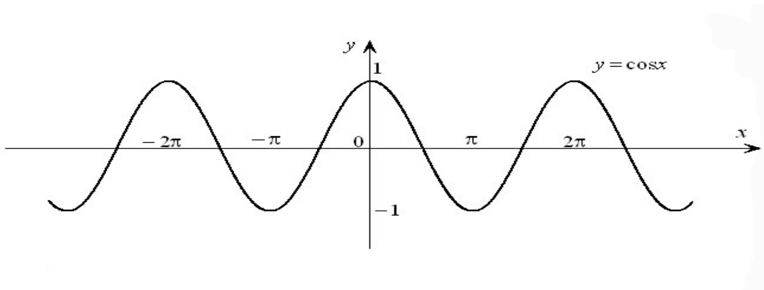


Рис. 1.5. График функции *косинус*

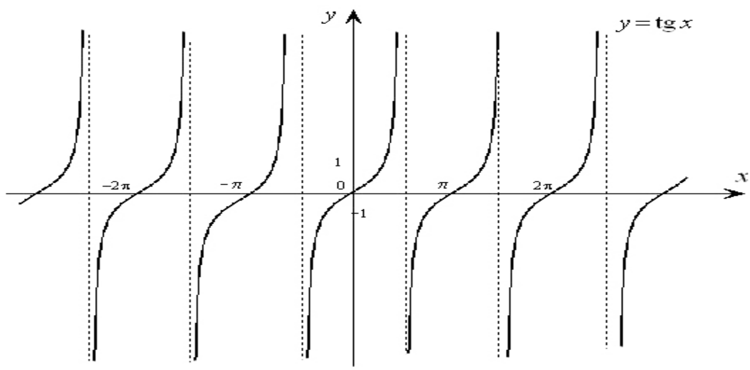


Рис. 1.6. График функции *тангенс*

2) Формулы суммы (разности) аргументов

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta\ ,$$

$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta\ ,$$

$$\operatorname{tg}(\alpha \pm \beta) = \frac{\operatorname{tg}\alpha \pm \operatorname{tg}\beta}{1 \mp \operatorname{tg}\alpha\operatorname{tg}\beta}\ ,$$

3) Формулы двойных аргументов

$$\sin 2\alpha = 2\sin\alpha\cos\alpha\ , \quad \cos 2\alpha = \cos^2\alpha - \sin^2\alpha\ ,$$

$$\operatorname{tg} 2\alpha = \frac{2\operatorname{tg}\alpha}{1 - \operatorname{tg}^2\alpha}\ .$$

4) Формулы преобразования суммы (разности) тригонометрических функций в произведение и обратно

$$\sin\alpha \pm \sin\beta = 2\sin\frac{\alpha \pm \beta}{2}\cos\frac{\alpha \mp \beta}{2}\ ,$$

$$\cos\alpha + \cos\beta = 2\cos\frac{\alpha + \beta}{2}\cos\frac{\alpha - \beta}{2}\ ,$$

$$\cos\alpha - \cos\beta = -2\sin\frac{\alpha + \beta}{2}\sin\frac{\alpha - \beta}{2}\ ,$$

$$\operatorname{tg}\alpha \pm \operatorname{tg}\beta = \frac{\sin(\alpha \pm \beta)}{\cos\alpha\cos\beta}\ ,$$

$$\sin\alpha\sin\beta = \frac{1}{2}\Big(\cos(\alpha - \beta) - \cos(\alpha + \beta)\Big)\ ,$$

$$\cos\alpha\cos\beta = \frac{1}{2}\Big(\cos(\alpha - \beta) + \cos(\alpha + \beta)\Big)\ ,$$

$$\sin\alpha\cos\beta = \frac{1}{2}\Big(\sin(\alpha + \beta) + \sin(\alpha - \beta)\Big)\ .$$

5) Наконец, могут оказаться полезными и соотношения вида

$$\sin\alpha = \frac{2\operatorname{tg}\dfrac{\alpha}{2}}{1 + \operatorname{tg}^2\dfrac{\alpha}{2}}\ , \qquad \cos\alpha = \frac{1 - \operatorname{tg}^2\dfrac{\alpha}{2}}{1 + \operatorname{tg}^2\dfrac{\alpha}{2}}\ ,$$

$$\operatorname{tg}\alpha = \frac{2\operatorname{tg}\dfrac{\alpha}{2}}{1-\operatorname{tg}^2\dfrac{\alpha}{2}}\,,$$

$$\sin 3\alpha = 3\cos^2\alpha\sin\alpha - \sin^3\alpha\ ,$$

$$\cos 3\alpha = \cos^3\alpha - 3\sin^2\alpha\cos\alpha\ .$$

В случае, когда $\alpha+\beta+\gamma=\pi$ (например, если α,β,γ суть углы некоторого треугольника) будут верны равенства

$$\sin\alpha+\sin\beta+\sin\gamma = 4\cos\frac{\alpha}{2}\cos\frac{\beta}{2}\cos\frac{\gamma}{2}\,,$$

$$\operatorname{tg}\alpha+\operatorname{tg}\beta+\operatorname{tg}\gamma = \operatorname{tg}\alpha\cdot\operatorname{tg}\beta\cdot\operatorname{tg}\gamma\ ,$$

В вычислительной практике достаточно часто возникает задача определения величины угла по значению какой-либо из его тригонометрических функций.

Для решения этой задачи используются *обратные тригонометрические функции*: *арксинус* $y=\arcsin x$, *арккосинус* $y=\arccos x$ и *арктангенс* $y=\operatorname{arctg} x$.

Напомним определения этих функций.

Арксинусом x при условии, что $|x|\le 1$, называется число y такое, что $|y|\le\dfrac{\pi}{2}$ и $\sin y = x$.

Арккосинусом x при условии, что $|x|\le 1$, называется число y такое, что $0\le y\le\pi$ и $\cos y = x$.

Арктангенсом x (для любого x) называется число y такое, что $|y|\le\dfrac{\pi}{2}$ и $\operatorname{tg} y = x$.

При решении прикладных задач часто оказывается полезной (легко проверяемой по приведенным выше определениям) формула

$$\arcsin x + \arccos x = \frac{\pi}{2}\,.$$

Отметим, что аргументы обратных тригонометрических функций не имеют размерности, в то время как их значения являются углами и изменяются как правило в радианной мере. Графики обратных тригонометрических функций приведены на рисунках 1.7 – 1.9.

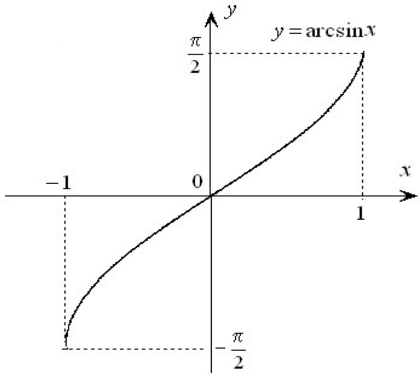


Рис. 1.7. График функции *арксинус*

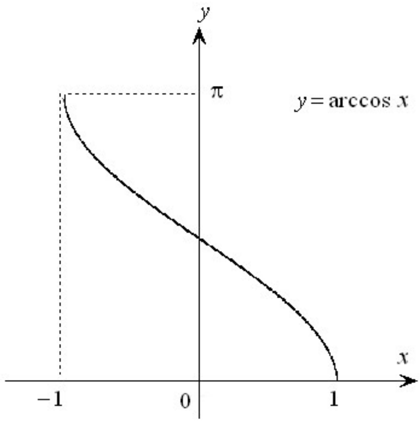


Рис. 1.8. График функции *арккосинус*

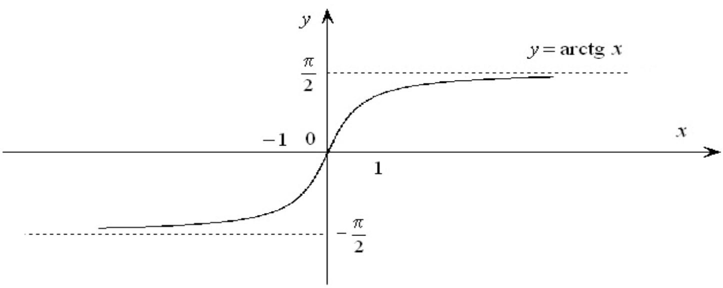


Рис. 1.9. График функции *арктангенс*

Обратные тригонометрические функции можно использовать при решении тригонометрических уравнений. Так, например, уравнение

$$\sin x = a \quad \text{при } |a| \le 1 \text{ имеет корни вида } x = (-1)^n \arcsin a + \pi n \ ,$$

$$\text{уравнение} \quad \cos x = a \quad \text{при } |a| \le 1 \text{ имеет корни } x = \pm \arccos a + 2\pi n \ ,$$

$$\text{уравнение} \quad \operatorname{tg} x = a \quad \text{при любом } a \text{ имеет корни } x = \operatorname{arctg} a + \pi n \ ,$$

причем во всех этих формулах n – любое целое число, то есть $n = 0, \pm 1, \pm 2, \pm 3, \ldots$

8°. **Множества. Элементы комбинаторики**

Под *множеством* в математике понимают совокупность объектов (или элементов), которые можно отличать как друг от друга, так и от объектов, не входящих в данную совокупность.

Тот факт, что объект x принадлежит множеству X принято обозначать как $x \in X$. Если объект x не принадлежит множеству X, то используется обозначение $x \notin X$. Для обозначения *пустого* множества, то есть не имеющего в своем составе ни одного объекта, используется символ $\emptyset$. Наконец, два множества X и Y, состоящие из одних и тех же объектов, называются *равными,* с обозначением факта равенства как $X = Y$.

Для множеств существуют операции *объединения*, обозначаемая символом $\bigcup$, и *пересечения* – $\bigcap$. Запись $x \in X \bigcup Y$ означает, что объект x принадлежит либо множеству X, либо множеству Y, либо им обоим одновременно. В свою очередь, запись $x \in X \bigcap Y$ означает, что объект x принадлежит одновременно как множеству X, так и множеству Y.

Если объектами, образующими множество X являются числа, то такое множество принято называть *числовым.* Множество состоящее из чисел x, удовлетворяющих неравенствам $a \le x \le b$, называется *отрезком* и обозначается $[a, b]$. Если же числовое множество состоит из чисел, для которых $a < x < b$, то оно называется *интервалом* и обозначается (a, b). Наконец, термин *промежуток* обозначает либо либо отрезок, либо интервал, либо *полуинтервал* вида $(a, b]$ или $[a, b)$. Промежуток, содержащий точку x, принято называть *окрестностью* этой точки.

Пусть имеется множество, состоящее из n элементов. Каждую *упорядоченную* выборку из этого множества, содержащую k элементов, называют *размещением* из n элементов по k элементов (иногда говорят: «размещение из n по k»). В рассматриваемом случае очевидно, что $0 \le k \le n$.

Число всех размещений из n по k обозначается A_n^k. Если учесть, что при $k = 0$ существует только одно размещение – пустое множество $\emptyset$, то справедливо равенство

$$A_n^k = \begin{cases} 1, & \text{если } k = 0, \\ n(n-1)(n-2)\ldots(n-k+1), & \text{если } k > 0. \end{cases}$$

Произведение всех натуральных чисел от 1 до n называется n-*факториалом*, его принято обозначать

$$1 \cdot 2 \cdot 3 \cdot \ldots \cdot n = n!$$

(читается «*эн-факториал*»). С помощью этого обозначения формула для полного числа размещений упрощается и принимает вид

$$A_n^k = \frac{n!}{(n-k)!}.$$

Заметим, что в силу $A_0^0 = 1$, имеет смысл считать (*по определению*) $0! = 1$. Тогда данная формула будет верной для любых $0 \le k \le n$.

Размещение из n элементов по n называется *перестановкой* из n элементов. Число всех перестановок из n элементов $P_n = n!$.

Согласно своему определению, размещения могут отличаться друг от друга как составом своих элементов, так и порядком их следования.

Если же порядок следования элементов в выборке не существенен, то такую выборку k элементов из n принято называть *сочетанием* из n элементов по k элементов (иногда говорят: «сочетание из n по k»). Формула для C_n^k – числа всех сочетаний из n по k имеет вид

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

Используя эту формулу, можно получить следующие полезные соотношения: $C_n^k = C_n^{n-k}$ и $(a+b)^n =$

$$= C_n^0 a^n + C_n^1 a^{n-1} b + C_n^2 a^{n-2} b^2 + \cdots + C_n^k a^{n-k} b^k + \cdots + C_n^{n-1} a b^{n-1} + C_n^n b^n.$$

Последнее равенство носит название *формулы бинома Ньютона* и является обобщением некоторых формул сокращенного умножения, приведенных в 2°.

9°. **Последовательности и прогрессии**

Будем говорить, что задана *числовая последовательность*, если указано правило, по которому каждому натуральному числу (номеру)

n поставлено сопоставлено единственное число x_n, называемое значением n-го члена последовательности. Числовую последовательность принято обозначать как $\{x_n\}$.

Числовая последовательность называется *арифметической прогрессией* $\{a_n\}$, если каждый ее член, начиная со второго, равен предыдущему, сложенному с постоянным числом d, называемым *разностью прогрессии*. Таким образом, для задания арифметической прогрессии следует указать ее первый член и разность, тогда для любого номера n будет справедливо равенство $a_{n+1} = a_n + d$.

В арифметической прогрессии значение ее n-го члена может быть найдено по формуле $a_n = a_1 + (n-1)d$. Кроме того, справедливо равенство

$$a_1 + a_2 + \dots + a_{n-1} + a_n = \frac{a_1 + a_n}{2} \cdot n\ ,$$

для суммы первых n членов арифметической прогрессии.

Числовая последовательность называется *геометрической прогрессией* $\{b_n\}$, если каждый ее член, начиная со второго, равен предыдущему, умноженному на постоянное число q, называемое *знаменателем прогрессии*. Для задания геометрической прогрессии следует указать ее первый член и знаменатель, тогда для любого номера n будет справедливо равенство $b_{n+1} = b_n q$.

В геометрической прогрессии значение ее n-го члена может быть найдено по формуле $b_n = b_1 q^{n-1}$. Кроме того, справедливо соотношение

$$b_1 + b_2 + \dots + b_{n-1} + b_n = \begin{cases} b_1 n, & \text{если } q = 1, \\ b_1 \dfrac{1-q^n}{1-q}, & \text{если } q \neq 1, \end{cases}$$

позволяющее находить сумму первых n членов геометрической прогрессии.

Наконец, если $|q| < 1$, то такая геометрическая прогрессия называется *бесконечно убывающей*, и в последнем равенстве можно перейти к пределу при неограниченном возрастании n. Величина этого предела, называемая суммой бесконечно убывающей прогрессии, будет равна $\dfrac{b_1}{1-q}$. Эта формула используется, например, для перевода записи дробных чисел из обыкновенной формы в десятичную и наоборот.

10°. **Специальные обозначения и символы**

В современных математических текстах допускаются некоторые стандартные обозначения, практически не применяемые в пособиях, используемых при изучении математики в средней школе. Поскольку это обстоятельство может до некоторой степени усложнить освоение студентами курса математики, представляется целесообразным привести краткое описание этих стандартов и правил их использования.

1) *Символы общности и существования*

Символ *общности* $\forall$ используется для замены слов «всякий», «любой», «для любого», в то время как символ *существования* $\exists$ заменяет слова «существует», «найдется»‘. Например, определение ограниченной числовой последовательности: последовательность $\{x_n\}$ называется ограниченной, если

найдется неотрицательное число C такое, что для любого номера n будет выполнено неравенство $|x_n| \leq C$,

может быть записано так: последовательность $\{x_n\}$ называется ограниченной, если

$$\exists C \geq 0 : \forall n : |x_n| \leq C \, .$$

2) *Символы суммирования и произведения*

Если необходимо записать выражение для суммы, в которой число слагаемых не имеет конкретного значения, но известно как зависит величина каждого слагаемого от его номера, то можно использовать специальный символ суммирования $\sum\limits_{k=1}^{n} a_k$, указав при этом общий вид слагаемого и диапазон изменения индекса суммирования. Можно считать, что этот символ заменяет слова «сумма слагаемых вида a_k по k в пределах от 1 до n».

Например, сумма

$$\sin 1 + \sin 2 + \sin 3 + \cdots + \sin(n-1) + \sin n$$

записывается как

$$\sum_{k=1}^{n} \sin k,$$

а сумма

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots + \frac{1}{(n-2)(n-1)} + \frac{1}{(n-1)n}$$

представляется в виде

$$\sum_{k=2}^{n} \frac{1}{(k-1)k}$$

.

Аналогично, формула бинома Ньютона с помощью символа суммирования может быть записана как

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^{n-k} b^k.$$

Например, при $n = 4$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4.$$

Подобный вид записи существует и для операции произведения. Например, равенство $1 \cdot 2 \cdot 3 \cdot \ldots \cdot n = n!$ можно записать в виде $\prod\limits_{k=1}^{n} k = n!$.

В заключение приведем следующие, используемые при решении многих задач высшей математики, формулы

$$1 + 2 + 3 + \cdots + (n-1) + n = \sum_{k=1}^{n} k = \frac{n(n+1)}{2}\ ;$$

$$1^2 + 2^2 + 3^2 + \cdots + (n-1)^2 + n^2 = \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}\ ;$$

$$1^3 + 2^3 + 3^3 + \cdots + (n-1)^3 + n^3 = \sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}\ ;$$

$$\sin\alpha + \sin 2\alpha + \sin 3\alpha + \cdots + \sin n\alpha = \sum_{k=1}^{n} \sin\alpha k = \frac{\sin\dfrac{(n+2)\alpha}{2} \cdot \sin\dfrac{n\alpha}{2}}{\sin\dfrac{\alpha}{2}}\ .$$

Отметим, что из этих соотношений следует любопытное равенство

$$\sum_{k=1}^{n} k^3 = \left(\sum_{k=1}^{n} k\right)^2 .$$

11°. **Полезные неравенства**

Для любых двух неотрицательных чисел a и b верно *неравенство Коши*

$$\frac{a+b}{2} \geq \sqrt{ab}\ ,$$

которое иногда используется в следующей форме: для любых двух вещественных чисел x и y справедливо соотношение

$$x^2 + y^2 \geq 2|xy| \ .$$

Неравенство Коши верно и для большего числа неотрицательных чисел:

$$\frac{a_1 + a_2 + a_3 + \ldots + a_{n-1} + a_n}{n} \geq \sqrt[n]{a_1 a_2 a_3 \cdots a_{n-1} a_n} \ .$$

Отметим, что с помощью символов суммирования и умножения последнее соотношение можно записать как

$$\frac{1}{n}\sum_{k=1}^{n} a_k \geq \sqrt[n]{\prod_{k=1}^{n} a_k} \ .$$

В большом числе прикладных задач необходимые оценки могут быть получены при помощи, вытекающего из формулы бинома Ньютона и верного для любого x и любого $a > -1$, *неравенства Бернулли*

$$(1 + a)^x \geq 1 + xa \ .$$

1.2. Замечания о роли точности определений и формулировок

В процессе изучения математики следует обращать особое внимание на полноту и точность *определений, формулировок теорем и описания свойств.* Недопустима как избыточность (излишняя многословность) подобных лексем, так и потеря каких-либо их деталей.

Проиллюстрируем это следующими примерами.

1°. *Арифметический квадратный корень.* Как это уже было отмечено, по определению арифметического квадратного корня считается, что $\sqrt{a^2} = |a|$. Может возникнуть вопрос: «Не проще ли положить, что $\sqrt{a^2} = a$?»

Что бы показать некорректность такого определения, рассмотрим следующую цепочку преобразований: для *любой* пары чисел x и y будут верными равенства

$$x^2 - 2xy + y^2 = y^2 - 2yx + x^2,$$

$$(x-y)^2 = (y-x)^2,$$

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

Если теперь применить определение вида $\sqrt{a^2} = a$, то мы получим

$$x - y = y - x \quad \Rightarrow \quad x = y,$$

что очевидно $\forall x, y$ неверно. В то время как использование определения $\sqrt{a^2} = |a|$ дает

$$|x-y| = |y-x| \quad \Rightarrow \quad 0 = 0,$$

что верно для любой пары чисел x и y.

2°. *Сколько корней может иметь квадратное уравнение?*

Рассмотрим три следующих утверждения A), B) и C):

A) Уравнение $ax^2 + bx + c = 0$ – квадратное.

B) Квадратное равнение не может иметь более двух корней.

C) Для любых, попарно неравных чисел α, β и γ уравнение

$$\frac{(x-\alpha)(x-\beta)}{(\gamma-\alpha)(\gamma-\beta)} + \frac{(x-\beta)(x-\gamma)}{(\alpha-\beta)(\alpha-\gamma)} + \frac{(x-\gamma)(x-\alpha)}{(\beta-\gamma)(\beta-\alpha)} = 1$$

может быть приведено к виду A) и при этом оно очевидно имеет три различных корня $x_1 = \alpha,\ x_2 = \beta,\ x_3 = \gamma$.

Понятно, что утверждения A), B) и C) противоречивы в свой совокупности. Иначе говоря, одно из них ошибочное и, на первый взгляд, наибольшие сомнения вызывает утверждение C). Однако, оно на самом деле верное, а ошибка содержится в утверждении A). Дело в том, что квадратным называется уравнение $ax^2 + bx + c = 0 \quad$ с $a \neq 0$. И именно для него верно утверждение B). В нашем же случае, если привести уравнение C) к виду, указанному в утверждении A), коэффициент при x^2 окажется равным нулю. Более того, это уравнение примет вид $1 = 1$, то есть является *тождеством* – верным равенством при любом значении x (в том числе и при $x_1 = \alpha,\ x_2 = \beta,\ x_3 = \gamma$).

3°. *Можно ли произвольно группировать слагаемые в сумме?* Казалось бы ассоциативность операции сложения для чисел позволяет дать положительный ответ на данный вопрос. Однако это верно лишь для сумм с *конечным* числом слагаемых. Если число слагаемых в сумме не ограничено, то возможно возникновение ситуации подобной следующей. Согласимся «на веру» с утверждением, что сумма неограниченного числа нулей равна нулю, и рассмотрим сумму вида

$$A = 1 + 2 - 3 + 1 + 2 - 3 + 1 + 2 - 3 + 1 + 2 - 3 + 1 + 2 - 3 + \ldots$$

сгруппировав слагаемые сначала как

$$A = (1 + 2 - 3) + (1 + 2 - 3) + (1 + 2 - 3) + (1 + 2 - 3) + (1 + 2 - 3) + \ldots$$

придем к заключению, что $A = 0$, поскольку каждая сумма в скобках дает ноль. Однако, при другом способе группировки

$$A = 1 + (2 - 3 + 1) + (2 - 3 + 1) + (2 - 3 + 1) + (2 - 3 + 1) + (2 - 3 + \ldots$$

получаем $A = 1$. Что означает неприменимость сочетательного правила для сумм с неограниченным числом слагаемых.

Последний пример наглядно демонстрирует, что с «бесконечностью» нельзя оперировать как с обычным числом.

Стоит также отметить, что методологически аналогичные проблемы могут возникнуть и в случае подмены понятий «отсутствие определенности» и «существование вероятности», которая нередко допускается при рассуждениях на интуитивном уровне.

Глава 2.

Элементы линейной алгебры

2.1. Матрицы

2.1.1. Определение и классификация матриц

Определение 2.1.1.1. *Матрицей* размера $m \times n$ называется упорядоченная прямоугольная таблица чисел, имеющая m строк и n столбцов.

Матрицу принято символически обозначать при помощи двойных вертикальных ограничителей: $\|A\|$, либо записывать в развернутом в виде

$$\left\| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \ddots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right\|,$$

хотя нередко для матриц можно встретить и обозначение вида

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \ddots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Каждое из чисел a_{ij} называется *элементом* матрицы $\|A\|$, а числа i и j – соответственно *строковыми* и *столбцовыми индексами* элемента a_{ij}. Число i показывает в какой строке матрицы расположен элемент a_{ij}, число j – в каком ее столбце этот элемент находится.

Матрицы классифицируются как по *числу строк и столбцов,* например,

квадратная, порядка n, имеющая n строк и n столбцов;
m-компонентная *строка*, имеющая одну строку ($m = 1$) и n столбцов; n-компонентный *столбец*, имеющий m строк и один столбец ($n = 1$)

$$\left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \ddots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{matrix} \right\| \quad \left\| \begin{matrix} a_{11} & a_{12} & \dots & a_{1n} \end{matrix} \right\| \quad \left\| \begin{matrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{matrix} \right\|,$$

так и по *значению их элементов,* например,

нулевая матрица (обозначаемая часто как $\|O\|$), *все* элементы которой имеют нулевое значение;
единичная матрица: квадратная, порядка n, вида

$$\|E\| = \left\| \begin{matrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & 1 \end{matrix} \right\|.$$

2.1.2. Операции с матрицами

С матрицами возможно выполнять операции: сравнения (установления факта равенства), сложения, умножения числа на матрицу, транспонирования, произведения и обращения.

Правила выполнения этих операций дают следующие определения.

Определение 2.1.2.1. **Сравнение.** Две матрицы $\|A\|$ и $\|B\|$ называются *равными* (что обозначается как $\|A\| = \|B\|$), если они

равных размеров и
имеют элементы с равными значениями для одинаковых пар индексов i и j.

Пример 2.1.2.1. Сравнение матриц:

а) $\left\| \begin{array}{cc} \sin\dfrac{\pi}{2} & \cos\dfrac{\pi}{2} \\ \cos\dfrac{\pi}{2} & \sin\dfrac{\pi}{2} \end{array} \right\| = \left\| \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right\|;$

б) $\left\| \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \right\| \neq \left\| \begin{array}{cc} 1 & 1 \\ 2 & 2 \\ 3 & 3 \\ 4 & 4 \end{array} \right\|$, поскольку сравниваемые матрицы имеют разное число столбцов.

Определение 2.1.2.2. **Сложение.** Матрица $\|C\|$ называется *суммой* матриц $\|A\|$ и $\|B\|$ (обозначается $\|A\| + \|B\| = \|C\|$), если все три матрицы равных размеров $m \times n$ и каждый элемент матрицы $\|C\|$ равен сумме соответствующих элементов матриц $\|A\|$ и $\|B\|$, то есть $c_{ij} = a_{ij} + b_{ij} \quad \forall i = [1, m], j = [1, n]$.

Пример 2.1.2.2. Сложение матриц:

$$a) \quad \left\| \begin{array}{rr} 7 & -3 \\ 4 & 0 \\ 5 & -1 \end{array} \right\| + \left\| \begin{array}{rr} 2 & -5 \\ -4 & 7 \\ 1 & 3 \end{array} \right\| = \left\| \begin{array}{lr} (7+2) & (-3-5) \\ (4-4) & (0+7) \\ (5+1) & (-1+3) \end{array} \right\| = \left\| \begin{array}{rr} 9 & -8 \\ 0 & 7 \\ 6 & 2 \end{array} \right\|.$$

$$b) \quad \left\| \begin{array}{rrrr} \lg 7 & \lg 6 & \lg 8 & \lg 18 \\ \lg 9 & \lg 5 & \lg 15 & 3\lg 2 \\ -\lg 4 & \lg 3 & 2\lg 5 & \lg 11 \end{array} \right\| + \left\| \begin{array}{rrrr} \lg 2 & \lg 6 & \lg 3 & -\lg 3 \\ \lg 2 & \lg 5 & -\lg 3 & -\lg 8 \\ \lg 20 & \lg 4 & \lg 3 & \lg 3 \end{array} \right\| =$$

$$= \left\| \begin{array}{rrrr} \lg 14 & 2\lg 6 & \lg 24 & \lg 6 \\ \lg 18 & \lg 25 & \lg 5 & 0 \\ \lg 5 & \lg 12 & \lg 75 & \lg 33 \end{array} \right\|.$$

(см. §1.1, п.6°.)

Определение 2.1.2.3. **Умножение числа на матрицу.** Матрица $\|B\|$ называется *произведением числа k на матрицу* $\|A\|$ (что обозначается как $k \cdot \|A\|$), если каждый элемент матрицы $\|B\|$ равен произведению числа k на соответствующий элемент матрицы $\|A\|$, то есть $b_{ij} = k \cdot a_{ij} \quad \forall i = [1, m], j = [1, n]$.

Пример 2.1.2.3. Умножение числа на матрицу:

$$a)\quad 3\cdot\left\|\begin{array}{rrrr} 2 & 0 & 5 & -1 \\ -3 & 1 & 0 & 6 \end{array}\right\| = \left\|\begin{array}{rrrr} (3\cdot 2) & (3\cdot 0) & (3\cdot 5) & (3\cdot(-1)) \\ (3\cdot(-3)) & (3\cdot 1) & (3\cdot 0) & (3\cdot 6) \end{array}\right\| =$$

$$= \left\|\begin{array}{rrrr} 6 & 0 & 15 & -3 \\ -9 & 3 & 0 & 18 \end{array}\right\|.$$

$$b)\quad \frac{1}{9}\cdot\left\|\begin{array}{lll} 3^5 & 3^{-1} & 3^2 \\ 3^{-2} & 3^4 & 3^3 \\ 3^4 & 3^{-3} & 3^6 \\ 3 & 3^2 & 3^3 \end{array}\right\| = \left\|\begin{array}{llr} 27 & 3^{-3} & 1 \\ 3^{-4} & 9 & 3 \\ 3^2 & 3^{-5} & 81 \\ \frac{1}{3} & 1 & 3 \end{array}\right\|.$$

(см. 1.1, п.5°.)

Определение 2.1.2.4. **Транспонирование матрицы.** Результатом *транспонирования* матрицы $\|A\|$ является новая матрица, обозначаемая $\|A\|^{\mathrm{T}}$, строками которой служат столбцы исходной матрицы $\|A\|$, записанные с сохранением порядка их следования. В этом случае $a_{ij}^{\mathrm{T}} = a_{ji} \quad \forall i = [1, m], j = [1, n]$.

Пример 2.1.2.4. Транспонирование матрицы:

$$\left\|\begin{array}{ccccc} 3 & 1 & 4 & 7 & 7 \\ 2 & 8 & 5 & 6 & 6 \end{array}\right\|^{\mathrm{T}} = \left\|\begin{array}{cc} 3 & 2 \\ 1 & 8 \\ 4 & 5 \\ 7 & 6 \\ 7 & 6 \end{array}\right\|$$

Определение 2.1.2.5. **Произведение матриц.** Матрица $\|C\|$ размера $m \times n$ называется *произведением* матриц $\|A\|$ размера $m \times p$ и $\|B\|$ размера $p \times n$ (оно обозначается как $\|A\| \cdot \|B\| = \|C\|$), где каждый элемент матрицы $\|C\|$ равен

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{i(l-1)}b_{(l-1)j} + a_{ip}b_{pj}$$

$$\text{или} \quad c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} \quad \forall i = [1, m], j = [1, n].$$

(см. §1.1, п.10°.)

Пример 2.1.2.5. Вычислить произведение матриц:

$$\|A\| = \begin{Vmatrix} 3 & -1 & 0 & 2 & -2 \\ 2 & 0 & -3 & 1 & 1 \end{Vmatrix} \qquad \text{и} \qquad \|B\| = \begin{Vmatrix} 3 & 2 & -2 \\ 1 & -1 & 3 \\ 3 & 0 & 2 \\ 5 & -4 & 0 \\ 0 & 3 & 1 \end{Vmatrix}.$$

В данном случае матрица $\|A\|$ размера 2×5, а матрица $\|B\|$ размера 5×3. Согласно определению 2.1.2.5 матрица $\|C\|$ будет иметь две строки и три столбца. Подсчитаем значения всех ее шести элементов, приняв во внимание, что в данном примере $m = 2$, $n = 3$ и $l = 5$.

$$\begin{aligned}
c_{11} &= 3\cdot 3 + (-1)\cdot 1 + 0\cdot 3 + 2\cdot 5 + (-2)\cdot 0 = \\ &= 18, \\
c_{12} &= 3\cdot 2 + (-1)\cdot(-1) + 0\cdot 0 + 2\cdot(-4) + (-2)\cdot 3 = \\ &= -7, \\
c_{13} &= 3\cdot(-2) + (-1)\cdot 3 + 0\cdot 2 + 2\cdot 0 + (-2)\cdot 1 = \\ &= -11, \\
c_{21} &= 2\cdot 3 + 0\cdot 1 + (-3)\cdot 3 + 1\cdot 5 + 1\cdot 0 = \\ &= 2, \\
c_{22} &= 2\cdot 2 + 0\cdot(-1) + (-3)\cdot 0 + 1\cdot(-4) + 1\cdot 3 = \\ &= 3, \\
c_{23} &= 2\cdot(-2) + 0\cdot 3 + (-3)\cdot 2 + 1\cdot 0 + 1\cdot 1 = \\ &= -9.
\end{aligned}$$

В итоге имеем $\|C\| = \|A\| \cdot \|B\| =$

$$= \begin{Vmatrix} 3 & -1 & 0 & 2 & -2 \\ 2 & 0 & -3 & 1 & 1 \end{Vmatrix} \cdot \begin{Vmatrix} 3 & 2 & -2 \\ 1 & -1 & 3 \\ 3 & 0 & 2 \\ 5 & -4 & 0 \\ 0 & 3 & 1 \end{Vmatrix} = \begin{Vmatrix} 18 & -7 & -11 \\ 2 & 3 & -9 \end{Vmatrix}.$$

Еще раз укажем, что

- перемножать можно матрицы, у которых число столбцов первого сомножителя равно числу строк второго сомножителя (в данном примере их по 5);
- число строк произведения равно числу строк первого сомножителя (то есть 2), а число столбцов произведения равно числу столбцов второго сомножителя (то есть 3).

Отметим следующие полезные свойства операции произведения матриц.

1. Результатом умножения любой матрицы на *нулевую* матрицу подходящего размера всегда будет *нулевая* матрица.
2. Умножение слева или справа на *единичную* матрицу (подходящего размера!) любую данную матрицу *не меняет.*
3. Если требуется перемножить три матрицы $\|A\| \cdot \|B\| \cdot \|C\|$, то начинать умножение можно *с любой пары стоящих рядом матриц.* То есть верно равенство

$$\big(\|A\| \cdot \|B\|\big) \cdot \|C\| = \|A\| \cdot \big(\|B\| \cdot \|C\|\big).$$

Данное свойство произведения матриц называется *ассоциативностью.*

При использовании операции произведения матриц следует помнить, что (в отличие от произведения чисел) **матричное произведение не обладает свойством коммутативности**, то есть при изменении порядка сомножителей результат произведения матриц может оказаться иным, или даже не существовать вовсе. Следующие примеры иллюстрируют это.

Попробуем изменить порядок сомножителей в примере 2.1.2.5. Нетрудно заметить, что произведение матриц

$$\left\| \begin{matrix} 3 & 2 & -2 \\ 1 & -1 & 3 \\ 3 & 0 & 2 \\ 5 & -4 & 0 \\ 0 & 3 & 1 \end{matrix} \right\| \cdot \left\| \begin{matrix} 3 & -1 & 0 & 2 & -2 \\ 2 & 0 & -3 & 1 & 1 \end{matrix} \right\|$$

существовать не будет, поскольку число столбцов первого сомножителя (*три*) в этом случае не равно числу строк второго (*два*). Однако в обратном порядке эти матрицы перемножать можно.

Рассмотрим другой пример. Допустим, что трехкомпонентную строку $\|A\| = \|\ a_1\ a_2\ a_3\ \|$ надо умножить на трехкомпонентный столбец $\|B\| = \left\| \begin{matrix} b_1 \\ b_2 \\ b_3 \end{matrix} \right\|$. Если первый сомножитель строка, второй – столбец, то произведение есть матрица размера 1×1, то есть просто, число.

$$\|A\| \cdot \|B\| = \|\ a_1\ a_2\ a_3\ \| \cdot \left\| \begin{matrix} b_1 \\ b_2 \\ b_3 \end{matrix} \right\| = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

С другой стороны, произведение столбца на строку также существует, однако это будет не число, а квадратная матрица третьего порядка. Действительно,

$$\|B\| \cdot \|A\| = \left\| \begin{matrix} b_1 \\ b_2 \\ b_3 \end{matrix} \right\| \cdot \| \, a_1 \; a_2 \; a_3 \, \| = \left\| \begin{matrix} b_1a_1 & b_1a_2 & b_1a_3 \\ b_2a_1 & b_2a_2 & b_2a_3 \\ b_3a_1 & b_3a_2 & b_3a_3 \end{matrix} \right\|.$$

Таким образом, выполняя произведение матриц, необходимо:

1. Убедиться в возможности этой операции (число столбцов первого сомножителя должно равняться числу строк второго).
2. Определить размеры матрицы, получаемой в результате произведения (число ее строк будет равно числу строк первого сомножителя, а число ее столбцов – числу столбцов второго).
3. Найти значения элементов матрицы-произведения по формуле указанной в определении 2.1.2.5.

Использование операций с матрицами позволяет в большом числе случаев упрощать как постановку, так и решение различных математических задач. Например, система трех линейных уравнений с четырьмя неизвестными

$$\left\{ \begin{array}{rcrcrcrcr} 3x_1 & - & x_2 & + & 5x_3 & - & 2x_4 & = & 5 \\ -x_1 & + & 2x_2 & - & x_3 & - & x_4 & = & -1 \\ 2x_1 & + & 3x_2 & + & x_3 & - & x_4 & = & 5 \end{array} \right.$$

может быть записана в виде $\|A\| \cdot \|X\| = \|B\|$, где

$$\|A\| = \left\| \begin{matrix} 3 & -1 & 5 & -2 \\ -1 & 2 & -1 & -1 \\ 2 & 3 & 1 & -1 \end{matrix} \right\|; \quad \|X\| = \left\| \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} \right\|; \quad \|B\| = \left\| \begin{matrix} 5 \\ -1 \\ 5 \end{matrix} \right\|.$$

Проверьте это самостоятельно.

Определение 2.1.2.6. **Обращение матрицы.** Результатом *обращения* квадратной матрицы $\|A\|$ является новая матрица, называемая *обратной* к $\|A\|$ и обозначаемая $\|A\|^{-1}$, для которой верны равенства

$$\|A\|^{-1} \cdot \|A\| = \|A\| \cdot \|A\|^{-1} = \|E\|,$$

где $\|E\|$ - единичная матрица (см. §2.1.1) того же размера, что и матрица $\|A\|$.

Пример 2.1.2.6. Обращение матрицы:

а) Очевидный пример

$$\left\| \begin{array}{ccc} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{array} \right\|^{-1} = \left\| \begin{array}{ccc} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{4} \end{array} \right\|.$$

б) Не столь очевидный, но, тем не менее, верный пример (см.1.1, п.7°).

$$\left\| \begin{array}{cc} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{array} \right\|^{-1} = \left\| \begin{array}{cc} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{array} \right\|.$$

Действительно, справедливо равенство

$$\left\| \begin{array}{cc} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{array} \right\| \cdot \left\| \begin{array}{cc} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{array} \right\| =$$

$$= \left\| \begin{array}{cc} \cos^2\alpha + \sin^2\alpha & \cos\alpha\sin\alpha - \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha - \cos\alpha\sin\alpha & \cos^2\alpha + \sin^2\alpha \end{array} \right\| = \left\| \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right\|.$$

Второе равенство определения 2.1.2.6 проверяется аналогичным образом.

Важно отметить, что не любая квадратная матрица имеет обратную. Например, нулевая матрица необратима. Более детально условия, при которых у квадратной матрицы имеется обратная, равно как и правила ее вычисления, будут рассмотрены в следующем параграфе.

2.1.3. Определители и ранги матриц

Помимо уже рассмотренных, матрицы также характеризуются такими количественными характеристиками, как детерминанты (определители) и ранги.

Детерминант (определитель) матрицы

Рассмотрим важные для приложений понятия определителей квадратных матриц второго и третьего порядков. [1]

[1] Если размер матрицы 1×1, то она имеет единственный элемент. Детерминант такой матрицы будем считать равным значению этого элемента.

Определение 2.1.3.1. *Детерминантом* или *определителем* квадратной матрицы второго порядка $\|A\| = \left\| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right\|$ называется число, находимое по формуле

$$a_{11}a_{22} - a_{12}a_{21}$$

и обозначаемое $\det \|A\| = \det \left\| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right\|$.

Определение 2.1.3.2. *Детерминантом* или *определителем* квадратной матрицы третьего порядка $\|A\| = \left\| \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} \right\|$ называется число, находимое по формуле

$$a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - \\ - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}$$

и обозначаемое $\det \|A\| = \det \left\| \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} \right\|$.

Добавим, что детерминант квадратной матрицы *первого порядка* по определению считается равным единственному элементу этой матрицы. Для квадратных же матриц более высокого порядка, чем третий, также можно ввести определение детерминанта, обобщающее определения 2.1.3.1 и 2.1.3.2. Однако этот вопрос выходит за рамки нашего курса.

Пример 2.1.3.1. Вычисление детерминантов квадратных матриц второго и третьего порядков:

а) $n = 2$

$$\det \left\| \begin{array}{cc} 3 & 7 \\ 5 & -1 \end{array} \right\| = 3 \cdot (-1) - 7 \cdot 5 = -38.$$

б) $n = 3$

$$\det \left\| \begin{array}{ccc} 5 & -2 & 3 \\ 0 & 6 & 11 \\ 0 & 0 & 7 \end{array} \right\| = 5{\cdot}6{\cdot}7 + 3{\cdot}0{\cdot}0 + (-2){\cdot}11{\cdot}0 - (-2){\cdot}0{\cdot}7 - 3{\cdot}6{\cdot}0 - 5{\cdot}11{\cdot}0 = 210.$$

На практике использование определения детерминанта квадратной матрицы второго порядка не вызывает затруднений: его значение равно *произведению двух элементов матрицы, расположенных на главной диагонали, минус произведение двух элементов на вспомогательной диагонали.*

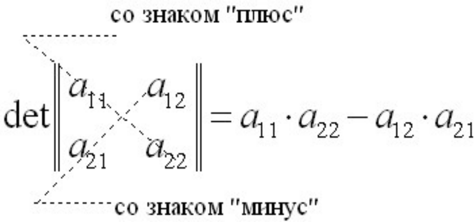


Рис. 2.1. Детерминант квадратной матрицы второго порядка

А для квадратных матриц третьего порядка вычисление детерминанта удобнее выполнять по так называемому правилу «треугольников», заметив, что в формуле 2.1.3.1 со знаком *плюс* входят произведения следующих выделенных троек элементов

$$\begin{Vmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & \boxed{a_{22}} & a_{23} \\ a_{31} & a_{32} & \boxed{a_{33}} \end{Vmatrix}, \quad \begin{Vmatrix} a_{11} & a_{12} & \boxed{a_{13}} \\ \boxed{a_{21}} & a_{22} & a_{23} \\ a_{31} & \boxed{a_{32}} & a_{33} \end{Vmatrix},$$

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

а со знаком *минус* – произведения троек

$$\begin{Vmatrix} a_{11} & a_{12} & \boxed{a_{13}} \\ a_{21} & \boxed{a_{22}} & a_{23} \\ \boxed{a_{31}} & a_{32} & a_{33} \end{Vmatrix}, \quad \begin{Vmatrix} a_{11} & \boxed{a_{12}} & a_{13} \\ \boxed{a_{21}} & a_{22} & a_{23} \\ a_{31} & a_{32} & \boxed{a_{33}} \end{Vmatrix},$$

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

Выбор троек элементов квадратной матрицы третьего порядка при вычислении ее определителя иллюстрирует рис. 2.1.

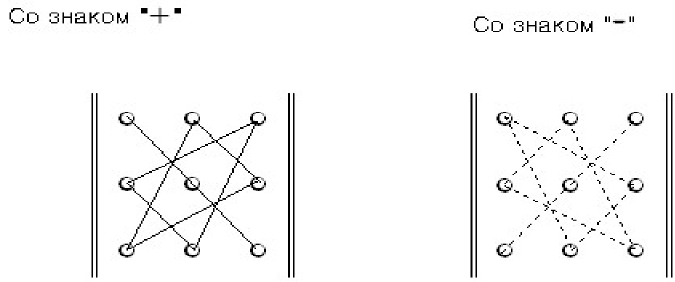


Рис. 2.2. Правило «треугольников»

Пример 2.1.3.2. Вычисление детерминанта по правилу «треугольников»:

$$\det\begin{Vmatrix} -3 & 0 & 2 \\ -1 & 8 & 1 \\ 5 & -4 & 0 \end{Vmatrix} =$$

$$= (-3)\cdot 8\cdot 0 + 2\cdot(-1)\cdot(-4) + 0\cdot 1\cdot 5 -$$
$$- 2\cdot 8\cdot 5 - 0\cdot(-1)\cdot 0 - (-3)\cdot 1\cdot(-4) = -84.$$

Альтернативой использования метода «треугольников» служат формулы, позволяющие выразить определитель третьего порядка через три определителя второго порядка. Например, формула *разложения по первой строке*

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

$$= a_{11}\cdot\det\begin{Vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{Vmatrix} - a_{12}\cdot\det\begin{Vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{Vmatrix} + a_{13}\cdot\det\begin{Vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{Vmatrix},$$

проверяемую непосредственно по определению 2.1.3.2.

Заметим, что в этой формуле

1) коэффициенты при определителях второго порядка есть элементы первой строки, взятые с чередующимися знаками «плюс» и «минус». Точнее говоря, знак коэффициентов при определителях второго порядка определяется по формуле $(-1)^{i+j}$, где i — номер строки, а j — номер столбца, в которых стоит коэффициент;
2) определители второго порядка есть детерминанты квадратных матриц, получаемых из исходной удалением строки и столбца, содержащих элемент, являющийся коэффициентом при определителе второго порядка.

Аналогичные формулы верны *для любых строк и столбцов* квадратной матрицы. Более того, любой определитель порядка n подобным способом можно выразить через определители порядка $n-1$.

Отметим, что использование формул разложения по строке (или столбцу) особенно эффективно, когда эта строка (или столбец) содержит нулевые элементы. Так, в примере 2.1.3.1.б целесообразно вычислять значение детерминанта, разложив его по *первому столбцу.*

$$\det\begin{Vmatrix} 5 & -2 & 3 \\ 0 & 6 & 11 \\ 0 & 0 & 7 \end{Vmatrix} =$$

$$= 5\cdot\det\begin{Vmatrix} \cancel{5} & -\cancel{2} & \cancel{3} \\ \cancel{0} & 6 & 11 \\ \cancel{0} & 0 & 7 \end{Vmatrix} - 0\cdot\det\begin{Vmatrix} \cancel{5} & -2 & 3 \\ \cancel{0} & \cancel{6} & \cancel{11} \\ \cancel{0} & 0 & 7 \end{Vmatrix} + 0\cdot\det\begin{Vmatrix} \cancel{5} & -2 & 3 \\ \cancel{0} & 6 & 11 \\ \cancel{0} & \cancel{0} & \cancel{7} \end{Vmatrix} =$$

$$= 5\cdot\det\begin{Vmatrix} 6 & 11 \\ 0 & 7 \end{Vmatrix} - 0\cdot\det\begin{Vmatrix} -2 & 3 \\ 0 & 7 \end{Vmatrix} + 0\cdot\det\begin{Vmatrix} -2 & 3 \\ 6 & 11 \end{Vmatrix} =$$

$$= 5\cdot(6\cdot 7 - 11\cdot 0) - 0\cdot((-2)\cdot 7 - 3\cdot 0) + 0\cdot((-2)\cdot 11 - 3\cdot 6) = 210;$$

Проверьте самостоятельно, что разложение этого определителя по *третьей строке* имеет следующий вид.

$$\det\begin{Vmatrix} 5 & -2 & 3 \\ 0 & 6 & 11 \\ 0 & 0 & 7 \end{Vmatrix} =$$

$$= 0\cdot\det\begin{Vmatrix} -2 & 3 \\ 6 & 11 \end{Vmatrix} - 0\cdot\det\begin{Vmatrix} 5 & 3 \\ 0 & 11 \end{Vmatrix} + 7\cdot\det\begin{Vmatrix} 5 & -2 \\ 0 & 6 \end{Vmatrix} = 210.$$

В приведенном примере для вычисления значения определителя третьего порядка оказалось достаточно найти значение одного определителя второго порядка, поскольку как первый столбец, так и третья строка имеют по два нулевых элемента.

Отметим теперь основные *свойства определителей.*

- при транспонировании квадратной матрицы ее определитель не меняется;
- если в матрице поменять местами две строки (или столбца), то ее определитель изменит знак на противоположный;

- из строки (столбца) определителя можно выносить общий множитель;
- определитель произведения квадратных матриц равен произведению определителей сомножителей:
  если $\|C\| = \|A\| \cdot \|B\|$, то $\det\|C\| = \det\|A\| \cdot \det\|B\|$;
- определитель матрицы равен нулю, если
  - в матрице есть нулевой столбец (нулевая строка);
  - матрица имеет две пропорциональные (в частности, равные) строки (столбцы).

Квадратную матрицу, определитель которой равен нулю, принято называть *вырожденной,* а матрицу с ненулевым детерминантом – *невырожденной.*

Используя понятие детерминанта можно сформулировать условие *обратимости* квадратной матрицы.

Для того,чтобы квадратная матрица имела обратную, необходимо и достаточно, чтобы чтобы *ее определитель был отличен от нуля.*

В этом случае матрица, обратная к матрице второго порядка

$$\|A\| = \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix},$$

будет определяться формулой

$$\|A\|^{-1} = \frac{1}{\det\|A\|} \cdot \begin{Vmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{Vmatrix}; \tag{2.1.3.1}$$

Действительно

$$\|A\|^{-1} \cdot \|A\| = \frac{1}{\det\|A\|} \cdot \begin{Vmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{Vmatrix} \cdot \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix} =$$

$$= \frac{1}{\det\|A\|} \cdot \begin{Vmatrix} a_{11}a_{22} - a_{12}a_{21} & -a_{11}a_{12} + a_{12}a_{11} \\ a_{21}a_{22} - a_{22}a_{21} & a_{11}a_{22} - a_{12}a_{21} \end{Vmatrix} =$$

$$= \frac{1}{\det\|A\|} \cdot \begin{Vmatrix} \det\|A\| & 0 \\ 0 & \det\|A\| \end{Vmatrix} = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} = \|E\| \;.$$

Равенство $\|A\| \cdot \|A\|^{-1} = \|E\|$ проверяется аналогичным способом.

Пример 2.1.3.3. Найти матрицу $\begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix}^{-1}$ и проверить соответствие результата определению 2.1.2.6.

Решение. Вначале убедимся, что матрица $\begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix}$ имеет обратную. Для этого найдем

$$\det \begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix} = 7 \cdot 5 - 6 \cdot 6 = -1 \neq 0.$$

Этот детерминант оказался не равным нулю. Значит, данная матрица обратима.

Теперь применим формулу (2.1.3.1). Получаем, с учетом определения 2.1.2.3, что

$$\begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix}^{-1} = \frac{1}{(-1)} \cdot \begin{Vmatrix} 5 & -6 \\ -6 & 7 \end{Vmatrix} = \begin{Vmatrix} -5 & 6 \\ 6 & -7 \end{Vmatrix}.$$

Теперь выполним проверку равенств $\|A\|^{-1} \cdot \|A\| = \|A\| \cdot \|A\|^{-1} = \|E\|$. Имеем

$$\begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix}^{-1} \cdot \begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix} = \begin{Vmatrix} -5 & 6 \\ 6 & -7 \end{Vmatrix} \cdot \begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix} =$$

$$= \begin{Vmatrix} (-5) \cdot 7 + 6 \cdot 6 & (-5) \cdot 6 + 6 \cdot 5 \\ 6 \cdot 7 + (-7) \cdot 6 & 6 \cdot 6 + (-7) \cdot 5 \end{Vmatrix} = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix} = \|E\| \ .$$

Убедитесь самостоятельно, что второе равенство

$$\begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix} \cdot \begin{Vmatrix} 7 & 6 \\ 6 & 5 \end{Vmatrix}^{-1} = \|E\|$$

также справедливо.

Пример 2.1.3.4. Найти матрицу $\|X\|$, если

a) $\|A\| \cdot \|X\| = \|B\|$,
где $\|A\| = \begin{Vmatrix} 2 & -3 \\ -1 & 2 \end{Vmatrix}, \|B\| = \begin{Vmatrix} -4 & -2 & 9 \\ 3 & 1 & -5 \end{Vmatrix}$;

b) $\|X\| \cdot \|A\| = \|B\|$,
где $\|A\| = \begin{Vmatrix} 7 & 5 \\ 4 & 3 \end{Vmatrix}, \|B\| = \begin{Vmatrix} 5 & 4 \\ 30 & 12 \\ 7 & 5 \end{Vmatrix}$.

Решение. a) Чтобы решить данное матричное уравнение, вначале определим какие размеры должна иметь искомая матрица.

Для того, чтобы произведение матриц $\|A\|\cdot\|X\|$ существовало необходимо равенство числа столбцов матрицы $\|A\|$ числу строк матрицы $\|X\|$, поэтому матрица $\|X\|$ должна иметь две строки.

С другой стороны, число столбцов произведения равно числу столбцов второго сомножителя, поэтому матрица $\|X\|$ будет иметь, как и матрица $\|B\|$, три столбца.

Итак, искомая матрица должна быть размера 2×3 .

Заметим, что матрица $\|A\|$ невырожденная, ее определитель отличен от нуля и, значит, она имеет обратную матрицу, равную (по формуле (2.1.3.1))

$$\|A\|^{-1} = \left\| \begin{array}{cc} 2 & 3 \\ 1 & 2 \end{array} \right\| .$$

Умножив *слева* на $\|A\|^{-1}$ обе части уравнения $\|A\| \cdot \|X\| = \|B\|$, получим

$$\|A\|^{-1} \cdot \|A\| \cdot \|X\| = \|A\|^{-1} \cdot \|B\| \ .$$

Затем, воспользовавшись определением 2.1.2.6 и свойствами матричного произведения, в силу $\|A\|^{-1} \cdot \|A\| = \|E\|$ и $\|E\| \cdot \|X\| = \|X\|$, упростим левую часть последнего равенства

$$\left(\|A\|^{-1} \cdot \|A\|\right) \cdot \|X\| = \|E\| \cdot \|X\| = \|X\| \ .$$

Таким образом, мы приходим к формуле, позволяющей найти искомую матрицу $\|X\|$.

$$\|X\| = \|A\|^{-1} \cdot \|B\| = \left\| \begin{array}{rr} 2 & -3 \\ -1 & 2 \end{array} \right\|^{-1} \cdot \left\| \begin{array}{rrr} -4 & -2 & 9 \\ 3 & 1 & -5 \end{array} \right\| =$$

$$= \left\| \begin{array}{cc} 2 & 3 \\ 1 & 2 \end{array} \right\| \cdot \left\| \begin{array}{rrr} -4 & -2 & 9 \\ 3 & 1 & -5 \end{array} \right\| = \left\| \begin{array}{rrr} 1 & -1 & 3 \\ 2 & 0 & -1 \end{array} \right\| .$$

b) Решение этой задачи аналогично пункту a) , за исключением того, что обе части исходного уравнения надо умножить на $\|A\|^{-1}$ не слева, а *справа.* Проверьте самостоятельно, что итоговая формула в данном случае имеет вид $\|X\| = \|B\| \cdot \|A\|^{-1}$ и дает решение

$$\|X\| = \left\| \begin{array}{rr} -1 & 3 \\ 2 & -4 \\ 1 & 0 \end{array} \right\| .$$

Ранг матрицы

Как уже отмечалось, детерминант является числовой характеристикой *квадратной* матрицы. Для неквадратной матрицы используется иная количественная характеристика, называемая ее *рангом.*

Пусть дана матрица $\|A\|$ размера $m \times n$. Выберем в ней произвольным (но не нарушающим порядок нумерации) способом k столбцов и k строк, на пересечении которых стоят элементы, образующие квадратную *подматрицу* k-го порядка. Определитель этой подматрицы называется *минором* k-го порядка матрицы $\|A\|$.

Пример 2.1.3.4. Для матрицы

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

найти сумму всех ее миноров 2-го порядка.

Решение. Вначале определим сколько миноров второго порядка имеет данная матрица. Поскольку минор 2-го порядка есть определитель квадратной подматрицы этого же порядка, то следует найти число квадратных подматриц 2-го порядка для данной матрицы. Иными словами, число миноров равно числу способов, которыми можно выбрать из трех строк данной матрицы две, а из трех ее столбцов – два.

Нетрудно заметить, что каждый их этих выборов можно осуществить лишь тремя способами, и потому полное число возможных вариантов оказывается равным девяти. Еще раз обратим внимание, что, во-первых, выборки строк и столбцов выполняются *независимо друг от друга,* и, во-вторых, *без нарушения порядка нумерации.*

Условимся, что при выборе пары строк с номерами i и j, а пары столбцов с номерами k и l, соответствующий минор 2-го порядка будет обозначаться как M_{ij}^{kl}. Элементы выбранной подматрицы будем выделять жирным шрифтом.

Подсчитаем значения миноров для всех девяти возможных случаев.

$$1) \qquad \left\| \begin{array}{ccc} \mathbf{2} & \mathbf{1} & 1 \\ \mathbf{0} & \mathbf{2} & 1 \\ 2 & 3 & 2 \end{array} \right\|, \qquad M_{12}^{12} = \det \left\| \begin{array}{cc} 2 & 1 \\ 0 & 2 \end{array} \right\| = 4\ ,$$

$$2) \qquad \left\| \begin{array}{ccc} \mathbf{2} & 1 & \mathbf{1} \\ \mathbf{0} & 2 & \mathbf{1} \\ 2 & 3 & 2 \end{array} \right\|, \qquad M_{12}^{13} = \det \left\| \begin{array}{cc} 2 & 1 \\ 0 & 1 \end{array} \right\| = 2\ ,$$

$$3) \qquad \left\| \begin{array}{ccc} 2 & \mathbf{1} & \mathbf{1} \\ 0 & \mathbf{2} & \mathbf{1} \\ 2 & 3 & 2 \end{array} \right\|, \qquad M_{12}^{23} = \det \left\| \begin{array}{cc} 1 & 1 \\ 2 & 1 \end{array} \right\| = -1\ ,$$

$$4)\qquad \left\| \begin{array}{ccc} \mathbf{2} & \mathbf{1} & 1 \\ 0 & 2 & 1 \\ \mathbf{2} & \mathbf{3} & 2 \end{array} \right\| , \qquad M_{13}^{12} = \det \left\| \begin{array}{cc} 2 & 1 \\ 2 & 3 \end{array} \right\| = 4 ,$$

$$5)\qquad \left\| \begin{array}{ccc} \mathbf{2} & 1 & \mathbf{1} \\ 0 & 2 & 1 \\ \mathbf{2} & 3 & \mathbf{2} \end{array} \right\| , \qquad M_{13}^{13} = \det \left\| \begin{array}{cc} 2 & 1 \\ 2 & 2 \end{array} \right\| = 2 ,$$

$$6)\qquad \left\| \begin{array}{ccc} 2 & \mathbf{1} & \mathbf{1} \\ 0 & 2 & 1 \\ 2 & \mathbf{3} & \mathbf{2} \end{array} \right\| , \qquad M_{13}^{23} = \det \left\| \begin{array}{cc} 1 & 1 \\ 3 & 2 \end{array} \right\| = -1 ,$$

$$7)\qquad \left\| \begin{array}{ccc} 2 & 1 & 1 \\ \mathbf{0} & \mathbf{2} & 1 \\ \mathbf{2} & \mathbf{3} & 2 \end{array} \right\| , \qquad M_{23}^{12} = \det \left\| \begin{array}{cc} 0 & 2 \\ 2 & 3 \end{array} \right\| = -4 ,$$

$$8)\qquad \left\| \begin{array}{ccc} 2 & 1 & 1 \\ \mathbf{0} & 2 & \mathbf{1} \\ \mathbf{2} & 3 & \mathbf{2} \end{array} \right\| , \qquad M_{23}^{13} = \det \left\| \begin{array}{cc} 0 & 1 \\ 2 & 2 \end{array} \right\| = -2 ,$$

$$9)\qquad \left\| \begin{array}{ccc} 2 & 1 & 1 \\ 0 & \mathbf{2} & \mathbf{1} \\ 2 & \mathbf{3} & \mathbf{2} \end{array} \right\| , \qquad M_{23}^{23} = \det \left\| \begin{array}{cc} 2 & 1 \\ 3 & 2 \end{array} \right\| = 1 ,$$

Таким образом мы получаем, что искомая сумма всех миноров 2-го порядка для данной матрицы составляет

$$4 + 2 + (-1) + 4 + 2 + (-1) + (-4) + (-2) + 1 = 5 \ .$$

Определение 2.1.3.3. **Рангом матрицы** $\|A\|$ называется *максимальный порядок ее миноров отличных от нуля.* Ранг обозначается $\operatorname{rg}\|A\|$.

Поскольку миноров порядка большего, чем число строк или столбцов матрицы, существовать не может, то из этого определения вытекает, что $\operatorname{rg}\|A\| \le \min\{m, n\}$.

Пример 2.1.3.5. Найдем ранги следующих матриц:

а)

$$\operatorname{rg} \left\| \begin{array}{rr} 3 & -5 \\ -2 & 4 \end{array} \right\| = 2.$$

б)

$$\operatorname{rg} \begin{Vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{Vmatrix} = 1.$$

в)

$$\operatorname{rg} \begin{Vmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 5 \end{Vmatrix} = 2.$$

Поясним указанные значения рангов. Для в матрице задачи а) можно выделить четыре ненулевых минора первого порядка и один минор второго порядка, совпадающий с определителем этой матрицы, значение которого $3 \cdot 4 - (-5) \cdot (-2) = 2 \neq 0$. Следовательно ранг матрицы, приведенной в примере а) равен 2.

В примере б) данная матрица имеет три столбца и четыре строки. Значит, ее ранг не может быть больше, чем 3. При этом, матрица имеет двенадцать ненулевых миноров первого порядка, а все миноры второго и третьего порядков равны нулю, как определители квадратных матриц с одинаковыми строками. Поэтому ранг данной матрицы равен 1.

Наконец, в примере в) у матрицы, данной в условии, есть четыре ненулевых минора первого порядка и один ненулевой минор второго порядка. А все миноры порядков три и четыре оказываются равными нулю, поскольку их матрицы содержат по крайней мере один нулевой столбец. Следовательно ранг данной матрицы равен 2.

Заметим, что из свойств детерминанта следует, что если у матрицы *все* миноры порядка k равны нулю, то и все миноры порядка $k+1$ (и выше) также будут нулевыми. Это свойство может быть полезным при нахождении ранга матрицы.

2.2. Системы линейных уравнений

2.2.1. Метод исключения неизвестных

В курсе элементарной математики рассматривается в качестве основного алгоритма решения систем линейных уравнений метод последовательного исключения неизвестных. Рассмотрим его применение для решения системы n линейных уравнений с n неизвестными

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nn}x_n & = & b_n \end{array}\right. \qquad (2.2.1.1)$$

в предположении, что эта система имеет единственное решение.

При использовании метода исключения, на первом шаге из последнего уравнения выражается x_n через $x_1, x_2, \dots x_{n-1}$. Затем полученное выражение подставляется в остальные $n-1$ уравнения. В результате мы получаем систему состоящую из $n-1$ уравнения с $n-1$ неизвестным. Последовательное повторение этой процедуры в итоге приведет к одному линейному уравнению с одним неизвестным x_1, решив которое, можно последовательно (в обратном порядке) также получить и значения остальных неизвестных.

Обобщением метода исключения является алгоритм решения системы (2.2.1.1), основанный на использовании матричных операций. Заметим, что эта система может быть записана в виде

$$\|A\| \cdot \|X\| = \|B\|, \qquad (2.2.1.2)$$

где

$$\|A\| = \left\| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{array} \right\|, \quad \|X\| = \left\| \begin{array}{c} x_1 \\ x_2 \\ \dots \\ x_n \end{array} \right\|, \quad \|B\| = \left\| \begin{array}{c} b_1 \\ b_2 \\ \dots \\ b_n \end{array} \right\|.$$

Договоримся матрицу $\|A\|$, элементами которой служат коэффициенты при неизвестных в системе (2.2.1.1), называть *основной* матрицей этой системы. В дальнейшем нам также понадобится матрица $\|A|B\|$, которая получается из основной приписыванием к ней справа столбца свободных членов $\|B\|$. Матрицу $\|A|B\|$ принято называть *расширенной* матрицей системы линейных уравнений.

В случае, когда матрица $\|A\|$ обратима, действуя так же как при решении задач примера 2.1.3.4, то есть, умножая обе части матричного уравнения (2.2.1.2) слева на $\|A\|^{-1}$, получим

$$\|A\|^{-1} \cdot \|A\| \cdot \|X\| = \|A\|^{-1} \cdot \|B\|$$

или

$$\|X\| = \|A\|^{-1} \cdot \|B\|, \tag{2.2.1.3}$$

поскольку

$$\|A\|^{-1} \cdot \|A\| = \|E\| \quad \text{и} \quad \|E\| \cdot \|A\| = \|A\|.$$

Пример 2.2.1.1. Решить систему уравнений:

$$\begin{cases} 2x_1 & - & 3x_2 & = & 1\,, \\ 3x_1 & + & x_2 & = & 7\,. \end{cases}$$

Решение. Для данной системы линейных уравнений

$$\|A\| = \left\| \begin{array}{rr} 2 & -3 \\ 3 & 1 \end{array} \right\|, \quad \|X\| = \left\| \begin{array}{c} x_1 \\ x_2 \end{array} \right\|, \quad \|B\| = \left\| \begin{array}{c} 1 \\ 7 \end{array} \right\|.$$

Основная матрица данной системы уравнений невырожденная, поскольку

$$\det \left\| \begin{array}{rr} 2 & -3 \\ 3 & 1 \end{array} \right\| = 2 \cdot 1 - (-3) \cdot 3 = 11 \neq 0.$$

Найдем для нее обратную по формуле (2.1.3.1).

$$\left\| \begin{array}{rr} 2 & -3 \\ 3 & 1 \end{array} \right\|^{-1} = \frac{1}{11} \cdot \left\| \begin{array}{rr} 1 & 3 \\ -3 & 2 \end{array} \right\|.$$

Наконец, по формуле (2.2.1.3) находим

$$\|X\| = \left\| \begin{array}{c} x_1 \\ x_2 \end{array} \right\| = \left\| \begin{array}{rr} 2 & -3 \\ 3 & 1 \end{array} \right\|^{-1} \cdot \left\| \begin{array}{c} 1 \\ 7 \end{array} \right\| = \frac{1}{11} \cdot \left\| \begin{array}{rr} 1 & 3 \\ -3 & 2 \end{array} \right\| \cdot \left\| \begin{array}{c} 1 \\ 7 \end{array} \right\| =$$

$$= \frac{1}{11} \cdot \left\| \begin{array}{r} 1 \cdot 1 + 3 \cdot 7 \\ (-3) \cdot 1 + 2 \cdot 7 \end{array} \right\| = \frac{1}{11} \cdot \left\| \begin{array}{c} 22 \\ 11 \end{array} \right\| = \left\| \begin{array}{c} 2 \\ 1 \end{array} \right\|.$$

Итак, решением данной системы уравнений является пара чисел $x_1 = 2$ и $x_2 = 1$.

2.2.2. Метод Крамера

Хотя формулы (2.2.1.1) и (2.2.1.3) по форме записи аналогичны, известным из курса элементарной алгебры формулам $ax = b$ и $x = a^{-1}b$ для линейного уравнения с одним неизвестным, их практическое применение не целесообразно из-за сравнительно большого объема требующихся вычислений. Существенно более эффективным методом решения систем линейных уравнений является метод Крамера, использование которого рассмотрим на примере системы двух линейных уравнений с двумя неизвестными вида

$$\left\{ \begin{array}{ccccc} a_{11}x_1 & + & a_{12}x_2 & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2, \end{array} \right. \qquad (2.2.2.1)$$

решением которой является упорядоченная пара чисел $\{x_1,\ x_2\}$, а во введенных ранее обозначениях

$$\|A\| = \left\| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right\|, \quad \|X\| = \left\| \begin{array}{c} x_1 \\ x_2 \end{array} \right\|, \quad \|B\| = \left\| \begin{array}{c} b_1 \\ b_2 \end{array} \right\|.$$

Получим два следствия из уравнений данной системы, каждое из которых будет содержать только по одному из неизвестных x_1 и x_2. Для этого сначала вычтем из первого уравнения, умноженного на a_{22}, второе, умноженное на a_{12}, что даст

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = b_1a_{22} - b_2a_{12},$$

затем из второго уравнения, умноженного на a_{11} , вычтем первое, умноженное на a_{21}, и получим

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = b_2a_{11} - b_1a_{21}.$$

Заметим, что оба полученных равенства, которые удовлетворяются любыми решениями системы (2.2.2.1), можно переписать в виде

$$\Delta \cdot x_1 = \Delta_1 \qquad \text{и} \qquad \Delta \cdot x_2 = \Delta_2 \qquad (2.2.2.2),$$

где

$$\Delta = \det \left\| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right\|, \quad \Delta_1 = \det \left\| \begin{array}{cc} b_1 & a_{12} \\ b_2 & a_{22} \end{array} \right\|, \quad \Delta_2 = \det \left\| \begin{array}{cc} a_{11} & b_1 \\ a_{21} & b_2 \end{array} \right\|.$$

и исследуем соотношения (2.2.2.2) в предположении, что на значения параметров $a_{11},\ a_{12},\ a_{21},\ a_{22},$ и $b_1,\ b_2$ не налагается никаких ограничений.

1°. Если $\Delta \neq 0$, существует единственная пара чисел x_1 и x_2, удовлетворяющая равенствам (2.2.2.2) вида

$$x_1 = \frac{\Delta_1}{\Delta} \quad \text{и} \quad x_2 = \frac{\Delta_2}{\Delta}\,. \tag{2.2.2.3}$$

Непосредственной проверкой можно убедиться, что эта пара чисел удовлетворяет и системе (2.2.2.1) Следовательно, при $\Delta \neq 0$ система линейных уравнений (2.2.2.1) имеет решение и притом единственное.

2°. Если $\Delta = 0$, а $\Delta_1 \neq 0$ или $\Delta_2 \neq 0$, то соотношения (2.2.2.2) не удовлетворяются ни при каких значениях x_1 и x_2. Значит, в этом случае система линейных уравнений (2.2.4) решений не имеет.

3°. Наконец, если $\Delta = 0$ и $\Delta_1 = \Delta_2 = 0$, то соотношения (2.2.2.2) будут удовлетворяться любой парой чисел x_1 и x_2. В этом случае система (2.2.2.1) имеет бесчисленное множество решений, поскольку коэффициенты ее уравнений оказываются пропорциональными.

Таким образом, можно утверждать следующее (это утверждение называется **теоремой Крамера**).

> Система n линейных уравнений с n неизвестными *имеет единственное решение* вида (2.2.2.3) тогда и только тогда, когда $\det\|A\| \neq 0$.

Пример 2.2.2.1. Решить методом Крамера систему двух линейных уравнений с двумя неизвестными:

$$\left\{\begin{array}{rcrcl} 2x_1 & - & 3x_2 & = & 1\,, \\ 3x_1 & + & x_2 & = & 7\,. \end{array}\right.$$

Решение. Найдем для данной системы линейных уравнений значение Δ.

$$\Delta = \det\|A\| = \det\left\|\begin{array}{rr} 2 & -3 \\ 3 & 1 \end{array}\right\| = 2\cdot 1 - (-3)\cdot 3 = 11 \neq 0,$$

значит, для решения данной системы мы можем использовать метод Крамера, то есть по формулам (2.2.2.2) вначале вычисляем значения Δ_1 и Δ_2.

$$\Delta_1 = \det\|A\| = \det\left\|\begin{array}{rr} 1 & -3 \\ 7 & 1 \end{array}\right\| = 1\cdot 1 - (-3)\cdot 7 = 22\,,$$

$$\Delta_2 = \det \|A\| = \det \begin{Vmatrix} 2 & 1 \\ 3 & 7 \end{Vmatrix} = 2 \cdot 7 - 1 \cdot 3 = 11 \; .$$

Наконец, по формулам (2.2.2.3) находим искомые значения неизвестных

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{22}{11} = 2 \quad \text{и} \quad x_2 = \frac{\Delta_2}{\Delta} = \frac{11}{11} = 1 \; .$$

Теорема Крамера справедлива в общем случае для системы n линейных уравнений с n неизвестными. В качестве примера продемонстрируем ее применение и для решения системы трех линейных уравнений с тремя неизвестными.

Пример 2.2.2.2. Решить систему линейных уравнений:

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

Решение. Вначале составим основную матрицу (то есть матрицу, составленную из коэффициентов при неизвестных) этой системы и найдем ее определитель, например методом «треугольников»

$$\det \|A\| = \det \begin{Vmatrix} -2 & 1 & -1 \\ 3 & -1 & 2 \\ 2 & 1 & -2 \end{Vmatrix} =$$

$$= (-2)\cdot(-1)\cdot(-2)+1\cdot2\cdot2+(-1)\cdot3\cdot1-(-1)\cdot(-1)\cdot2-1\cdot3\cdot(-2)-(-2)\cdot2\cdot1 =$$

$$= -4+4-3-2+6+4 = 5 \neq 0.$$

Полученный результат означает, что система имеет единственное решение, определяемое формулами:

$$x_1 = \frac{\Delta_1}{\Delta} \, , \quad x_2 = \frac{\Delta_2}{\Delta} \, , \quad x_3 = \frac{\Delta_3}{\Delta} \, ,$$

где $\Delta = \det \|A\| = 5$, а числа Δ_1, Δ_2 и Δ_3 согласно теореме Крамера равны

$$\Delta_1 = \det \begin{Vmatrix} 1 & 1 & -1 \\ -3 & -1 & 2 \\ -2 & 1 & -2 \end{Vmatrix} = -5 \; ; \quad \Delta_2 = \det \begin{Vmatrix} -2 & 1 & -1 \\ 3 & -3 & 2 \\ 2 & -2 & -2 \end{Vmatrix} = -10 \; ;$$

$$\Delta_3 = \det \begin{Vmatrix} -2 & 1 & 1 \\ 3 & -1 & -3 \\ 2 & 1 & -2 \end{Vmatrix} = -5 \; .$$

Напомним, что матрицы для определителей Δ_1, Δ_2 и Δ_3 получаются из матрицы

$$\|A\| = \left\| \begin{array}{rrr} -2 & 1 & -1 \\ 3 & -1 & 2 \\ 2 & 1 & -2 \end{array} \right\|$$

последовательной заменой ее столбцов на $\|B\| = \left\| \begin{array}{r} 1 \\ -3 \\ -2 \end{array} \right\|$ – столбец правых частей решаемой системы уравнений (выделенный в формулах затенением).

Подставив найденные значения чисел Δ_1, Δ_2 и Δ_3 , получим, что

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{-\,5}{5} = -1\ , \quad x_2 = \frac{\Delta_2}{\Delta} = \frac{-\,10}{5} = -2\ , \quad x_3 = \frac{\Delta_3}{\Delta} = \frac{-\,5}{5} = -1\ .$$

2.2.3. Системы m линейных уравнений с n неизвестными

Рассмотрим теперь случай, когда число уравнений и число неизвестных системы линейных уравнений произвольны, то есть систему m линейных уравнений с n неизвестными вида

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array} \right. \qquad (2.2.3.1)$$

при произвольном соотношении между натуральными числами m и n.

Столбец неизвестных, столбец свободных членов, основная и расширенная матрицы в этом случае соответственно будут иметь вид

$$\|X\| = \left\| \begin{array}{c} x_1 \\ x_2 \\ \ldots \\ x_n \end{array} \right\| , \quad \|B\| = \left\| \begin{array}{c} b_1 \\ b_2 \\ \ldots \\ b_m \end{array} \right\| , \quad \|A\| = \left\| \begin{array}{cccc} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{array} \right\| ,$$

$$\|A|B\| = \left\| \begin{array}{cccccc} a_{11} & a_{12} & \ldots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & | & b_2 \\ \ldots & \ldots & \ldots & \ldots & | & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & | & b_m \end{array} \right\| .$$

Условие разрешимости системы линейных уравнений (2.2.3.1) (то есть существования решения) дается следующей теоремой (носящей название **теоремы Кронекера-Капелли**).

> Для того,чтобы система m линейных уравнений с n неизвестными имела решение, необходимо и достаточно, чтобы ранг ее *основной* матрицы равнялся рангу *расширенной* матрицы.

Или, во введенных ранее обозначениях, $\mathrm{rg}\,\|A\| = \mathrm{rg}\,\|A|B\|$.

Метод решения системы (2.2.3.1) основывается на теореме, утверждающей, что *максимальное число независимых уравнений этой системы равно* $\mathrm{rg}\,\|A\|$, и (в случае, если она совместна) состоит в следующих, последовательно выполняемых действиях:

1° - отбрасывании зависимых уравнений;

2° - разделении всех неизвестных на две группы: *основные* и *свободные*, к первым из которых(общим числом $\mathrm{rg}\,\|A\|$), оказывается применимой теорема Крамера, а значения вторых могут иметь произвольные значения;

3° - построении *упрощенной* системы и ее решении по правилу Крамера, сводящемуся к выражению основных неизвестных через свободные.

Пример 2.2.3.1. Проиллюстрируем использование данного метода на примере задачи: исследовать на совместность и найти все решения системы линейных уравнений

$$\left\{\begin{array}{ccccccccccc} 3x_1 & + & 2x_2 & + & x_3 & - & 7x_4 & + & x_5 & = & 0, \\ x_1 & + & x_2 & + & x_3 & - & 4x_4 & + & x_5 & = & 0 \end{array}\right.$$

или в матричном виде

$$\left\|\begin{array}{ccccc} 3 & 2 & 1 & -7 & 1 \\ 1 & 1 & 1 & -4 & 1 \end{array}\right\| \cdot \left\|\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array}\right\| = \left\|\begin{array}{c} 0 \\ 0 \end{array}\right\|.$$

Решаемая система линейных уравнений есть частный случай системы (2.2.3.1) с $m = 2$ и $n = 5$. Найдем ранги ее основной и расширенной матриц.

$$\|A|B\| = \left\|\begin{array}{cccccc} 3 & 2 & 1 & -7 & 1 & | & 0 \\ 1 & 1 & 1 & -4 & 1 & | & 0 \end{array}\right\|$$

Во-первых, очевидно, что ранг основной матрицы не меньше, чем 1, поскольку в матрице есть элементы с ненулевыми значениями, и не больше, чем 2, так как число строк матрицы $m = 2$. Поэтому достаточно подсчитать значения всех миноров второго порядка и выяснить, имеется ли среди них хотя бы один отличный от нуля. Всего в основной матрице миноров второго порядка 24. В их числе, например, имеется не равный нулю минор, матрицу которого образуют первые два столбца. Действительно,

$$\det \begin{Vmatrix} 3 & 2 \\ 1 & 1 \end{Vmatrix} = 3 \cdot 1 - 2 \cdot 1 = 1 \neq 0.$$

Значит, для данной системы линейных уравнений $\mathrm{rg}\,\|A\| = 2$. С другой стороны, ранг расширенной матрицы $\mathrm{rg}\,\|A|B\|$ также очевидно равен 2, поскольку столбец правых частей системы содержит только нули и с «участием» этого столбца нельзя получить ненулевой минор. Следовательно, система совместна – она имеет решение.

Отметим, что для данной системы уравнений к заключению о ее совместности можно было придти и без теоремы Кронекера-Капелли, так как у нее есть очевидное решение: $x_1 = x_2 = x_3 = x_4 = x_5 = 0$. Подобная ситуация будет иметь место для любой *однородной* системы линейных уравнений, то есть в случае, когда правые части всех уравнений равны нулю.

Однако применение теоремы Кронекера-Капелли необходимо, поскольку ранг расширенной матрицы также показывает какое максимальное число независимых уравнений содержит решаемая система. Зависимые же уравнения можно не принимать во внимание. В нашем случае

$$\mathrm{rg}\,\|A\| = \mathrm{rg}\,\|A|B\| = 2 = m,$$

и, согласно пункту 1°, оба уравнения следует оставить в системе.

При выполнении пункта 2° вначале следует определить сколько основных и сколько свободных неизвестных имеет система. В нашем случае число основных неизвестных равно $\mathrm{rg}\,\|A\| = 2$ – максимальному числу независимых уравнений, ибо правило Крамера применимо для систем, в которых число неизвестных равно числу уравнений. Число же свободных неизвестных, таким образом, оказывается равным $n - \mathrm{rg}\,\|A\| = 5 - 2 = 3$.

Теперь из имеющихся в исходной системе пяти неизвестных надо выбрать два в качестве основных. В общем случае *не любая* пара неизвестных для этого годится. Выбор должен позволить применение теоремы Крамера, то есть детерминант основной матрицы упрощенной системы (получающейся после переноса свободных неизвестных в правую часть уравнений) не должен быть равен нулю. В нашем случае, например, пара x_1 и x_2 отвечает этому требованию, а пара x_3 и x_5 – нет.

Если принять за основные неизвестные первую пару, то упрощенная система будет иметь вид:

$$\left\{\begin{array}{rcrcrcrcrcr} 3x_1 & + & 2x_2 & = & - & x_3 & + & 7x_4 & - & x_5, \\ x_1 & + & x_2 & = & - & x_3 & + & 4x_4 & - & x_5. \end{array}\right.$$

Присвоим свободным неизвестным произвольные фиксированные значения:

$$x_3 = t\ , \quad x_4 = p\ , \quad x_5 = q\ .$$

Теперь значения основных неизвестных однозначно находятся из упрощенной системы. Например, по формулам Крамера, или почленным вычитанием удвоенного второго уравнения из первого, получаем

$$\left\{\begin{array}{rcrrcrcr} x_1 & = & & t & - & p & + & q, \\ x_2 & = & - & 2t & + & 5p & - & 2q. \end{array}\right.$$

Следовательно, решение исходной системы линейных уравнений в развернутой форме может быть записано в виде

$$\left\{\begin{array}{rcrrcrcr} x_1 & = & & t & - & p & + & q, \\ x_2 & = & - & 2t & + & 5p & - & 2q, \\ x_3 & = & & t, & & & & \\ x_4 & = & & & & p, & & \\ x_5 & = & & & & & & q, \end{array}\right. \qquad \forall t,\ p,\ q.$$

Полученные формулы можно также записать и при помощи матричных операций.

$$\left\| \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \right\| = \left\| \begin{array}{rrcrcr} & t & - & p & + & q \\ - & 2t & + & 5p & - & 2q \\ & t & & & & \\ & & & p & & \\ & & & & & q \end{array} \right\|$$

или

$$\left\| \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \right\| = t \cdot \left\| \begin{array}{r} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{array} \right\| + p \cdot \left\| \begin{array}{r} -1 \\ 5 \\ 0 \\ 1 \\ 0 \end{array} \right\| + q \cdot \left\| \begin{array}{r} 1 \\ -2 \\ 0 \\ 0 \\ 1 \end{array} \right\| \qquad \forall t,\ p,\ q.$$

Пример 2.2.3.2. Рассмотрим еще один пример: исследовать на совместность и найти все решения системы линейных уравнений

$$\left\{\begin{array}{rcrcrcr} 2x_1 & - & x_2 & + & 2x_3 & = & 3 \\ x_1 & - & x_2 & + & 2x_3 & = & 2 \\ -3x_1 & + & 3x_2 & - & 6x_3 & = & -6 \\ 2x_1 & - & 2x_2 & + & 4x_3 & = & 4 \\ -x_1 & + & x_2 & - & 2x_3 & = & -2 \end{array}\right. \qquad (2.2.3.2)$$

или в матричном виде

$$\left\| \begin{array}{rrr} 2 & -1 & 2 \\ 1 & -1 & 2 \\ -3 & 3 & -6 \\ 2 & -2 & 4 \\ -1 & 1 & -2 \end{array} \right\| \cdot \left\| \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right\| = \left\| \begin{array}{r} 3 \\ 2 \\ -6 \\ 4 \\ -2 \end{array} \right\| .$$

Обратите внимание на размеры матриц в данной форме записи системы уравнений.

Очевидно, что система (2.2.3.2) есть частный случай системы (2.2.3.1) с $m = 5$ и $n = 3$. Построим для нее расширенную матрицу $\|A|B\|$ и найдем как ее ранг, так и ранг основной матрицы $\|A\|$. Имеем

$$\|A|B\| = \left\| \begin{array}{rrrcr} 2 & -1 & 2 & | & 3 \\ 1 & -1 & 2 & | & 2 \\ -3 & 3 & -6 & | & -6 \\ 2 & -2 & 4 & | & 4 \\ -1 & 1 & -2 & | & -2 \end{array} \right\|$$

Несложно заметить, что как в основной, так и в расширенной матрицах третья строка есть вторая умноженная на -3. Это означает, что третье уравнение системы (2.2.3.2) является следствием второго уравнения и его надо удалить из списка независимых уравнений. Аналогично к зависимым следует отнести также четвертое и пятое уравнения, поскольку они могут быть получены умножением второго уравнения соответственно на 2 и -1.

С другой стороны, в первых двух строках расширенной матрицы, в столбцах один и два расположена квадратная подматрица второго порядка

$$\left\| \begin{array}{rr} 2 & -1 \\ 1 & -1 \end{array} \right\| ,$$

определитель которой $\det \left\| \begin{array}{cc} 2 & -1 \\ 1 & -1 \end{array} \right\| = 2 \cdot (-1) - 1 \cdot (-1) = -1 \neq 0$. А это означает, что $\operatorname{rg} \|A\| = \operatorname{rg} \|A|B\| = 2$ и, следовательно, система (2.2.3.2) совместна, то есть она имеет решение.

Найдем эти решения по схеме 1°–3°. После удаления зависимых уравнений, система примет вид

$$\left\{ \begin{array}{rcrcrcl} 2x_1 & - & x_2 & + & 2x_3 & = & 3 \\ x_1 & - & x_2 & + & 2x_3 & = & 2 \end{array} \right. \tag{2.2.3.3}$$

Теперь следует разделить неизвестные на основные и свободные. Число основных неизвестных должно быть $\operatorname{rg} \|A\| = 2$, однако не любая пара неизвестных годится для использования в качестве основных. Напомним: для этой пары неизвестных должна быть применимой теорема Крамера. Другими словами, минор основной матрицы не должен равняться нулю. В нашем примере за основные неизвестные можно выбрать пары $\{x_1, x_2\}$ или $\{x_1, x_3\}$, но не $\{x_2, x_3\}$, так как для последней пары соответствующий минор равен $\det \left\| \begin{array}{cc} -1 & 2 \\ -1 & 2 \end{array} \right\| = (-1) \cdot 2 - 2 \cdot (-1) = 0$.

Примем за основные неизвестные пару $\{x_1, x_2\}$, а неизвестное x_3 за свободное, значение которого t произвольно. Тогда система (2.2.8) может быть записана в виде

$$\left\{ \begin{array}{rcrcrcl} 2x_1 & - & x_2 & = & 3 & - & 2t, \\ x_1 & - & x_2 & = & 2 & - & 2t, \end{array} \right. \tag{2.2.3.4}$$

где $x_3 = t, \quad t \in (-\infty, +\infty)$, то есть t – параметр с произвольным значением.

Решим систему (2.2.3.4) по схеме Крамера. Для этого предварительно найдем Δ, Δ_1 и Δ_2.

$$\Delta = \det \left\| \begin{array}{cc} 2 & -1 \\ 1 & -1 \end{array} \right\| = 2 \cdot (-1) - 1 \cdot (-1) = -1;$$

$$\Delta_1 = \det \left\| \begin{array}{cc} (3-2t) & -1 \\ (2-2t) & -1 \end{array} \right\| = (3-2t) \cdot (-1) - (2-2t) \cdot (-1) = -1;$$

$$\Delta_2 = \det \left\| \begin{array}{cc} 2 & (3-2t) \\ 1 & (2-2t) \end{array} \right\| = 2 \cdot (3-2t) - 1 \cdot (2-2t) = 1 - 2t.$$

Теперь, учитывая, что $x_1 = \dfrac{\Delta_1}{\Delta} = -1$ и $x_2 = \dfrac{\Delta_2}{\Delta} = 2t - 1$ (по теореме Крамера), можно получить ответ задачи: все решения системы линейных уравнений (2.2.3.4) описываются формулами

$$\begin{cases} x_1 &= 1, \\ x_2 &= 2t-1, \qquad t \in (-\infty, +\infty) \\ x_3 &= t. \end{cases}$$

или, при помощи матричных операций,

$$\left\| \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right\| = \left\| \begin{array}{r} 1 \\ -1 \\ 0 \end{array} \right\| + t \cdot \left\| \begin{array}{c} 0 \\ 2 \\ 1 \end{array} \right\|, \qquad t \in (-\infty, +\infty)$$

2.3. Иллюстративный пример: «задача о жуках»

Продемонстрируем использование матричных операций и систем линейных уравнений на примере следующей задачи.

Рассматривается колония жуков определенного вида, средой обитания которых при нормальных условиях может быть одна из трех следующих: «на берегу», «в воздухе» и «на воде». Последовательно проводимые наблюдения показали, что:

- жук, находящийся «на берегу», при последующем наблюдении, остается «на берегу» в тридцати процентах случаев, в сорока процентах случаев обнаруживается «в воздухе», и, наконец, в тридцати процентах случаев оказывается «на воде»;
- жук, находящийся «в воздухе», при последующем наблюдении оказывается «на берегу» в половине всех случаев, обнаруживается «в воздухе» в одной десятой части случаев или находится «на воде» в двух пятых случаев;
- жук, находящийся «на воде», при последующем наблюдении обнаруживается «на берегу» в двух пятых всех случаев, находится «в воздухе » в одной пятой случаев или оказывается «на воде» в двух пятых случаев.

В предположении, что полная численность колонии жуков постоянна, требуется найти:

1) распределение жуков по средам обитания, если предшествующее наблюдение показало, что «на берегу« находилось 30% жуков, «в воздухе» - 10% жуков, а «на воде» 60%.

2) распределение жуков по средам обитания, предшедствующее наблюдению, которое показало, что «на берегу« находилось 45% жуков, «в воздухе» - 15% жуков, а «на воде» 40%.

Решение. 1) Пусть упорядоченная тройка чисел $\{x_1, x_2, x_3\}$ есть исходное распределение жуков по средам обитания: «на берегу», «в воздухе » и соответственно «на воде» (в процентах), а тройка чисел $\{y_1, y_2, y_3\}$ является искомым аналогичным распределением. Тогда связь между исходным и искомым распределением, как нетрудно заметить, задается формулами:

$$\begin{aligned} y_1 &= 0.3 \cdot x_1 + 0.5 \cdot x_2 + 0.4 \cdot x_3 \\ y_2 &= 0.4 \cdot x_1 + 0.1 \cdot x_2 + 0.2 \cdot x_3 \\ y_3 &= 0.3 \cdot x_1 + 0.4 \cdot x_2 + 0.4 \cdot x_3 \end{aligned} \tag{2.3.1}$$

Действительно, например y_2 - процентная доля жуков «в воздухе » образуется из:

- доли жуков бывших «на берегу» и взлетевших «в воздух» (которых, согласно условию задачи 40% от x_1),
- доли жуков бывших «в воздухе» и там же оставшихся (а таких, согласно условию задачи 10% от x_2),
- и, наконец, доли жуков бывших «на воде» и оттуда взлетевших «в воздух» (таких, согласно условию задачи 20% от x_3).

Просуммировав эти три доли, мы и получаем искомое значение для y_2. Величины y_1 и y_3 находятся аналогично.

Теперь, подставив в соотношения (2.3.1) данные в условии задачи значения $x_1 = 10\%, x_2 = 65\%$ и $x_3 = 25\%$, получим путем несложных арифметических вычислений ответ на первый вопрос задачи: $y_1 = 45.5\%, y_2 = 15.5\%$ и $y_3 = 39\%$.

Заметим, что упорядоченные тройки чисел, задающие распределение жуков по средам обитания, можно записать в виде трехкомпонентных столбцов

$$\|X\| = \left\| \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right\|; \quad \|Y\| = \left\| \begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} \right\|,$$

а доли, характеризующие изменение среды обитания жуков, в виде матрицы

$$\|A\| = \left\| \begin{array}{ccc} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{array} \right\|.$$

Обратите внимание, что столбцы этой матрицы соответствуют среде, в которой жук находился, а строки – среде в которую он переместился.

Теперь, согласно определению операции произведения матриц - 2.1.2.5, соотношения (2.3.1) можно записать как $\|Y\| = \|A\| \cdot \|X\|$ или в развернутом виде

$$\left\| \begin{matrix} y_1 \\ y_2 \\ y_3 \end{matrix} \right\| = \left\| \begin{matrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{matrix} \right\| \cdot \left\| \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix} \right\|. \tag{2.3.2}$$

Отметим сразу, что, так же как и соотношения (2.3.1), эта формула позволяет получить ответ на вопросы задачи. Действительно, если в (2.3.2) подставить $x_1 = 10\%, x_2 = 65\%$ и $x_3 = 25\%$, то мы получим согласно определению 2.1.2.5.

$$\left\| \begin{matrix} y_1 \\ y_2 \\ y_3 \end{matrix} \right\| = \left\| \begin{matrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{matrix} \right\| \cdot \left\| \begin{matrix} 10 \\ 65 \\ 25 \end{matrix} \right\| = \left\| \begin{matrix} 45.5 \\ 15.5 \\ 39.0 \end{matrix} \right\|,$$

что естественно совпадает с ответом на первый вопрос задачи, полученным нами по формулам (2.3.1.)

Получим теперь ответ на второй вопрос задачи. Нам снова необходимо решить уравнение $\|Y\| = \|A\| \cdot \|X\|$, но только теперь неизвестным будет столбец $\|X\|$, в то время как столбец $\|Y\|$ задан и имеет компоненты $y_1 = 45\%, y_2 = 15\%$ и $y_3 = 40\%$.

Для нахождения $\|X\|$ вполне возможно подставить значения компонентов столбца $\|Y\|$ в соотношения (2.3.1) и попытаться решить полученную систему трех линейных уравнений с тремя неизвестными, например, методом исключения, который был рассмотрен в начале данного параграфа. Однако возможен и другой подход, основанный на том факте, что согласно определению 2.1.3.2.

$$\det \|A\| = \det \left\| \begin{matrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{matrix} \right\| = -0.01 \neq 0, \tag{2.3.3}$$

а это, в свою очередь, означает что матрица $\|A\|$ имеет обратную, которая равна

$$\|A\|^{-1} = \left\| \begin{matrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{matrix} \right\|^{-1} = \left\| \begin{matrix} 4 & 4 & -6 \\ 10 & 0 & -10 \\ -13 & -3 & 17 \end{matrix} \right\|.$$

Этот результат мы приводим без доказательства, однако рекомендуем самостоятельно проверить справедливость равенств (2.3.3), а также соотношений $\|A\|^{-1} \cdot \|A\| = \|A\| \cdot \|A\|^{-1} = \|E\|$, где $\|E\|$ *единичная* матрица (см. §2.1.1.)

Как было показано ранее в данном параграфе, если $\det \|A\| \neq 0$, то уравнение $\|A\| \cdot \|X\| = \|Y\|$ равносильно уравнению $\|X\| = \|A\|^{-1} \cdot \|Y\|$, или в развернутом виде

$$\begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = \begin{Vmatrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{Vmatrix}^{-1} \cdot \begin{Vmatrix} y_1 \\ y_2 \\ y_3 \end{Vmatrix}.$$

После подстановки конкретных значений для компонентов матрицы $\|A\|^{-1}$ и столбца $\|Y\|$, и выполнения операции умножения матриц, получим

$$\begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = \begin{Vmatrix} 4 & 4 & -6 \\ 10 & 0 & -10 \\ -13 & -3 & 17 \end{Vmatrix} \cdot \begin{Vmatrix} 45 \\ 15 \\ 40 \end{Vmatrix} = \begin{Vmatrix} 0 \\ 50 \\ 50 \end{Vmatrix}.$$

То есть ответ на второй вопрос задачи будет таким: во время предшедствовавшего наблюдения «на берегу» жуков не было, а «в воздухе» и «на воде» находилось по 50% состава колонии.

2.4. Собственные векторы и собственные значения квадратных матриц

2.4.1. Определение и метод вычисления

При решении большого числа задач в прикладных разделах математики, включая и математическую статистику, возникает так называемая «задача на собственные значения» имеющая следующую постановку.

Пусть дана квадратная, порядка n матрица $\|A\|$ вида

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

Определение 2.4.1.1. **Собственным вектором** квадратной матрицы $\|A\|$ называется *ненулевой* n-компонентный столбец $\|X\|$ такой, что выполнено равенство

$$\|A\|\,\|X\| = \lambda\,\|X\|\,, \tag{2.4.1.1}$$

где λ – некоторое число, называемое **собственным значением** матрицы $\|A\|$.

Принято говорить, что в случае выполнения равенства (2.4.1.1), собственный вектор $\|X\|$ *отвечает* собственному значению λ, а собственное значение λ *отвечает* собственному вектору $\|X\|$. «Задача на собственные значения» заключается в отыскании для данной квадратной матрицы $\|A\|$ всех ее собственных векторов и отвечающих им собственных значений λ.

Пусть столбец $\|X\|$ представлен в развернутой форме как

$$\|X\| = \left\|\begin{matrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{matrix}\right\|\,,$$

тогда равенство (2.4.1.1) имеет вид

$$\left\|\begin{matrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{matrix}\right\| \left\|\begin{matrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{matrix}\right\| = \lambda \left\|\begin{matrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{matrix}\right\|. \tag{2.4.1.2}$$

Если выполнить все операции с матрицами как в левой, так и в правой части данного равенства, то мы придем к системе уравнений вида

$$\left\{\begin{matrix} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & \lambda x_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & \lambda x_2, \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & = & \lambda x_n. \end{matrix}\right. \tag{2.4.1.3}$$

Таким образом, решение задачи на собственные значения сводится к отысканию всех $\{x_1, x_2, \ldots, x_n\}$ и всех чисел λ, удовлетворяющих уравнениям системы (2.4.1.3). Отметим, что эта система n *нелинейных* уравнений с $n+1$ неизвестным, поскольку неизвестными являются все $x_1, x_2, \ldots, x_n$ и λ, а каждое из ее уравнений содержит *произведение* λ на одно из неизвестных $x_1, x_2, \ldots, x_n$.

Решим задачу на собственные значения методом, который не зависит от значения n, и потому (ради наглядности) ограничимся случаем $n = 2$, когда система (2.4.1.2) имеет вид

$$\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix} \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = \lambda \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} \quad \text{или же, в форме (2.4.1.3),}$$

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

Преобразуем ее к виду

$$\begin{Vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{Vmatrix} \begin{Vmatrix} x_1 \\ x_2 \end{Vmatrix} = \begin{Vmatrix} 0 \\ 0 \end{Vmatrix} \quad \text{или}$$

$$\begin{cases} (a_{11} - \lambda)x_1 & + & a_{12}x_2 & = 0, \\ a_{21}x_1 & + & (a_{22} - \lambda)x_2 & = 0. \end{cases} \qquad (2.4.1.4)$$

Предположим, что собственное значение λ является параметром в системе (2.4.1.4), то есть его значение фиксировано, хотя и остается пока не известным. Тогда эту систему можно рассматривать как систему линейных уравнений с неизвестными $\{x_1, x_2\}$.

Легко видеть, что эта система *однородная* (правые части всех ее уравнений равны нулю) и имеет решение $x_1 = x_2 = 0$. Однако столбец

$$\|X\| = \begin{Vmatrix} 0 \\ 0 \end{Vmatrix} .$$

не может быть собственным вектором, поскольку последний должен быть *ненулевым.* С другой стороны, по теореме Крамера система n линейных уравнений с n неизвестными имеет единственное решение тогда и только тогда, когда определитель ее основной матрицы отличен от нуля. Значит, если определитель матрицы, составленной из коэффициентов при неизвестных в системе (2.4.1.4), отличен от нуля, решение у этой системы единственное и притом нулевое. Следовательно, задача на собственные значения в этом случае не имеет решений.

Если же определитель основной матрицы системы (2.4.1.4) равен нулю, то, как было показано в §2.2.2, решений либо вовсе нет, либо их неограничено много. А поскольку нулевое решение у системы (2.4.1.4) всегда имеется, то *равенство нулю определителя основной матрицы этой системы – достаточное условие существования ненулевых решений, то есть собственных векторов.*

Определитель основной матрицы системы (2.4.1.4) будет равен нулю, если

$$\det \left\| \begin{array}{cc} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{array} \right\| = 0.$$

или $(a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} = 0$. Откуда получаем, что значения параметра λ, для которых система (2.4.1.4) имеет ненулевые решения x_1 и x_2, должны быть корнями квадратного уравнения вида

$$\lambda^2 - (a_{11}+a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0, \qquad (2.4.1.5)$$

называемого *характеристическим.*

Таким образом решение задачи на собственные значения для квадратной матрицы второго порядка сводится к следующей последовательности действий.

1) Составляется и решается характеристическое уравнение (2.4.1.5). Квадратное уравнение, как известно, может иметь два различных корня, иметь один корень, либо не иметь решений вовсе. Если корней нет, то решение задачи на собственные значения завершается констатацией факта их отсутствия.

2) Если корни имеются, то для *каждого* из них решается система (2.4.1.4) и находятся все ее ненулевые решения, которые и являются искомыми собственными векторами.

Пример 2.4.1.1. Решить задачу на собственные значения для матрицы

$$\|A\| = \left\| \begin{array}{cc} 2 & 1 \\ 1 & 2 \end{array} \right\|.$$

Решение. 1) Характеристическое уравнение будет

$$\det \left\| \begin{array}{cc} 2-\lambda & 1 \\ 1 & 2-\lambda \end{array} \right\| = 0.$$

или $\lambda^2 - 4\lambda + 3 = 0$. Его корни $\lambda_1 = 1$ и $\lambda_2 = 3$.

2) Для собственного значения $\lambda_1 = 1$ система (2.4.1.4) имеет вид

$$\left\| \begin{array}{cc} 2-\lambda_1 & 1 \\ 1 & 2-\lambda_1 \end{array} \right\| \left\| \begin{array}{c} x_1 \\ x_2 \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \end{array} \right\| \quad \text{или} \quad \left\{ \begin{array}{ccccc} x_1 & + & x_2 & = 0, \\ x_1 & + & x_2 & = 0, \end{array} \right.$$

решением которой будет столбец вида $\|X\| = \left\| \begin{array}{r} t \\ -t \end{array} \right\|$, где t – любое число. Следовательно собственному значению $\lambda_1 = 1$ будет отвечать собственный вектор $\|X\| = t \left\| \begin{array}{r} 1 \\ -1 \end{array} \right\|$ при любом $t \neq 0$.

Аналогично получаем, что для собственного значения $\lambda_2 = 3$ система (2.4.1.4) имеет вид

$$\left\| \begin{array}{cc} 2-\lambda_2 & 1 \\ 1 & 2-\lambda_2 \end{array} \right\| \left\| \begin{array}{c} x_1 \\ x_2 \end{array} \right\| = \left\| \begin{array}{c} 0 \\ 0 \end{array} \right\| \qquad \text{или} \qquad \left\{ \begin{array}{rcrl} -x_1 & + & x_2 & = 0, \\ x_1 & + & -x_2 & = 0, \end{array} \right.$$

решением которой будет столбец вида $\|X\| = \left\| \begin{array}{c} t \\ t \end{array} \right\|$, где t – любое число. Следовательно собственному значению $\lambda_2 = 3$ будет отвечать собственный вектор $\|X\| = t \left\| \begin{array}{c} 1 \\ 1 \end{array} \right\|$ при любом $t \neq 0$.

Пример 2.4.1.2. Решить задачу на собственные значения для матрицы

$$\|A\| = \left\| \begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array} \right\|.$$

Решение. 1) Характеристическое уравнение будет

$$\det \left\| \begin{array}{rr} 1-\lambda & 1 \\ 0 & 1-\lambda \end{array} \right\| = 0.$$

или $(1-\lambda)^2 = 0$. У него корень единственный: $\lambda = 1$.
2) Для собственного значения $\lambda = 1$ система (2.4.1.4) имеет вид

$$\left\{ \begin{array}{rcrl} 0x_1 & + & x_2 & = 0, \\ 0x_1 & + & 0x_2 & = 0, \end{array} \right.$$

решением которой будет столбец вида $\|X\| = \left\| \begin{array}{c} t \\ 0 \end{array} \right\|$, где t – любое число. Следовательно собственному значению $\lambda = 1$ будет отвечать собственный вектор $\|X\| = t \left\| \begin{array}{c} 1 \\ 0 \end{array} \right\|$ при любом $t \neq 0$.

Пример 2.4.1.3. Решить задачу на собственные значения для матрицы

$$\|A\| = \left\| \begin{array}{rr} 1 & 1 \\ -1 & 1 \end{array} \right\|.$$

Решение. 1) Характеристическое уравнение будет

$$\det \begin{Vmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{Vmatrix} = 0.$$

или $\lambda^2 - 2\lambda + 2 = 0$, которое корней не имеет. Следовательно, данная матрица не имеет собственных значений и собственных векторов.

Иногда задачу нахождения собственных значений удается решить более просто, используя какую-либо особенность матрицы $\|A\|$. Например, заметив, что характеристическое уравнение (2.4.1.5) можно записать в виде

$$\lambda^2 - (a_{11} + a_{22})\lambda + \det \|A\| = 0,$$

в случае $\det \|A\| = 0$ очевидно, что $\lambda_1 = 0$ и $\lambda_2 = a_{11} + a_{22}$.

2.4.2. Иллюстративные примеры

Вернемся теперь к «задаче о жуках» (см. п.2.3) и выясним существует ли распределение числа жуков по средам обитания, которое не будет меняться от наблюдения к наблюдению, то есть будет *стационарным.*

Математически условие стационарности некоторого состояния колонии жуков

$$\|X\| = \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix}$$

с матрицей, характеризующей их поведение, вида

$$\|A\| = \begin{Vmatrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{Vmatrix}.$$

может быть записано как $\|A\| \cdot \|X\| = \lambda \|X\|$ или в развернутой форме

$$\begin{Vmatrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{Vmatrix} \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix} = \lambda \begin{Vmatrix} x_1 \\ x_2 \\ x_3 \end{Vmatrix}. \qquad (2.4.2.1)$$

Таким образом, мы пришли к задаче вида (2.4.1.1), которой надо найти *собственный вектор* матрицы $\|A\|$, отвечающий собственному значению $\lambda = 1$. Решим эту задачу, приняв для определенности общую численность колонии жуков равной 238.

Проверим вначале, что $\lambda = 1$ является собственным значением матрицы $\|A\|$. Это условие, как было показано выше, имеет вид

$$\det\left(\|A\| - \lambda\,\|E\|\right) = 0.$$

В нашем случае

$$\det\left(\left\|\begin{array}{ccc} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{array}\right\| - 1\cdot\left\|\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right\|\right) = \det\left\|\begin{array}{rrr} -0.7 & 0.5 & 0.4 \\ 0.4 & -0.9 & 0.2 \\ 0.3 & 0.4 & -0.6 \end{array}\right\| =$$

$$= 10^{-3}\cdot\det\left\|\begin{array}{rrr} -7 & 5 & 4 \\ 4 & -9 & 2 \\ 3 & 4 & -6 \end{array}\right\| =$$

используя правило «треугольников» для подсчета значения определителя

$$= 10^{-3}\cdot\Big((-7)\cdot(-9)\cdot(-6)+5\cdot2\cdot3+4\cdot4\cdot4-4\cdot(-9)\cdot3-5\cdot4\cdot(-6)-(-7)\cdot2\cdot4\Big) =$$

$$= 10^{-3}\cdot\Big(-378+30+64+108+56+120\Big) = 10^{-3}\cdot\Big(-378+378\Big) = 0\,.$$

Найдем теперь решение матричного уравнения (2.4.2.1) с $\lambda = 1$. Для чего запишем это уравнение в развернутом виде

$$\left\{\begin{array}{ccccccc} 0.3x_1 & + & 0.5x_2 & + & 0.4x_3 & = & x_1, \\ 0.4x_1 & + & 0.1x_2 & + & 0.2x_3 & = & x_2, \\ 0.3x_1 & + & 0.4x_2 & + & 0.4x_3 & = & x_3, \end{array}\right.$$

или

$$\left\{\begin{array}{ccccccc} -0.7x_1 & + & 0.5x_2 & + & 0.4x_3 & = & 0, \\ 0.4x_1 & - & 0.9x_2 & + & 0.2x_3 & = & 0, \\ 0.3x_1 & + & 0.4x_2 & - & 0.6x_3 & = & 0. \end{array}\right. \qquad (2.4.2.2)$$

Нам нужно найти *ненулевое* решение системы линейных уравнений (2.4.2.2.) Заметим, что существование для нее ненулевого решения гарантируется (что было показано выше) равенством нулю определителя ее основной матрицы

$$\det\left\|\begin{array}{rrr} -0.7 & 0.5 & 0.4 \\ 0.4 & -0.9 & 0.2 \\ 0.3 & 0.4 & -0.6 \end{array}\right\| = 0.$$

Ранг основной матрицы этой системы равен 2, поскольку у нее есть ненулевой минор второго порядка (расположенный в левом верхнем углу), а единственный минор третьего порядка – определитель основной матрицы нулевой. Значит, в системе (2.4.2.2.) только два независимых уравнения. Найдем ее решение, заменив последнее уравнение на условие $x_1 + x_2 + x_3 = 238$, означающее, что общее число жуков остается постоянным. В итоге получим систему уравнений

$$\left\{ \begin{array}{rcrcrcr} -0.7x_1 & + & 0.5x_2 & + & 0.4x_3 & = & 0, \\ 0.4x_1 & - & 0.9x_2 & + & 0.2x_3 & = & 0, \\ x_1 & + & x_2 & + & x_3 & = & 238, \end{array} \right. \tag{2.4.2.3}$$

решение которой – искомое стационарное распределение числа жуков по средам обитания, имеет вид

$$\|X\| = \left\| \begin{array}{c} 92 \\ 60 \\ 86 \end{array} \right\|.$$

Проверьте самостоятельно, что определитель основной матрицы системы (2.4.2.3) отличен от нуля, и следовательно, в силу теоремы Крамера, найденное стационарное распределение числа жуков по средам обитания *единственное.*

Глава 3.

Функции, последовательности и пределы

3.1. Определение функции

На практике достаточно часто приходится иметь дело с так называемыми *переменными величинами*, то есть численными характеристиками, могущими принимать различные значения. Такие количественные характеристики принято называть просто *переменными*. Например, для описания конкретного человека можно использовать переменные: возраст, рост, вес, коэффициент интеллекта IQ и т.п. При этом нередко оказывается, что значения одной переменной связаны со значениями другой. Скажем, вес человека зависит от его роста, рост – от возраста, обменный курс валюты – от времени и т.д.

В некоторых, вообще говоря, довольно не частых случаях, зависимость одной переменной величины от другой оказывается *однозначной*, то есть для каждого допустимого значения второй переменной значение первой существует и единственно. Например, площадь круга однозначно зависит от его радиуса, возраст человека имеет единственное значение в каждый момент времени, масса однородного тела однозначно определяется его объемом.

Зависимости между переменными величинами, обладающие такими свойствами, принято называть *функциональными.* Они являются объектом изучения математического анализа - раздела курса высшей математики, и играют важную роль в большом числе теоретических и прикладных дисциплин.

Определение 3.1.1. Будем говорить, что задана *функциональная зависимость* или, просто *функция*, если указано **правило**, по которому **каждому** числу x, принадлежащему числовому множеству X, поставлено в соответствие **единственное** число y, принадлежащему числовому множеству Y.

Множество X принято называть *областью определения* функции, а множество Y - *областью ее значений.* Саму функцию принято обозначать

$$y = f(x),\ x \in X,\ y \in Y.$$

Наконец, x - независимая переменная, называется *аргументом*, а y - зависимая переменная, *значением функции* или, просто, *функцией.*

Схематически функциональную зависимость можно представить как объект состоящий из трех компонентов: области определения, области значений и правила, по которому каждому числу из множества определения поставлено в соответствие единственное число из области значений. (См. рис. 3.1). Как область определения, так и область значения – это числовые множества. Правило соответствия может иметь различные формы представления: в виде таблицы, математической формулы, графика или являться некоторой математической задачей определения y по значению x . Наконец, это правило может быть просто описано словесно.

Следует иметь в виду, что достаточно часто функцию задают только формулой соответствия. В этом случае *предполагается*, что областью определения является множество чисел, для которых выполнимы все использованные в записи этой формулы операции. За область значений при этом принимается множество чисел, являющихся значениями функции соответствующие *всем возможным* значениям аргумента.

В соответствии с этим соглашением, можно сказать, что область определения X, например, функции $y = \sqrt{x-3}$, составляют все дей-

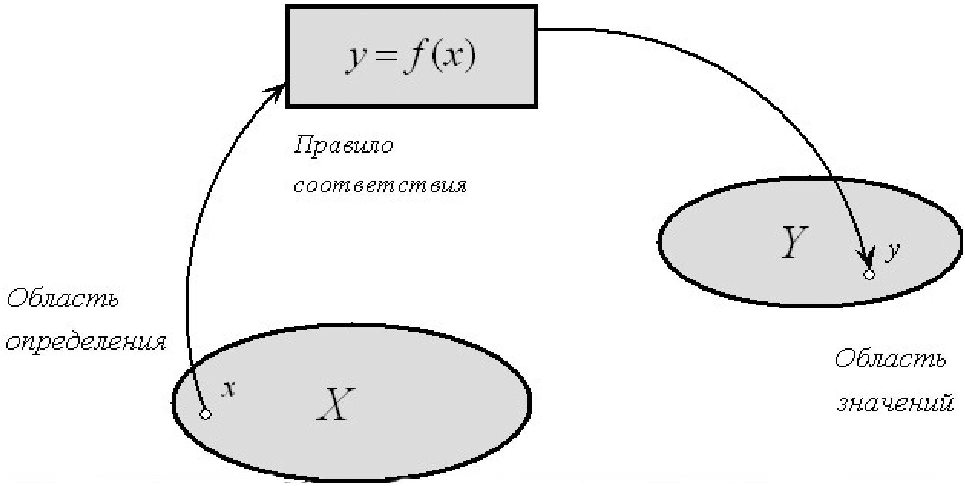


Рис. 3.1. Определение функциональной зависимости

ствительные числа не меньшие, чем 3, поскольку извлечь арифметический квадратный корень (согласно его определению, см. п.2° §1.1) можно только из неотрицательного числа. Множество значений Y согласно этому же определению содержит все неотрицательные числа. Символически это можно записать в виде:

$$X:\{\forall x\geq 3\},\ Y:\{\forall y\geq 0\}\text{ или же }X:\{[3,+\infty)\},\ Y:\{[0,+\infty)\}.$$

Заметим, что задача построения области определения и области значений не всегда оказывается столь тривиальной. Проиллюстрируем это следующими примерами.

Пример 3.1.1. Найти область определения и область значений для функций:

а) $y=\sqrt{\dfrac{2x+3}{x-2}}$

Область определения: решив неравенство $\dfrac{2x+3}{x-2}\ \geq\ 0$, получим $\left[\begin{array}{l} x\leq -\dfrac{3}{2} \\ x>2, \end{array}\right.$ то есть,

$$X:\ \left\{(-\infty,-\frac{3}{2}]\bigcup(2,+\infty)\right\}\ ,$$

поскольку извлечение квадратного корня возможно только из неотрицательного числа. Отметим также, что число 2 не принадлежит области определения, поскольку при таком значении x знаменатель подкоренного выражения обратится в 0 , а деление на нуль невозможно.
Область значений: чтобы найти область значений, рассмотрим формулу $y = \sqrt{\dfrac{2x+3}{x-2}}$ как *уравнение с неизвестным* x *и параметром* y . Выясним, при каких значениях y существует x — вещественный корень этого уравнения. Несложные выкладки приводят к $x = \dfrac{2y^2+3}{y^2-2}$, что означает существование вещественного x при любых $y \neq \pm\sqrt{2}$. С другой стороны, значение функции в рассматриваемом примере является арифметическим квадратным корнем и, значит, неотрицательно. Объединив найденные ограничения на величину y , получим, что

$$Y: \ \left\{[0,\sqrt{2})\bigcup(\sqrt{2},+\infty)\right\} .$$

б) $y = x + \dfrac{1}{x}$
Область определения: очевидно $X : \{\forall x \neq 0\}$ или, что то же самое,

$$\left\{(-\infty,0)\bigcup(0,+\infty)\right\}.$$

Область значений: оценку области значений данной функции выполним в два этапа. Сначала рассмотрим случай $x > 0$. Для любых таких значений x будет справедливо неравенство

$$\left(\sqrt{x}-\frac{1}{\sqrt{x}}\right)^2 \geq 0 \text{ , или } \left(x-2\sqrt{x}\frac{1}{\sqrt{x}}+\frac{1}{x}\right) \geq 0 \text{ , откуда } \quad x+\frac{1}{x} \geq 2.$$

Для случая $x < 0$ оценку области значений можно получить, воспользовавшись равенством

$$(-x)+\frac{1}{(-x)} = -\left(x+\frac{1}{x}\right),$$

из которого в силу неравенства

$$x+\frac{1}{x} \geq 2 \text{ для всех } x > 0$$

имеем

$$x + \frac{1}{x} \leq -2 \text{ для всех } x < 0.$$

Окончательно получаем, что областью значений данной функции является множество $Y : \{(-\infty, -2] \bigcup [2, +\infty)\}$.

в) $y = \dfrac{2x^2 + 2x + 1}{x^2 + 3x + 2}$

Область определения: находится из условия

$$x^2 + 3x + 2 \neq 0 \quad \Leftrightarrow \quad \left[\begin{array}{l} x \neq -2 \\ x \neq -1, \end{array} \right.$$

поскольку знаменатель дроби не может принимать нулевых значений. Других ограничений на вычисление значений функции нет, поэтому область определения будет

$$X : \Big\{(-\infty, -2) \bigcup (-2, -1) \bigcup (-1, +\infty)\Big\} .$$

Область значений: область значений данной функции удобно находить, используя тот факт, что ее область определения образуется произвольными вещественными числами (за исключением -2 и -1.) Рассмотрим формулу $y = \dfrac{2x^2 + 2x + 1}{x^2 + 3x + 2}$ как уравнение с неизвестной x и решим его, считая y некоторым фиксированным параметром. Для этого преобразуем данное равенство к виду, стандартному для квадратных уравнений

$$(y - 2)x^2 + (3y - 2)x + (2y - 1) = 0 \,, \tag{3.1.1}$$

корни которого определяются формулой

$$x_{1,2} = \frac{-\,(3y - 2) \pm \sqrt{D}}{2(y - 2)} \quad y \neq 2 \,,$$

где дискриминант $D = (3y - 2)^2 - 4(y - 2)(2y - 1)$. Условие существования вещественных значений x будет $D \geq 0$, или, в нашем случае,

$$(3y - 2)^2 - 4(y - 2)(2y - 1) =$$

$$= y^2 + 8y - 4 \geq 0 \quad \Leftrightarrow \quad y \in \Big\{ \, (-\infty, -2\sqrt{5} - 4] \bigcup [2\sqrt{5} - 4, +\infty)\Big\} \;.$$

Иначе говоря, x может принимать вещественные значения лишь либо при $y \leq -2\sqrt{5}-4 \approx -8.4$, либо при $y \geq 2\sqrt{5}-4 \approx 0.4$. Следовательно, область значений данной функции образуют числа y, удовлетворяющие либо первому, либо второму из полученных неравенств и не равные 2.

Наконец заметим, что, хотя приведенные рассуждения не применимы для $y = 2$, ибо в этом случае уравнение (3.1.1) не квадратное, а линейное – вида $4x+3 = 0$, тем не менее число 2 принадлежит области значений, поскольку у этого линейного уравнения имеется вещественное решение $x = -\dfrac{3}{4}$, являющееся значением аргумента при котором значение функции равно 2. Следовательно,

$$Y : \left\{ (-\infty, -2\sqrt{5}-4] \bigcup [2\sqrt{5}-4, +\infty) \right\} .$$

В заключение обсуждения понятия функциональной зависимости отметим, что функции принято классифицировать по наличию или отсутствую у нее свойства *периодичности* и свойства *четности.*

Определение 3.1.2. Функция $y = f(x)$ называется *периодической*, если существует число $T \neq 0$ такое, что для любого $x \in X$ выполнено $x \pm T \in X$ и $f(x+T) = f(x)$. Число T в этом случае называется *периодом* функции $y = f(x)$.

Пример 3.1.2. К периодическим относятся следующие функции:

$$\begin{aligned}
&y = \sin x && \text{с периодом } T = 2\pi,\\
&y = \cos 3x && \text{с периодом } T = \frac{2\pi}{3},\\
&y = \operatorname{tg} x && \text{с периодом } T = \pi.
\end{aligned}$$

Определение 3.1.3. Пусть X – область определения функции $y = f(x)$, симметрична относительно точки $x = 0$, тогда эта функция называется:
четной, если $\forall x \in X$ выполнено $f(-x) = f(x)$,
нечетной, если $\forall x \in X$ выполнено $f(-x) = -f(x)$,

Пример 3.1.3. Классификация функций по четности:

$$\begin{aligned}
&y = x^2 && \text{– четная,}\\
&y = x^3 && \text{– нечетная,}\\
&y = \sin x && \text{– нечетная,}\\
&y = \cos x && \text{– четная,}\\
&y = 3^x && \text{– не является ни четной, ни нечетной.}
\end{aligned}$$

Следует иметь в виду, что, хотя существуют функции не относящиеся ни к четным, ни к нечетным, в симметричной области определения каждую их них можно представить как сумму некоторой четной функции и некоторой нечетной. Для этого можно использовать, например, формулу

$$f(x) = \frac{f(x)+f(-x)}{2} + \frac{f(x)-f(-x)}{2}\,.$$

Так для функции $y = 3^x$ разложение в сумму четной и нечетной будет иметь вид

$$y = \frac{3^x+3^{-x}}{2} + \frac{3^x-3^{-x}}{2}\,.$$

3.2. Последовательности и их пределы

3.2.1. Числовые последовательности

Будем говорить, что задана *числовая последовательность*, если указано правило, согласно которому каждому натуральному числу (номеру) n поставлено в соответствие единственное число x_n, называемое значением n-го члена последовательности. Числовую последовательность принято обозначать как $\{x_n\}$.

Согласно этому определению числовую последовательность можно рассматривать как *функцию натурального ряда чисел,* то есть как функцию, областью определения которой является множество всех натуральных чисел.

Числовую последовательность можно задать одним из следующих трех способов:

1) *Перечислением значений ее членов.* Например, последовательность $\{x_n\}$, у которой все члены с четными номерами равны 1, а все члены с нечетными -1, может быть представлена в виде $\{-1, 1, -1, 1, -1, 1, \ldots\}$
2) *Функциональным правилом,* которое для каждого члена последовательности позволяет однозначно определить его значение по

его номеру. Например, для рассмотренной в 1) последовательности, таким правилом могут быть формулы

$$x_n = (-1)^n \quad \text{или} \quad x_n = \sin\left(\frac{\pi}{2} + \pi n\right) \ , \quad n = 1, 2, \ldots$$

3) *Рекурсивным правилом,* по которому значение каждого члена последовательности может быть однозначно определено по значению одного или нескольких предыдущих членов. Например, для рассмотренной в 1) последовательности, таким правилом могут служить соотношения

$$x_{n+1} = (-1) \cdot x_n \ , \quad x_1 = -1 \ , \quad n = 1, 2, \ldots$$

3.2.2. Классификация числовых последовательностей

Числовые последовательности принято различать по множеству значений их членов. Например,

– последовательность, все члены которой имеют значение одного знака называется *знакопостоянной*,
– последовательность, все члены которой имеют значение, не превосходящего по модулю некоторого фиксированного числа, называется *ограниченной.*

Заметим, что определение ограниченной последовательности удобно давать, используя логические символы (см. п.10° §1.1.) Например, последовательность $\{x_n\}$ называется *ограниченной*, если

$$\exists C \geq 0 : \forall n : |x_n| \leq C \ ,$$

то есть, найдется неотрицательное число C такое, что для любого номера n будет выполнено неравенство $|x_n| \leq C$. Если последнее неравенство имеет вид $x_n \leq C$ (или $x_n \geq C$), то говорят об *ограниченной сверху* (или, соответственно, *ограниченной снизу*) числовой последовательности.

Сформулируем теперь определение *неограниченной* числовой последовательности, имея в виду, что *отрицание* некоторого определения должно строиться с соблюдением правил формальной логики. Например, формулировка «не существует число C такое, что ...» не является ошибочной, но для определения она не подходит, ибо нельзя убедиться в том, что это число *не существует* (полный перебор физически не возможен!). Конструктивным вариантом определения неограниченной последовательности может служить, скажем, следующее: последовательность $\{x_n\}$ называется *неограниченной*, если

$$\forall C \geq 0 : \exists N_C : |x_{N_C}| > C \ ,$$

то есть, для каждого неотрицательного числа C найдется номер N_C такой, что будет выполнено неравенство $|x_{N_C}| > C$.

Числовые последовательности также можно различать по характеру изменения значений их членов при изменении номера. Например,

– последовательность, в которой изменение номера на 1 меняет знак ее члена на противоположный, называется *знакопеременной,*

– последовательность, для которой при любом n выполняется неравенство $x_{n+1} > x_n$ называется *монотонно возрастающей,* а в случае выполнения неравенства $x_{n+1} < x_n$ – *монотонно убывающей.*

Поясним данные определения следующими примерами.

1) Числовая последовательность $x_n = 1 - \dfrac{1}{n}$ является ограниченной, поскольку $\forall n : \quad 0 \leq x_n < 1$. Кроме того, она будет монотонно возрастающей в силу неравенства

$$x_{n+1} - x_n = \left(1 - \frac{1}{n+1}\right) - \left(1 - \frac{1}{n}\right) = -\frac{1}{n+1} + \frac{1}{n} = \frac{1}{n(n+1)} > 0 \quad \forall n.$$

2) Числовая последовательность $x_n = n^{(-1)^n}$, для которой

$$x_1 = 1, \ x_2 = 2, \ x_3 = \frac{1}{3}, \ x_4 = 4, \ x_5 = \frac{1}{5}, \ x_6 = 6, \ x_7 = \frac{1}{7}, \ \ldots,$$

ограничена снизу (числом ноль), не ограничена сверху и не является ни монотонно возрастающей, ни монотонно убывающей.

3.2.3. Предел числовой последовательности и его свойства

Как следует из определения числовой последовательности, для ее описания необходимо указать правило, которое позволяет находить значения членов последовательности по их номерам. Помимо этого числовая последовательность может характеризоваться также некоторым числом, не связанным ни с каким конкретным номером. Эта характеристика называется пределом числовой последовательности и определяется следующим образом.

Определение 3.2.3.1. Число A называется *пределом числовой последовательности* $\{x_n\}$, если *для любого* положительного числа ε *существует* номер N такой, что *для всех* членов последовательности с номерами $n \geq N$ выполнено неравенство $|x_n - A| < \varepsilon$.

Тот факт, что число A является пределом числовой последовательности $\{x_n\}$, символически записывается в виде $\lim\limits_{n\to\infty} x_n = A$. Иногда также используется обозначение $\{x_n\} \underset{n\to\infty}{\longrightarrow} A$. Вообще говоря, не всякая числовая последовательность имеет предел. Если числовая последовательность имеет предел, то она называется *сходящейся,* иначе – *расходящейся.*

Пример 3.2.3.1. Числовая последовательность $\left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right\}$, для которой $x_n = \frac{1}{n}$, имеет предел, равный нулю. То есть $\lim\limits_{n\to\infty} \frac{1}{n} = 0$.

Докажем это, использовав определение 3.2.3.1. Заметим во-первых, что данная числовая последовательность является монотонно убывающей, поскольку для любого номера n справедливо неравенство $\frac{1}{n} > \frac{1}{n+1}$, то есть $x_n > x_{n+1}$. С другой стороны,

$$|x_n - A| = \left|\frac{1}{n} - 0\right| = \frac{1}{n}.$$

Следовательно, для любого заданного положительного числа ε можно выбрать номер $N > \frac{1}{\varepsilon}$, для которого $\frac{1}{N} < \varepsilon$ и

$$|x_N - 0| = \left|\frac{1}{N} - 0\right| = \frac{1}{N} < \varepsilon.$$

Но тогда, в силу монотонного убывания рассматриваемой последовательности, *для всех* номеров $n \geq N$ также будет верным неравенство

$$|x_n - 0| = \frac{1}{n} < \varepsilon.$$

Значит, число 0 является пределом числовой последовательности $x_n = \frac{1}{n}$.

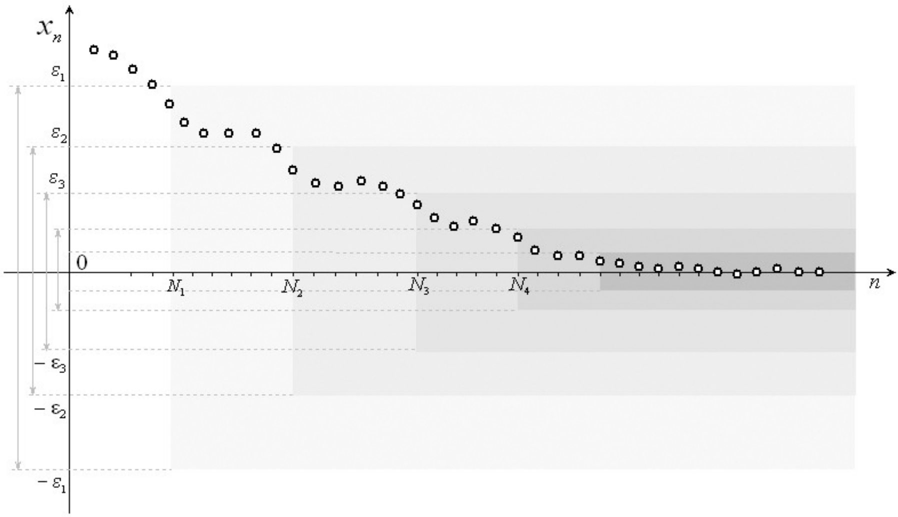


Рис. 3.2. График числовой последовательности с нулевым пределом.

Определение 3.2.3.1 можно интерпретировать как «игру», в которой один игрок задает *произвольное* (сколь угодное малое) положительное число ε, а его «противник» – второй игрок, по данному значению этого числа подбирает (или просто угадывает) номер N такой, что *для всех* членов последовательности с номерами $n \geq N$ их значения принадлежат интервалу $(A-\varepsilon, A+\varepsilon)$ (см. рис. 3.2, иллюстрирующий случай с $A = 0$) или, что то же самое, удовлетворяют условию $A-\varepsilon < x_n < A+\varepsilon$, которое равносильно неравенству $|x_n - A| < \varepsilon$. Если «победителем» в этой игре оказывается второй игрок, то число A называется пределом числовой последовательности $\{x_n\}$.

Стоит отметить, что правило выбора номера N (согласно определению 3.2.3.1) может быть различным при разных ε, что облегчает «победу второму игроку». С другой стороны, с точки зрения формальной логики тот факт, что «второму игроку» не удается по предложенному ему ε найти значение N с требуемым свойством, еще не означает, что предела у данной последовательности нет (может, этот «игрок» просто плохой). Определение *отсутствия предела* у последовательности $\{x_n\}$ должно формулироваться в виде условия «выигрыша первого игрока», например как

Определение 3.2.3.2. Число A *не является пределом числовой последовательности* $\{x_n\}$, если *существует* положительное число ε_0 такое, что *для любого* номера N *найдется* член данной последовательности с номером $n_0 \geq N$, для которого будет выполнено неравенство $|x_{n_0} - A| \geq \varepsilon_0$.

Последовательности, имеющие своим пределом число 0 (то есть для которых $\lim\limits_{n\to\infty} x_n = 0$), принято называть *бесконечно малыми.* Следует также различать случаи последовательностей *расходящихся*, то есть не имеющих никакого предела, и последовательностей *бесконечно больших*, члены которых (начиная с определенного номера) принимают значения по модулю большие, чем любое наперед заданное число, значения. Примером расходящейся может служить последовательность $x_n = (-1)^n$, то есть $\{-1,\ 1, -1,\ 1, -1,\ 1, \ldots\}$, а примером бесконечно большой – последовательность $x_n = n$. Для бесконечно больших последовательностей принято обозначение $\lim\limits_{n\to\infty} x_n = \infty$.

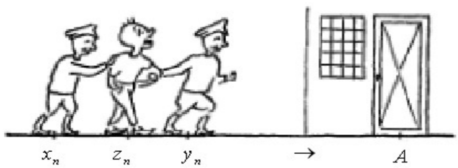


Рис. 3.3. Теорема «о двух милиционерах».

Отметим основные, полезные для решения практических задач, свойства пределов числовых последовательностей. Пусть все, использованные в записях 1°–6°, пределы существуют и C – некоторая константа, тогда

1°. $\lim\limits_{n\to\infty} (x_n + y_n) = \lim\limits_{n\to\infty} x_n + \lim\limits_{n\to\infty} y_n$.

2°. $\lim\limits_{n\to\infty} (C \cdot x_n) = C \cdot \lim\limits_{n\to\infty} x_n$.

3°. $\lim\limits_{n\to\infty} (x_n \cdot y_n) = \lim\limits_{n\to\infty} x_n \cdot \lim\limits_{n\to\infty} y_n$.

4°. Если, кроме того,

$$y_n \neq 0 \; \forall n \;\text{ и }\; \lim_{n\to\infty} y_n \neq 0, \text{ то } \lim_{n\to\infty} \frac{x_n}{y_n} = \frac{\lim\limits_{n\to\infty} x_n}{\lim\limits_{n\to\infty} y_n}\,.$$

5°. Если последовательность монотонно возрастает (убывает) и ограничена сверху (снизу), то она имеет предел.

6°. Если $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} y_n = A$ и $\forall n \quad x_n \geq z_n \geq y_n$, то $\lim\limits_{n\to\infty} z_n = A$.

Обратите внимание, что свойство 5°, являющееся *достаточным* условием существования предела, позволяет делать заключение о сходимости последовательности, *не находя значения ее предела.* Например, числовая последовательность, значения которой равны периметру правильного вписанного многоугольника, при неограниченном удвоении числа его сторон имеет предел, поскольку геометрически очевидно, что она монотонно возрастающая (в силу правила «треугольника») и ограничена сверху, скажем, периметром, описанного около той же окружности, квадрата. Значение предела этой последовательности принимается (по определению!) за *длину окружности,* которая обозначается как $2\pi R$.

В заключение стоит также заметить, что студенческий фольклёр называет свойство 6° теоремой «о двух милиционерах» (рис 3.3).

3.2.4. Нахождение пределов числовых последовательностей

Поиск значения предела числовых последовательностей, основанный лишь на его определении может оказаться весьма сложной вычислительной процедурой. На практике оказывается гораздо удобнее использовать *свойства пределов* последовательностей в сочетании с некоторым небольшим набором, найденных заранее, пределов.

В рамках настоящего курса окажется достаточным сочетание набора свойств 1°–6° предыдущего параграфа и трех следующих пределов:

$$\lim_{n\to\infty} \frac{1}{n} = 0\,; \qquad \lim_{n\to\infty} n \cdot \sin\frac{1}{n} = 1 \quad \text{и} \quad \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e\,.$$

Справедливость первого равенства была показана в примере 3.2.3.1. Рассмотрим второе равенство, часто называемое *первым замечательным пределом.*

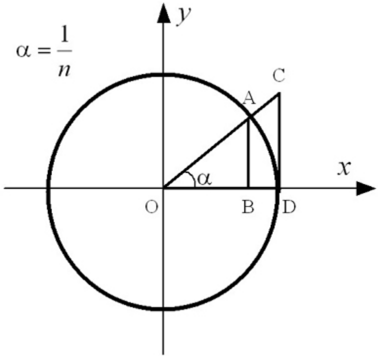


Рис. 3.4. К доказательству равенства $\lim\limits_{n\to\infty} n \cdot \sin \dfrac{1}{n} = 1$.

На тригонометрическом круге единичного радиуса отложим угол, величина которого (в радианной мере) равна $\alpha = \dfrac{1}{n}$ (рис. 3.3), и построим прямоугольные треугольники OAB и OCD. Заметим, что круговой сектор OAD, с одной стороны, содержит в себе треугольник OAB, а с другой – сам содержится в треугольнике OCD. Это означает, что для *площадей* этих трех фигур справедливы неравенства

$$S_{\triangle \mathrm{OAB}} \le S_{\cup \mathrm{OAD}} \le S_{\triangle \mathrm{OCD}} \,.$$

Поскольку $S_{\triangle \mathrm{OAB}} = \dfrac{1}{2}\cdot|\mathrm{OB}|\cdot|\mathrm{AB}|$, $\quad S_{\triangle \mathrm{OCD}} = \dfrac{1}{2}\cdot|\mathrm{OD}|\cdot|\mathrm{CD}|$, а площадь кругового сектора $S_{\cup \mathrm{OAD}} = \dfrac{1}{2}\cdot|\mathrm{OD}|\cdot\alpha$, то с учетом $|\mathrm{OD}| = 1$ приходим к неравенствам

$$\frac{1}{2}\cdot\sin\alpha\cdot\cos\alpha \quad \le \quad \frac{1}{2}\cdot 1\cdot\alpha \quad \le \quad \frac{1}{2}\cdot 1\cdot \operatorname{tg}\alpha \,.$$

или

$$\frac{1}{2}\cdot\sin\frac{1}{n}\cdot\cos\frac{1}{n} \quad \le \quad \frac{1}{2}\cdot 1\cdot\frac{1}{n} \quad \le \quad \frac{1}{2}\cdot 1\cdot \operatorname{tg}\frac{1}{n}$$

Далее преобразуя, получаем

$$\frac{1}{\sin\frac{1}{n}\cdot\cos\frac{1}{n}} \quad \ge \quad n \quad \ge \quad \frac{\cos\frac{1}{n}}{\sin\frac{1}{n}}.$$

Откуда, окончательно, следует, что

$$\frac{1}{\cos\dfrac{1}{n}} \quad \geq \quad n\cdot\sin\frac{1}{n} \quad \geq \quad \cos\frac{1}{n}.$$

Теперь можно воспользоваться свойствами пределов числовых последовательностей. Будем считать, что

$$x_n = \frac{1}{\cos\dfrac{1}{n}}\,; \quad z_n = n\cdot\sin\frac{1}{n}\,; \quad y_n = \cos\frac{1}{n}\,.$$

Тогда, в силу очевидного равенства $\lim\limits_{n\to\infty}\cos\dfrac{1}{n} = 1$ и теоремы «о двух милиционерах» – свойства 6°, получаем, что $\lim\limits_{n\to\infty} n\cdot\sin\dfrac{1}{n} = 1$.

К необходимости вычисления последнего из указанных выше пределов $\lim\limits_{n\to\infty}\left(1+\dfrac{1}{n}\right)^n$ приводит задача «о добром банке и жадном вкладчике», имеющая следующую формулировку.

Некий «добрый» банк предлагает своим вкладчикам 100% годовых по срочным вкладам, с равномерным во времени начислением процентов по вкладу. [1] У одного из его клиентов к началу года имеется денежная сумма размером в один рубль, которую он хочет положить в банк до начала следующего года. Очевидный расчет показывает, что вкладчик получит в конце года сумму в рублях: свой вклад 1 руб плюс 100%, то есть еще 1 руб. Итак,

$$S_1 = 1 + 1 = 2\,.$$

Однако вкладчика этот итог кажется недостаточным и он рассуждает так: «если я положу свой рубль на первое полугодие, то 30 июня у меня будет $\left(1+\dfrac{1}{2}\right)$ руб, которые я положу на оставшиеся полгода.» Тогда за год вкладчик будет иметь

$$S_2 = \left(1+\frac{1}{2}\right) + \frac{1}{2}\cdot\left(1+\frac{1}{2}\right) = \left(1+\frac{1}{2}\right)^2 = 2\frac{1}{4}.$$

[1] В жизни, конечно, никакой банк так никогда не делал и не делает.

Хотя эффект данной операции очевиден, «жадному» вкладчику и этого мало. Следующие его рассуждения таковы: «если я положу свой рубль на первые четыре месяца, то к 1 мая у меня будет на руках $\left(1+\dfrac{1}{3}\right)$ руб, которые я положу на следующие четыре месяца и получу 1 сентября

$$\left(1+\frac{1}{3}\right)+\frac{1}{3}\cdot\left(1+\frac{1}{3}\right)=\left(1+\frac{1}{3}\right)^{2}.$$

Эту сумму я вкладываю на оставшиеся четыре месяца и получаю в итоге

$$S_3=\left(1+\frac{1}{3}\right)^{2}+\frac{1}{3}\cdot\left(1+\frac{1}{3}\right)^{2}=\left(1+\frac{1}{3}\right)^{3}=2\frac{10}{27},$$

что больше, чем S_2.»

Нетрудно видеть, что, если весь год разделить на n равных периодов, то полученная сумма составит

$$S_n=\left(1+\frac{1}{n}\right)^{n-1}+\frac{1}{n}\cdot\left(1+\frac{1}{n}\right)^{n-1}=\left(1+\frac{1}{n}\right)^{n}.$$

Исследуем теперь свойства числовой последовательности $S_n = \left(1+\dfrac{1}{n}\right)^{n}$.

Во-первых, покажем, что $\forall n : S_{n+1} > S_n$. Действительно,

$$\frac{S_{n+1}}{S_n}=\frac{\left(1+\dfrac{1}{n+1}\right)^{n+1}}{\left(1+\dfrac{1}{n}\right)^{n}}=\frac{\left(1+\dfrac{1}{n+1}\right)^{n+1}}{\left(1+\dfrac{1}{n}\right)^{n+1}}\cdot\left(1+\frac{1}{n}\right)=$$

$$=\left[\frac{\dfrac{n+2}{n+1}}{\dfrac{n+1}{n}}\right]^{n+1}\cdot\left(1+\frac{1}{n}\right)=\left[\frac{n(n+2)}{(n+1)^2}\right]^{n+1}\cdot\left(1+\frac{1}{n}\right)=$$

$$=\left[\frac{n^2+2n}{(n+1)^2}\right]^{n+1}\cdot\left(1+\frac{1}{n}\right)=\left[\frac{n^2+2n+1-1}{(n+1)^2}\right]^{n+1}\cdot\left(1+\frac{1}{n}\right)=$$

и, согласно неравенству Бернулли (см. п.11° §1.1),

$$= \left[1 - \frac{1}{(n+1)^2}\right]^{n+1} \cdot \left(1 + \frac{1}{n}\right) > \left(1 - \frac{1}{n+1}\right) \cdot \left(1 + \frac{1}{n}\right) =$$

$$= \left(\frac{n}{n+1}\right) \cdot \left(1 + \frac{1}{n}\right) = 1.$$

Таким образом, $\dfrac{S_{n+1}}{S_n} > 1$ и последовательность S_n – *монотонно возрастающая*, то есть «шустрость» вкладчика оправдана.

Теперь убедимся, что, сколь угодно большой суммы вкладчик получить тем не менее не сможет. Выполним следующую оценку воспользовавшись формулой бинома Ньютона (см. п.8° §1.1),

$$S_n = \left(1 + \frac{1}{n}\right)^n =$$

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

$$\leq 1 + 1 + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \ldots \leq 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + + \frac{1}{2^3} + \ldots \leq$$

и, по формуле суммы всех членов бесконечно убывающей геометрической прогрессии (см. п.9° §1.1), получаем

$$\leq 1 + \frac{1}{1 - \frac{1}{2}} = 3.$$

Это означает, что последовательность $\{S_n\}$ *ограничена сверху* и как бы вкладчик не суетился, даже трех рублей ему получить не удастся.

Согласно свойству 5°, числовая последовательность монотонно возрастающая и ограниченная сверху имеет предел. Значит, $\{S_n\}$ сходится. Пределом числовой последовательности $\left\{\left(1 + \frac{1}{n}\right)^n\right\}$ является иррациональное (подобно π или $\sqrt{2}$) число, обозначаемое как e и равное приближенно $e \approx 2.718281828459045 \cdots$. Иначе говоря,

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Это равенство принято называть *вторым замечательным пределом.*

Заметим, что непосредственное применение свойства 3° при вычислении, например, первого замечательного предела невозможно, поскольку из двух последовательностей $x_n = n$ и $y_n = \sin\frac{1}{n}$ сходится только вторая. Ее предел равен 0, в то время как первая неограничено возрастает. Подобный случай принято называть неопределенностью вида «$0 \cdot \infty$» и для нахождения предела потребовалось специальное исследование. Аналогичные проблемы возникают для неопределенностей типа «$\frac{0}{0}$» , «$\frac{\infty}{\infty}$» , «$\infty - \infty$» , «1^∞» .

Второй замечательный предел является примером неопределенности последнего типа.

Преобразования формульной записи общего члена числовой последовательности в тех случаях, когда непосредственное использование свойств числовых последовательностей 1°–6° невозможно, принято называть методом «раскрытия неопределенностей.»

Рассмотрим следующие задачи.

Пример 3.2.4.1. Найти $\lim\limits_{n\to\infty} \dfrac{(3n+2)^2}{5n^2+3}$.

Формула значения члена последовательности является дробью, однако использование свойства 4° невозможно, поскольку

$$\lim_{n\to\infty}(3n+2)^2 = +\infty \quad \text{и} \lim_{n\to\infty}\left(5n^2+3\right) = +\infty\ .$$

То есть мы имеем случай неопределенности вида «$\frac{\infty}{\infty}$» . Для ее «раскрытия» (до перехода к пределу!) преобразуем числитель по формуле «квадрат суммы двух чисел», а затем разделим как числитель, так и знаменатель на n^2 и в итоге получим

$$\lim_{n\to\infty}\frac{(3n+2)^2}{5n^2+3} = \lim_{n\to\infty}\frac{9n^2+12n+4}{5n^2+3} = \lim_{n\to\infty}\frac{9+\dfrac{12}{n}+\dfrac{4}{n^2}}{5+\dfrac{3}{n^2}}.$$

Теперь, в силу $\lim\limits_{n\to\infty}\dfrac{1}{n} = 0$ и $\lim\limits_{n\to\infty}\dfrac{1}{n^2} = \lim\limits_{n\to\infty}\dfrac{1}{n}\cdot\lim\limits_{n\to\infty}\dfrac{1}{n} = 0$, очевидно, что пределы числителя и знаменателя существуют, и можно восползоваться свойствами 4°, 2° и 1° .

$$\lim_{n\to\infty}\frac{9+\dfrac{12}{n}+\dfrac{4}{n^2}}{5+\dfrac{3}{n^2}}=\frac{\lim\limits_{n\to\infty}\left(9+\dfrac{12}{n}+\dfrac{4}{n^2}\right)}{\lim\limits_{n\to\infty}\left(5+\dfrac{3}{n^2}\right)}=\frac{9+12\lim\limits_{n\to\infty}\dfrac{1}{n}+4\lim\limits_{n\to\infty}\dfrac{1}{n^2}}{5+3\lim\limits_{n\to\infty}\dfrac{1}{n^2}}=\frac{9}{5}.$$

Пример 3.2.4.2. Найти $\lim\limits_{n\to\infty}\left(\sqrt{4n^2+3n}-2n\right)$.

Формула общего члена последовательности $a_n = \sqrt{4n^2+3n}-2n$ представляет собой разность двух выражений, однако использовать свойство 1° мы не можем, поскольку

$$\lim_{n\to\infty}\sqrt{4n^2+3n}=+\infty \quad \text{и} \quad \lim_{n\to\infty}2n=+\infty$$

и мы имеем дело с неопределенностью вида «$\infty-\infty$». Для ее «раскрытия», не переходя пока еще к пределу, умножим и одновременно разделим эту разность на сумму $\sqrt{4n^2+3}+2n$, с последующим использованием формулы $(a-b)(a+b)=a^2-b^2$

$$\lim_{n\to\infty}\left(\sqrt{4n^2+3n}-2n\right)=\lim_{n\to\infty}\frac{(\sqrt{4n^2+3n}-2n)(\sqrt{4n^2+3n}+2n)}{\sqrt{4n^2+3n}+2n}=$$

$$=\lim_{n\to\infty}\frac{(4n^2+3n)-4n^2}{\sqrt{4n^2+3n}+2n}=\lim_{n\to\infty}\frac{3n}{\sqrt{4n^2+3n}+2n}.$$

Полученный предел есть неопределенность вида «$\dfrac{\infty}{\infty}$», «раскрыть» которую можно делением числителя и знаменателя на n.

$$\lim_{n\to\infty}\frac{3n}{\sqrt{4n^2+3n}+2n}=3\cdot\lim_{n\to\infty}\frac{1}{\sqrt{4+\dfrac{3}{n}}+2}=\frac{3}{4},$$

поскольку $\lim\limits_{n\to\infty}\sqrt{4+\dfrac{3}{n}}=2.$

Действительно, с одной стороны, $\sqrt{4+\dfrac{3}{n}}\geq 2$, но, с другой

$$\sqrt{4+\frac{3}{n}}=\sqrt{4+2\cdot2\cdot\frac{3}{4n}+\frac{9}{16n^2}-\frac{9}{16n^2}}=$$

$$= \sqrt{\left(2+\frac{3}{4n}\right)^2 - \frac{9}{16n^2}} \le \sqrt{\left(2+\frac{3}{4n}\right)^2} = 2+\frac{3}{4n}.$$

То есть,

$$2 \le \sqrt{4+\frac{3}{n}} \le 2+\frac{3}{4n}\,,$$

и на основании теоремы «о двух милиционерах», приходим к заключению о том, что $\lim\limits_{n\to\infty}\sqrt{4+\dfrac{3}{n}} = 2$.

Пример 3.2.4.3. Найти $\lim\limits_{n\to\infty}\left(\dfrac{n+2}{n+1}\right)^{3n}$.

Здесь мы имеем $\lim\limits_{n\to\infty}\dfrac{n+2}{n+1} = 1$ и $\lim\limits_{n\to\infty} 3n = \infty$, то есть неопределенность вида «1^∞». Чтобы «раскрыть» ее, преобразуем выражение под знаком предела следующим образом.

$$\lim_{n\to\infty}\left(\frac{n+2}{n+1}\right)^{3n} = \lim_{n\to\infty}\left(1+\frac{1}{n+1}\right)^{3n} = \lim_{k\to\infty}\left[\left(1+\frac{1}{k}\right)^{k+1}\right]^3 =$$

где $k = n+1$ и $n = k-1 \Rightarrow \lim\limits_{n\to\infty} k = \infty$

$$= \left[\lim_{k\to\infty}\left(1+\frac{1}{k}\right)^{k}\right]^3 \cdot \lim_{k\to\infty}\left(1+\frac{1}{k}\right)^3 = e^3\cdot 1 = e^3.$$

3.2.5. Последовательности нечисловых объектов

В завершение обсуждения темы «числовые последовательности» отметим, что последовательности можно образовывать не только из чисел, но и из математических объектов более сложного вида, скажем, функций или матриц. Поясним это следующими примерами.

Пример 3.2.5.1. Рассмотрим множество функций вида $y_n(x) = \dfrac{1}{x^2+n^2}$, где n – произвольное натуральное число.

У всех этих функций область определения одинаковая X_n : $\{(-\infty,+\infty)\}$ – множество всех вещественных чисел, в то время как область значений Y_n : $\left\{\left(0,\dfrac{1}{n^2}\right]\right\}$ – различная для разных номеров n.

Для каждого номера n функция указанного вида существует и единственна, поэтому для описания совокупности всех таких функций можно использовать термин *функциональная последовательность*, который можно рассматривать как обобщение понятия числовой последовательности поскольку для каждого фиксированного $x \in X_n$ формула $y_n(x) = \dfrac{1}{x^2+n^2}$ задает именно числовую последовательность.

Легко видеть, что $\lim\limits_{n\to\infty}\dfrac{1}{x^2+n^2} = 0$ при любых $x \in X_n$. Поэтому естественно функцию $y^*(x)$, равную 0 для всех x назвать *предельной функцией* для функциональной последовательности $y_n(x)$ и записывать этот факт в виде $\lim\limits_{n\to\infty} y_n(x) = y^*(x)$.

Подобная ситуация имеет место и для матриц. Действительно, раз элементами матриц являются числа, то матричные последовательности можно строить, заменяя эти элементы членами числовых последовательностей, имеющими одинаковые номера.

Пример 3.2.5.2. Пусть элементами квадратной матрицы второго порядка $\|A_n\| = \begin{Vmatrix} a_{11,n} & a_{12,n} \\ a_{21,n} & a_{22,n} \end{Vmatrix}$ являются члены числовых последовательностей

$$a_{11,n} = \frac{n}{n+1}\,, \quad a_{12,n} = \frac{1}{n+2}\,, \quad a_{21,n} = \frac{1}{n+1}\,, \quad a_{22,n} = \frac{n}{n+2}\,,$$

тогда присваивая номерам значения $1, 2, 3, \ldots$ получим последовательность состоящую из матриц

$$\left\{ \begin{Vmatrix} \frac{1}{2} & \frac{1}{3} \\[2ex] \frac{1}{2} & \frac{1}{3} \end{Vmatrix}, \begin{Vmatrix} \frac{2}{3} & \frac{1}{4} \\[2ex] \frac{1}{3} & \frac{1}{2} \end{Vmatrix}, \begin{Vmatrix} \frac{3}{4} & \frac{1}{5} \\[2ex] \frac{1}{4} & \frac{3}{5} \end{Vmatrix}, \begin{Vmatrix} \frac{4}{5} & \frac{1}{6} \\[2ex] \frac{1}{5} & \frac{2}{3} \end{Vmatrix}, \ldots \right\}$$

Поскольку последовательности $\{a_{11,n}\}, \{a_{12,n}\}, \{a_{21,n}\}, \{a_{22,n}\}$ сходящиеся и

$$\lim_{n\to\infty}\frac{n}{n+1} = 1\,, \quad \lim_{n\to\infty}\frac{1}{n+2} = 0\,, \quad \lim_{n\to\infty}\frac{1}{n+1} = 0\,, \quad \lim_{n\to\infty}\frac{n}{n+2} = 1\,,$$

то можно сказать, что и матричная последовательность $\|A_n\|$ сходится и притом к единичной матрице второго порядка, то есть

$$\lim_{n\to\infty} \|A_n\| = \|E\| = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}.$$

Пример 3.2.5.3. Еще раз вспомним «задачу о жуках».
Пусть столбец $\|X_n\| = \begin{Vmatrix} x_{1,n} \\ x_{2,n} \\ x_{3,n} \end{Vmatrix}$ описывает распределение жуков по средам обитания во время n-го наблюдения. Тогда распределение при следующим, $n+1$-ом наблюдении, согласно соотношению (2.3.2) будет определяться формулой

$$\|X_{n+1}\| = \|A\| \cdot \|X_n\|,$$

или в развернутом виде

$$\begin{Vmatrix} x_{1,n+1} \\ x_{2,n+1} \\ x_{3,n+1} \end{Vmatrix} = \begin{Vmatrix} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{Vmatrix} \cdot \begin{Vmatrix} x_{1,n} \\ x_{2,n} \\ x_{3,n} \end{Vmatrix}.$$

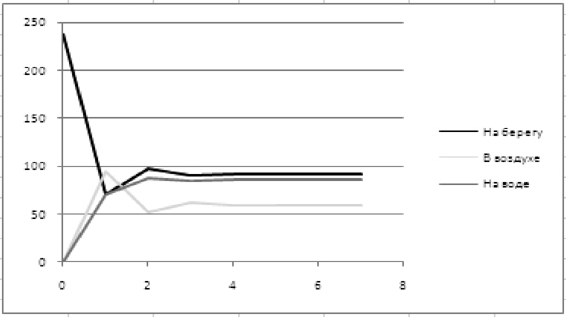
Нетрудно видеть, что в рассматриваемой задаче возникает матричная последовательность $\{X_n\}$, заданная рекуррентно, каждый член которой является трехкомпонентным столбцом, описывающем распределение числа жуков по средам обитания во время n-го наблюдения.

Получение аналитической формы записи членов этой последовательности является весьма сложной математической задачей и выходит за рамки нашего курса. Однако, представление о характере поведения ее членов с ростом n можно получить, проведя численные расчеты и построив графические диаграммы.

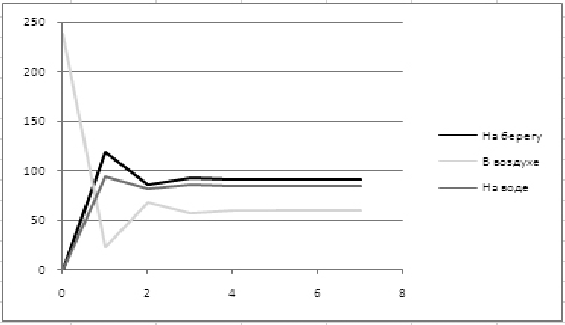
Предполагая, что полная численность колонии постоянна и равна 238 особям, найдем несколько начальных членов этой последовательности для различных исходных вариантов распределения жуков по средам. Рассмотрим случаи «в начале все жуки были на берегу», затем - «в начале все жуки были в воздухе», «в начале все жуки были на воде» и, наконец, случай, когда число жуков в различных средах в начале было примерно одинаковым. Соответствующие столбцы этих начальных распределений имеют вид:

$$\begin{Vmatrix} 238 \\ 0 \\ 0 \end{Vmatrix} \quad \begin{Vmatrix} 0 \\ 238 \\ 0 \end{Vmatrix} \quad \begin{Vmatrix} 0 \\ 0 \\ 238 \end{Vmatrix} \quad \begin{Vmatrix} 79 \\ 79 \\ 80 \end{Vmatrix}.$$

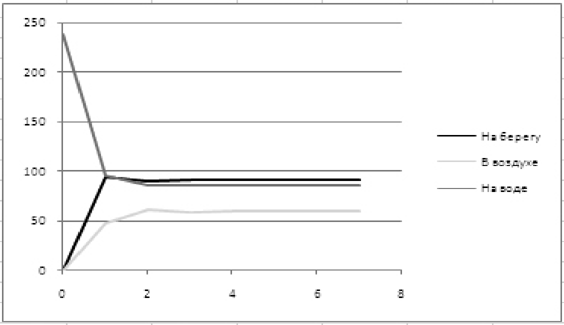
№ наблюд.	На берегу	В воздухе	На воде
0	238	0	0
1	71	95	71
2	98	52	88
3	91	62	85
4	92	60	86
5	92	60	86
6	92	60	86
7	92	60	86



№ наблюд.	На берегу	В воздухе	На воде
0	0	238	0
1	119	24	95
2	86	69	83
3	94	58	87
4	92	60	86
5	92	60	86
6	92	60	86
7	92	60	86



№ наблюд.	На берегу	В воздухе	На воде
0	0	0	238
1	95	48	95
2	90	62	86
3	92	60	86
4	92	60	86
5	92	60	86
6	92	60	86
7	92	60	86



№ наблюд.	На берегу	В воздухе	На воде
0	79	79	80
1	95	56	87
2	91	61	86
3	92	60	86
4	92	60	86
5	92	60	86
6	92	60	86
7	92	60	86

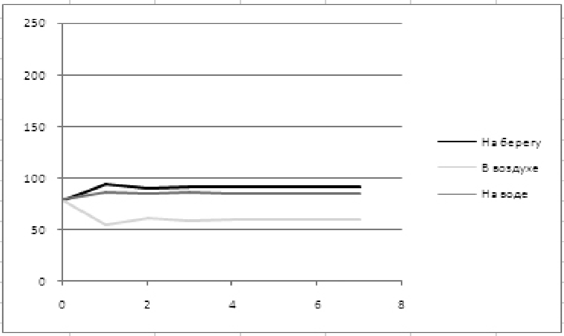


Рис. 3.5. Изменение распределения числа жуков по средам обитания

Результаты расчетов и графики изменения распределений числа жуков по средам обитания показаны на рис.3.4.

Представляется интересным тот факт, что последовательности оказались сходящимися (и довольно быстро) к некоторому «равновесному» распределению, которое совпадает с *собственным вектором* матрицы $\|A\|$, отвечающему ее *собственному значению* $\lambda = 1$, и имеющему вид (см. §2.4.2)

$$\left\| \begin{array}{c} 92 \\ 60 \\ 86 \end{array} \right\| .$$

3.3. Предел и непрерывность функции

3.3.1. Предел функции и его свойства

Рассматривая значения некоторой функции $f(x)$ в малой окрестности точки $x = a$, в большом числе случаев можно заметить, что эти значения оказываются тем ближе к некоторому числу A, чем меньше значение x отличается от a. При этом величина A может отличаться от значения $f(a)$ и даже существовать в тех случаях, когда точка a не принадлежит области определения функции $f(x)$.

Примером может служить функция $f(x) = \dfrac{\sin x}{x}$, которая не имеет никакого значения в точке $x = 0$, но определена в любой ее окрестности и значение которой тем меньше отличается от единицы, чем меньше абсолютная величина x.

Если такое число A существует, его называют *пределом функции $f(x)$ при x стремящемся к a*, и данный факт символически обозначают как $\lim\limits_{x\to a} f(x) = A$.

Приведем теперь строгое определение понятия предела функции.

Определение 3.3.1.1. Число A называется *пределом функции $f(x)$ при x стремящемся к a*, если *для любой* числовой последовательности $\{x_n\}$ такой, что $\lim\limits_{n\to\infty} x_n = a$ и $x_n \neq a$, числовая последовательность $\{f(x_n)\}$

сходится к A, то есть выполняется равенство $\lim\limits_{n\to\infty} f(x_n) = A$.

В символической форме это определение можно записать так: $\lim\limits_{x\to a} f(x) = A$, если $\forall\{x_n\} \underset{n\to\infty}{\longrightarrow} a \ : \ \{f(x_n)\} \underset{n\to\infty}{\longrightarrow} A$.

Отметим, что определение 3.3.1.1 применимо, когда a обозначает или некоторое конечное число, или является одним из следующих трех символов ∞ , $+\infty$ или $-\infty$.

Подчеркнем еще раз: предел функции при x стремящемся к a (так же как и $f(a)$ – ее значение) является *локальной* числовой характеристикой функции, то есть относящейся к точке a. Для одной и той же точки значение функции и значение её предела, равно как и их существование, независимы друг от друга: они могут существовать одновременно и быть равными или не равными друг другу, и также могут не существовать, как вместе, так и по отдельности. Поясним это следующими примерами.

Пример 3.3.1.1. Найти, используя определение 3.3.1.1, $\lim\limits_{x\to 3} \dfrac{2x}{x^2+1}$.

Возьмем произвольную числовую последовательность $\{x_n\}$ такую, что $\lim\limits_{n\to\infty} x_n = 3$, тогда n-й член последовательности $\{f(x_n)\}$ будет иметь вид $\dfrac{2x_n}{x_n^2+1}$. Найдем ее предел, пользуясь известными нам свойствами пределов числовых последовательностей.

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \frac{2x_n}{x_n^2+1} = \frac{\lim\limits_{n\to\infty} 2x_n}{\lim\limits_{n\to\infty} (x_n)^2+1} =$$

$$\frac{2\cdot \lim\limits_{n\to\infty} x_n}{(\lim\limits_{n\to\infty} x_n)\cdot(\lim\limits_{n\to\infty} x_n)+1} = \frac{2\cdot 3}{3\cdot 3+1} = \frac{3}{5}\,.$$

Нетрудно заметить, что функция $f(x) = \dfrac{2x}{x^2+1}$ определена (имеет значение) для любого конечного x, в том числе и для $x=3$. В рассматриваемом случае $f(3) = \dfrac{2\cdot 3}{3^2+1} = \dfrac{3}{5}$, то есть значение функции и ее предел в точке $x=3$ совпадают.

Однако, если рассмотреть поведение этой же функции при x стремящимся к ∞, то мы получим иной случай. С одной стороны, $f(x) = \dfrac{2x}{x^2+1}$ не определена, то есть не имеет никакого значения при $x=\infty$.

Но, с другой стороны, для $\lim\limits_{n\to\infty} x_n = \infty$,

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} \frac{2x_n}{x_n^2+1} = \lim_{n\to\infty} \frac{1}{x_n} \cdot \frac{2}{1+\dfrac{1}{x_n^2}} =$$

$$= \lim_{n\to\infty} \frac{1}{x_n} \cdot \lim_{n\to\infty} \frac{2}{1+\dfrac{1}{x_n^2}} = 0 \cdot 2 = 0 \; .$$

Итак, эта функция при $x \longrightarrow \infty$ имеет предел, равный 0, то есть $f(x) \longrightarrow 0$, но не имеет никакого значения при $x = \infty$.

Пример 3.3.1.2. Рассмотрим функцию, называемую *сигнатурой числа,* обозначаемую как $y = \operatorname{sgn} x$ и определяемую формулой (см. рис.3.5)

$$\operatorname{sgn} x = \left[\begin{array}{rl} 1, & \text{если } x > 0, \\ 0, & \text{если } x = 0, \\ -1, & \text{если } x < 0. \end{array} \right.$$

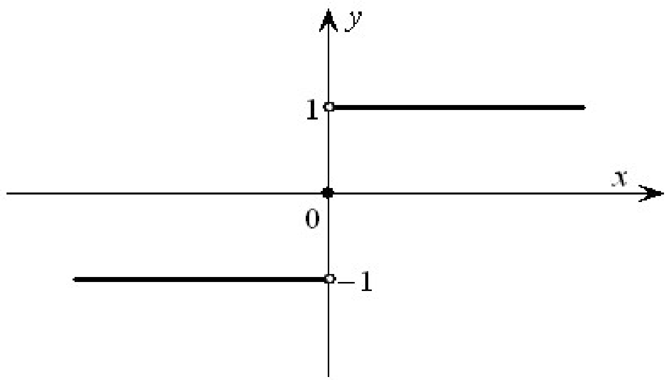


Рис. 3.6. График функции $y = \operatorname{sgn} x$

Эта функция при $x = 0$ имеет нулевое значение, то есть $\operatorname{sgn} 0 = 0$, однако предел $\lim\limits_{x\to 0} \operatorname{sgn} x$ не существует. Действительно, возьмем две числовые последовательности с $x_n = \dfrac{1}{n}$ и $y_n = -\dfrac{1}{n}$. Они обе имеют своим пределом число 0. Но, по определению сигнатуры, $\lim\limits_{n\to\infty} \operatorname{sgn} x_n =$

$\lim\limits_{n\to\infty} \operatorname{sgn}\left(\dfrac{1}{n}\right) = 1$ и $\lim\limits_{n\to\infty} \operatorname{sgn} y_n = \lim\limits_{n\to\infty} \operatorname{sgn}\left(-\dfrac{1}{n}\right) = -1$, а это противоречит определению 3.3.1.1, поскольку значение предела $\operatorname{sgn} x_n$ должно быть *одинаковым для всех* последовательностей $\{x_n\} \underset{n\to\infty}{\longrightarrow} 0$.

Наконец, возможен случай, когда у функции в некоторой точке нет значения и не существует предел. Проверьте самостоятельно, что такая ситуация имеет место для $f(x) = \dfrac{1}{x}$ при $x = 0$.

Свойства пределов функций аналогичны свойствам пределов последовательностей. Приведем основные из них, предполагая, что, используемые в формулировках, пределы существуют.

1°. $\lim\limits_{x\to a} (f(x) \pm g(x)) = \lim\limits_{x\to a} f(x) \pm \lim\limits_{x\to a} g(x)$.

2°. $\lim\limits_{x\to a} (C \cdot f(x)) = C \cdot \lim\limits_{x\to a} f(x)$.

3°. $\lim\limits_{x\to a} (f(x) \cdot g(x)) = \lim\limits_{x\to a} f(x) \cdot \lim\limits_{x\to a} g(x)$.

4°. Если, кроме того $\lim\limits_{x\to a} g(x) \neq 0$, то $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = \dfrac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)}$.

5°. Аналог теоремы «о двух милиционерах»: если для всех x, принадлежащих некоторой окрестности точки a, верно $f(x) \geq h(x) \geq g(x)$ и $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = A$, то $\lim\limits_{x\to a} h(x) = A$.

Наконец, для решения задач полезными оказываются так называемые «замечательные пределы» функций:

1) *Первый* замечательный предел: $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$.

2) *Второй* замечательный предел:

$$\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x = e \quad \text{или} \quad \lim_{x\to 0} (1 + x)^{\frac{1}{x}} = e \ .$$

3.3.2. Нахождение пределов функций. Раскрытие неопределенностей

Как и для задачи нахождения предела числовой последовательности, при поиске пределов функций сочетание использования свойств пределов и набора «замечательных пределов« позволяет в ряде случаев выполнять «раскрытие неопределенностей«, основные из которых:

$$«0\cdot\infty» \quad , \qquad «\frac{0}{0}» \quad , \qquad «\frac{\infty}{\infty}» \quad , \qquad «\infty-\infty» \quad , \qquad «1^{\infty}» \ .$$

Продемонстрируем соответствующие приемы на следующих примерах.

Пример 3.3.2.1. Найти $\lim\limits_{x\to 2}\dfrac{x^2-3x+2}{x^3-8}$.

В этой задаче необходимо раскрыть неопределенность вида $«\frac{0}{0}»$.

$$\lim_{x\to 2}\frac{x^2-3x+2}{x^3-8}=\lim_{x\to 2}\frac{(x-1)(x-2)}{(x-2)(x^2+2x+4)}=\lim_{x\to 2}\frac{x-1}{x^2+2x+4}=\frac{1}{12}\,.$$

Пример 3.3.2.2. Найти $\lim\limits_{x\to\infty}\dfrac{3x^2+2}{(2x+1)^2}$.

Здесь имеет место неопределенность вида $«\frac{\infty}{\infty}»$.

$$\lim_{x\to\infty}\frac{3x^2+2}{(2x+1)^2}=\lim_{x\to\infty}\frac{3x^2+2}{4x^2+4x+1}=\lim_{x\to\infty}\frac{3+\dfrac{2}{x^2}}{4+\dfrac{4}{x}+\dfrac{1}{x^2}}=\frac{3}{4}\,.$$

Пример 3.3.2.3. Найти $\lim\limits_{x\to+\infty}\left(\sqrt{x^2+3x}-x\right)$.

Тип неопределенности в этом примере – $«\infty-\infty»$.

$$\lim_{x\to+\infty}\left(\sqrt{x^2+3x}-x\right)=\lim_{x\to+\infty}\frac{\sqrt{x^2+3x}-x}{1}=$$

$$=\lim_{x\to+\infty}\frac{(\sqrt{x^2+3x}-x)(\sqrt{x^2+3x}+x)}{(\sqrt{x^2+3x}+x)}=$$

$$=\lim_{x\to+\infty}\frac{x^2+3x-x^2}{(\sqrt{x^2+3x}+x)}=\lim_{x\to+\infty}\frac{3x}{(\sqrt{x^2+3x}+x)}=$$

теперь мы имеем дело с неопределенностью вида $«\frac{\infty}{\infty}»$ – разделим числитель и знаменатель на x :

$$=\lim_{x\to+\infty}\frac{3}{\sqrt{1+\dfrac{3}{x}}+1}=\frac{3}{2}\,.$$

В ряде случаев для использования значений «замечательных» пределов оказывается целесообразным выполнить замену переменной величины.

Пример 3.3.2.4. Найти $\lim\limits_{x \to 0} \dfrac{\sin 5x}{x}$.

Для раскрытия неопределенности вида «$\dfrac{0}{0}$» в этой задаче удобно ввести новую переменную $t = 5x$, которая будет очевидно стремиться к нулю, когда x стремится к нулю. Поэтому, в силу первого «замечательного» предела

$$\lim_{x \to 0} \frac{\sin 5x}{x} = \lim_{t \to 0} \frac{\sin t}{\left(\dfrac{t}{5}\right)} = \lim_{t \to 0} 5 \cdot \frac{\sin t}{t} = 5 \cdot \lim_{t \to 0} \frac{\sin t}{t} = 5 \cdot 1 = 5\ .$$

Пример 3.3.2.5. Найти $\lim\limits_{x \to \infty} \left(1 - \dfrac{3}{x}\right)^x$.

Данная задача приводит к необходимости раскрытия неопределенности типа «1^∞». Выполним замену переменной, положив $t = -\dfrac{x}{3} \Rightarrow x = -3t$. Подставляя, получаем

$$\lim_{x \to \infty} \left(1 - \frac{3}{x}\right)^x = \lim_{t \to \infty} \left(1 + \frac{1}{t}\right)^{-3t} = \lim_{t \to \infty} \left(\left(1 + \frac{1}{t}\right)^t\right)^{-3} =$$

$$= \left(\lim_{t \to \infty} \left(1 + \frac{1}{t}\right)^t\right)^{-3} = e^{-3} = \frac{1}{e^3}.$$

3.3.3. Непрерывность функции в точке и на множестве. Классификация точек разрыва

Определение 3.3.3.1. Функция $f(x)$ называется *непрерывной в точке* $x = a$, если

$$\lim_{x \to a} f(x) = f(a).$$

Если это условие не выполнено, то говорят, что функция $f(x)$ имеет *разрыв в точке* $x = a$.

Если функция $f(x)$ непрерывна в точке $x = a$, то для любой числовой последовательности $x_n \longrightarrow a$ верно равенство

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(a).$$

а, значит, если $f(x)$ непрерывна и в точке $x = g(a)$, где $g(x)$ другая непрерывная в точке $x = a$ функция, то

$$\lim_{x\to a} f(g(x)) = f(\lim_{x\to a} g(x)) = f(g(a)), \tag{3.3.3.1}$$

Определение 3.3.3.2. Функция $f(x)$ называется *непрерывной на некотором числовом множестве*, если она непрерывна в *каждой* точке этого множества. Функция $f(x)$ называется *разрывной на некотором числовом множестве*, если она не не является непрерывной *хотя бы в одной* из точек этого множества.

Пример 3.3.3.1. Функция $f(x) = \sin x$ непрерывна на $X\ :\ (-\infty, +\infty)$, а функция $f(x) = \dfrac{1}{x}$ непрерывна как на $X\ :\ (-\infty, 0)$, так и на $X\ :\ (0, +\infty)$, но разрывна на $X\ :\ ((-\infty, +\infty))$.

Определение 3.3.3.3. Говорят, что функция $f(x)$ имеет *устранимый разрыв* в точке $x = a$, если $\lim\limits_{x\to a} f(x)$ существует, но $\lim\limits_{x\to a} f(x) \neq f(a)$. Если же $\lim\limits_{x\to a} f(x)$ не существует, когда точка $x = a$ принадлежит области определения функции $f(x)$, то точка $x = a$ называется *неустранимой точкой разрыва* функции $f(x)$.

Пример 3.3.3.2. Исследовать на непрерывность и выполнить классификацию ее точек разрыва функций:

1) $y = \operatorname{sgn} x$.

   У данной функции точка разрыва $x = 0$ неустранимая, так как предел в этой точке не существует (см. рис. 3.5).

2) $y = |\operatorname{sgn} x|$.

   Для этой функции точка разрыва $x = 0$ устранимая, так как предел в этой точке существует, $\lim\limits_{x\to 0} |\operatorname{sgn} x| = 1$, но не равен значению функции: $|\operatorname{sgn} 0| = 0 \neq 1$.

3)

$$y = \left[\begin{array}{ll} \dfrac{\sin x}{x}, & \text{если } x \neq 0, \\ a \quad , & \text{если } x = 0. \end{array} \right.$$

Данная функция непрерывна $\forall x$, если $a = 1$, поскольку $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$ и имеет в $x = 0$ устранимую точку разрыва $\forall a \neq 1$.

4)

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

Эта функция имеет в $x = 1$ устранимую точку разрыва, а в $x = 2$ – неустранимую. Действительно, если преобразовать запись данной функции к виду $y = \dfrac{\sin(x-1)}{x-1} \cdot \dfrac{1}{x-2}$, то для $x = 1$, в силу «первого замечательного предела«, существует предел

$$\lim_{x\to 1} \frac{\sin(x-1)}{x-1} \cdot \frac{1}{x-2} = \lim_{x\to 1} \frac{\sin(x-1)}{x-1} \cdot \lim_{x\to 1} \frac{1}{x-2} = 1 \cdot \frac{1}{-1} = -1,$$

в то время как в точке $x = 2$, хотя $\lim\limits_{x\to 2} \dfrac{\sin(x-1)}{x-1} = \sin 1$, но $\lim\limits_{x\to 2} \dfrac{1}{x-2}$ – не существует, и мы имеем разрыв неустранимого типа.

Использование свойства непрерывности во многих случаях позволяет упростить процедуру раскрытия неопределенностей при нахождении пределов функций.

Пример 3.3.3.3. Найти $\lim\limits_{x\to 0} \dfrac{\ln(1+x)}{x}$.

Данная задача приводит к неопределенности типа «$\dfrac{0}{0}$». Для ее раскрытия выполним, приняв во внимание непрерывность логарифмической функции (см. формулу (3.3.3.1)) и «второй замечательный предел», следующие преобразования

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} = \lim_{x\to 0} \left(\ln(1+x)^{\frac{1}{x}} \right) = \ln \left(\lim_{x\to 0} (1+x)^{\frac{1}{x}} \right) = \ln e = 1 \ .$$

Глава 4.

Элементы аналитической геометрии

Помимо табличного и формульного способов описания функциональных зависимостей на практике достаточно часто используется также и графический метод их представления. Основой этого подхода к представлению функций и исследованию их свойств являются системы координат и координатные описания геометрических объектов, знакомству с которыми и посвящена эта глава.

4.1. Векторы и действия с ними

Определение 4.1.1. *Направленным отрезком* $\overline{AB}$ называется отрезок прямой, для которого указаны начало (точка A) и конец (точка B). Длиной направленного отрезка $\overline{AB}$ называется расстояние между точками A и B, обозначаемое $\left|\overline{AB}\right|$. Направленный отрезок, у которого начало и конец совпадают, называется *нулевым* .

Определение 4.1.2. **Операция сравнения** Два направленных отрезка называются *равными*, если их соответствующие концевые точки можно совместить параллельным переносом любого из сравниваемых отрезков.

Определение 4.1.3. **Операция сложения** *Суммой* направленных отрезков $\overline{AB}$ и $\overline{CD}$ называется направленный отрезок $\overline{AD}$, получаемый после совмещения точек B и C параллельным переносом любого из слагаемых.

Данное определение иногда называют *правилом треугольника*, или *правилом параллелограмма*, поскольку сложение направленных отрезков также можно выполнять, совместив их начала и построив параллелограмм, сторонами которого данные отрезки являются. Суммой при этом (в соответствии с определением 4.1.3) оказывается направленный отрезок $\overline{AC}$, совпадающий с диагональю построенного параллелограмма (см. рис 4.1.)

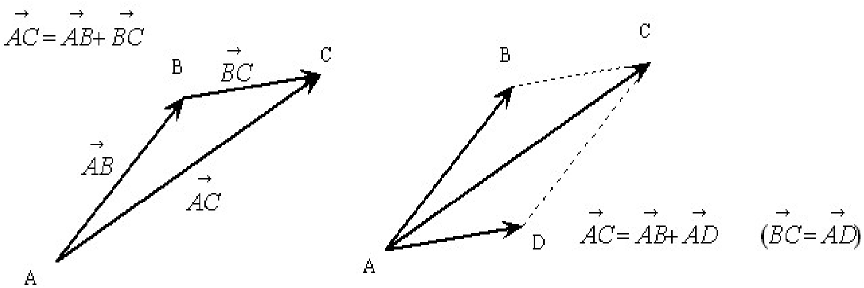


Рис. 4.1. Операция сложения направленных отрезков

Определение 4.1.4. **Операция умножения числа на направленный отрезок** *Произведением числа* λ на направленный отрезок $\overline{AB}$ называется параллельный ему направленный отрезок $\overline{A_1B_1}$ такой, что

$$\left|\overline{A_1B_1}\right| = |\lambda| \cdot \left|\overline{AB}\right|,$$

сонаправленный с $\overline{AB}$, если $\lambda > 0$, и противонаправленный с $\overline{AB}$, если $\lambda < 0$.

На рис. 4.2. приведены примеры умножения числа на направленный отрезок.

Отметим следующие свойства операции умножения числа на направленный отрезок:

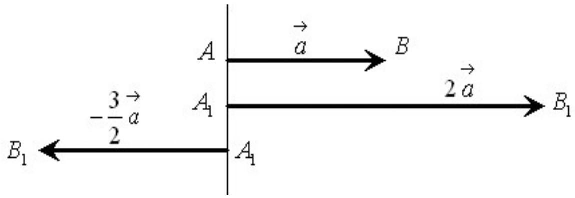


Рис. 4.2. Операция умножения числа на направленный отрезок

- при сложении с нулевым направленным отрезком любой другой направленный отрезок не меняется;
- при умножении числа ноль на любой направленный отрезок получается нулевой направленный отрезок.

Определение 4.1.5. *Множеством векторов* назовем совокупность всех направленных отрезков, для которых введены операции сравнения, сложения и умножение числа на отрезок. Каждый элемент этого множества назовем *вектором.*

Векторы, в отличие от направленных отрезков, будем выделять «верхней стрелкой». Например, $\overrightarrow{AB}$ или $\overrightarrow{a}$. Длину вектора принято обозначать $\left|\overrightarrow{a}\right|$.

Различие между направленным отрезком и вектором состоит в том, что первый может существовать сам по себе, без какой-либо связи с другими направленными отрезками. В то время как понятие вектора предполагает выполнимость операций сравнения, сложения и умножение числа на вектор. Поэтому каждый вектор является направленным отрезком, но не всякий направленный отрезок есть вектор.

Иногда говорят, что вектор это объект, который характеризуется как величиной, так и направлением. Однако следующий пример показывает, что такое определение, вообще говоря, не является корректным.

Рассмотрим две расположенные под прямым углом улицы, соединяющиеся в третью, идущую под углом 45° по отношению к первым двум (рис. 4.3). Допустим, что по каждой из двух первых улиц в сторону их соединения движется поток автомобилей 500 авт./час. Поток автомобилей естественно характеризуется как величиной, так и направлением. Если предположить, что он есть вектор, то суммарный поток по третьей улице согласно определению 4.1.3 составит $500\sqrt{2} \approx 700$ авт./час., а не ожидаемые 1000 авт./час. Данное противоречие объяс-

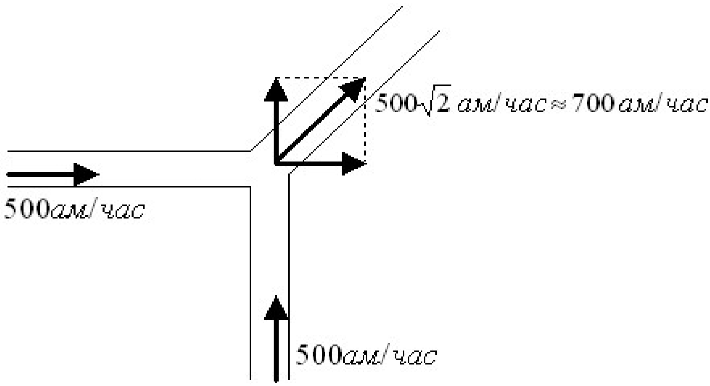


Рис. 4.3. Не всякий направленный отрезок есть вектор

няется тем, что, хотя поток автомобилей характеризуется величиной и направлением, он не является вектором в смысле определения (4.1.5.)

В заключение рассмотрим еще одну векторную операцию, используемую при решении большого числа прикладных задач.

Определение 4.1.6. **Скалярным произведением** двух ненулевых векторов называется число, равное произведению длин этих векторов на косинус угла между ними. Если хотя бы один из сомножителей есть нулевой вектор, то скалярное произведение считается равным нулю.

Скалярное произведение векторов $\overrightarrow{a}$ и $\overrightarrow{b}$ принято обозначать $\left(\overrightarrow{a}\,,\overrightarrow{b}\right)$, значит, в силу определения 4.1.6, имеет место равенство

$$\left(\overrightarrow{a}\,,\overrightarrow{b}\right) = \left[\begin{array}{ll} \left|\overrightarrow{a}\right| \cdot \left|\overrightarrow{b}\right| \cdot \cos \widehat{\overrightarrow{a}\,\overrightarrow{b}}\,, & \text{если } \overrightarrow{a} \neq \overrightarrow{o} \text{ и если } \overrightarrow{b} \neq \overrightarrow{o}\,, \\ 0\,, & \text{если } \overrightarrow{a} = \overrightarrow{o} \text{ или } \overrightarrow{b} = \overrightarrow{o}\,. \end{array}\right.$$

Перечислим основные свойства скалярного произведения, вытекающие непосредственно из его определения.

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

2°. $\left(\lambda\overrightarrow{a}\,,\overrightarrow{b}\right) = \lambda \cdot \left(\overrightarrow{a}\,,\overrightarrow{b}\right)$,

3°. $\left(\overrightarrow{a_1} + \overrightarrow{a_2}, \overrightarrow{b}\right) = \left(\overrightarrow{a_1}, \overrightarrow{b}\right) + \left(\overrightarrow{a_2}, \overrightarrow{b}\right)$,

4°. $\left|\overrightarrow{a}\right| = \sqrt{\left(\overrightarrow{a}\,,\overrightarrow{a}\right)}$, 5°. Если $\vec{a} \neq \vec{o}$ и $\vec{b} \neq \vec{o}$, то $\cos \widehat{\vec{a}\,\vec{b}} = \dfrac{\left(\vec{a}, \vec{b}\right)}{\left|\vec{a}\right| \cdot \left|\vec{b}\right|}$,

6°. $\left(\vec{a}, \vec{a}\right) = 0 \quad \Leftrightarrow \quad \vec{a} = \vec{o}$.

4.2. Линейная зависимость векторов

Возможность выполнения векторных операций позволяет придти к выводу, что *любой* набор векторов должен обладать одним из двух взаимоисключающих свойств: *линейной зависимостью* или *линейной независимостью.*

Суть линейной зависимости заключается в возможности выразить один из векторов данного набора через другие при помощи операций сравнения, сложения и умножения числа на вектор. Линейная независимость означает невозможность такого выражения. Строгие определения этих свойств формулируются в следующем виде.

Определение 4.2.1. Набор векторов $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$ называется *линейно зависимым*, если существует линейная функция этих векторов, коэффициенты которой не равны нулю одновременно, равная нулевому вектору. Иначе говоря, выполняется условие

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \ldots + \lambda_k \vec{a}_k = \vec{o},$$

где числа $\lambda_1, \lambda_2, \ldots, \lambda_k$ не равны нулю одновременно.

Определение 4.2.2. Набор векторов $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$ называется *линейно независимым*, если из условия

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \ldots + \lambda_k \vec{a}_k = \vec{o},$$

следует, что $\lambda_1 = \lambda_2 = \ldots = \lambda_k = 0$.

Пример 4.2.1. Три вектора $\overrightarrow{AB}, \overrightarrow{AD}$ и $\overrightarrow{BD}$, совпадающие с двумя смежными сторонами и противолежащей углу A диагональю прямоугольника (рис.4.4.*a*), линейно зависимы.

Действительно, согласно правилу треугольника $\overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AD}$. Откуда следует, что

$$1 \cdot \overrightarrow{AB} + (-1) \cdot \overrightarrow{AD} + 1 \cdot \overrightarrow{BD} = \overrightarrow{o} \ ,$$

что, согласно определению 4.2.1 и означает линейную зависимость данной тройки векторов.

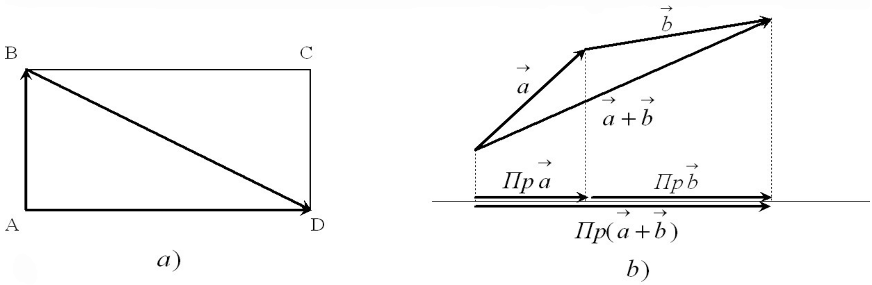


Рис. 4.4. Линейная зависимость и линейная независимость векторов.

Пример 4.2.2. Покажем теперь, что рассмотренные в предыдущем примере векторы $\overrightarrow{AB}$ и $\overrightarrow{AD}$, будут линейно независимыми.

Пусть

$$\lambda_1 \cdot \overrightarrow{AB} + \lambda_2 \cdot \overrightarrow{AD} = \overrightarrow{o} \ . \tag{4.2.1}$$

Обозначим ортогональную проекцию некоторого вектора $\overrightarrow{a}$ на прямую, проходящую через точки A и D, как $Пр\,\overrightarrow{a}$. Тогда из (4.2.1) будет следовать равенство

$$Пр\left(\lambda_1 \cdot \overrightarrow{AB} + \lambda_2 \cdot \overrightarrow{AD}\right) = Пр\,\overrightarrow{o} \ ,$$

ибо очевидно, что равные векторы должны иметь равные ортогональные проекции на одну и ту же прямую.

Поскольку ортогональная проекция суммы векторов равна сумме ортогональных проекций слагаемых (см. рис. 4.4b), то последнее равенство можно записать в виде

$$\lambda_1 \cdot Пр\overrightarrow{AB} + \lambda_2 \cdot Пр\overrightarrow{AD} = Пр\,\overrightarrow{o} \ ,$$

что, в свою очередь, в силу очевидных соотношений

$$Пр\overrightarrow{AD} = \overrightarrow{AD} \ , \quad Пр\overrightarrow{AB} = \overrightarrow{o} \ , \quad Пр\,\overrightarrow{o} = \overrightarrow{o} \ ,$$

приводит к равенству

$$\lambda_1 \cdot \overrightarrow{o} + \lambda_2 \cdot \overrightarrow{AD} = \overrightarrow{o}\ ,$$

которое является верным лишь при $\lambda_2 = 0$. Но, подставив $\lambda_2 = 0$ в (4.2.1), получим и $\lambda_1 = 0$.

Таким образом, из условия (4.2.1) следует $\lambda_1 = \lambda_2 = 0$, что в силу определения 4.2.2 означает линейную независимость векторов $\overrightarrow{AB}$ и $\overrightarrow{AD}$.

4.3. Координатное представление геометрических объектов

4.3.1. Прямоугольная декартова система координат

Векторы оказываются удобным средством описания геометрических объектов – линий, фигур и тел, графического представления свойств функций, большого числа физических законов, а также разнообразных процессов и явлений социально-экономической природы. При этом часто возникает задача выражения некоторого конкретного вектора через набор известных векторов при помощи операций сравнения, сложения и умножения числа на вектор.

Рассмотрим следующий способ описания точки на плоскости.

Выберем на плоскости некоторую фиксированную точку O, которую будем называть *началом координат.* Тогда на этой плоскости любая точка M может быть однозначно задана вектором $\overrightarrow{OM}$, начало которого находится в точке O, а конец – в точке M.

Пусть имеется пара векторов $\overrightarrow{\imath}$ и $\overrightarrow{\jmath}$ таких, что

1) каждый из них *единичной* длины;
2) они взаимно *перпендикулярны*;
3) их начала находятся в точке O.

Тогда *существует единственная* пара чисел x и y таких, что

$$\overrightarrow{OM} = x \cdot \overrightarrow{\imath} + y \cdot \overrightarrow{\jmath}\,.$$

Определение 4.3.1.1. Совокупность точки O и пары векторов $\overrightarrow{\imath}$ и $\overrightarrow{\jmath}$, (именуемых часто *ортами* или *базисными векторами*), называется *декартовой прямоугольной системой координат*, а числа x и y называются *декартовыми прямоугольными координатами* (или просто *координатами*) точки M в системе координат $\left\{O, \overrightarrow{\imath}, \overrightarrow{\jmath}\right\}$.

Определение 4.3.1.2. Вектор $\overrightarrow{OM}$ называется *радиусом-вектором* точки M в системе координат $\left\{O, \overrightarrow{\imath}, \overrightarrow{\jmath}\right\}$, а равенство

$$\overrightarrow{OM} = x \cdot \overrightarrow{\imath} + y \cdot \overrightarrow{\jmath}$$

называется *координатным разложением* вектора $\overrightarrow{OM}$.

Покажем теперь, что, с одной стороны, зная радиус-вектор точки M, можно найти ее координаты и, с другой стороны, по координатам точки M однозначно находится ее радиус-вектор.

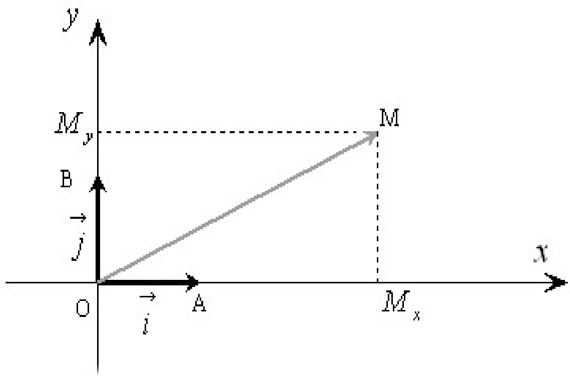


Рис. 4.5. Определение координат точки M.

1) Пусть имеется система координат $\left\{O, \overrightarrow{\imath}, \overrightarrow{\jmath}\right\}$ и задана точка M. Проведем через точку O и векторы $\overrightarrow{\imath}$ и $\overrightarrow{\jmath}$ прямые Ox и Oy как показано на рис. 4.5. Превратим эти прямые в *числовые оси*, приняв точку O за нулевую, и приписав концам векторов $\overrightarrow{\imath}$ и $\overrightarrow{\jmath}$ – точкам A и B значения $+1$.

Опустив из точки M перпендикуляры на Ox и Oy, получим пару точек M_x и M_y, которым на осях будут соответствовать числовые значения x и y. Таким образом, любой точке плоскости может быть поставлена в соответствие упорядоченная пара чисел $\{x, y\}$. А, поскольку основания опущенных перпендикуляров определяются однозначно (как точки пересечения непараллельных прямых), то и координатная пара чисел x и y для точки M единственная.

2) Рассмотрим теперь обратную задачу: найти радиус-вектор точки M по значениям ее координат в системе $\left\{O, \vec{\imath}, \vec{\jmath}\right\}$. Зная пару чисел x и y – координаты точки M, построим векторы $\overrightarrow{OC}$ и $\overrightarrow{OD}$ такие, что $\overrightarrow{OC} = x \cdot \vec{\imath}$ и $\overrightarrow{OD} = y \cdot \vec{\jmath}$. См. рис. 4.6.

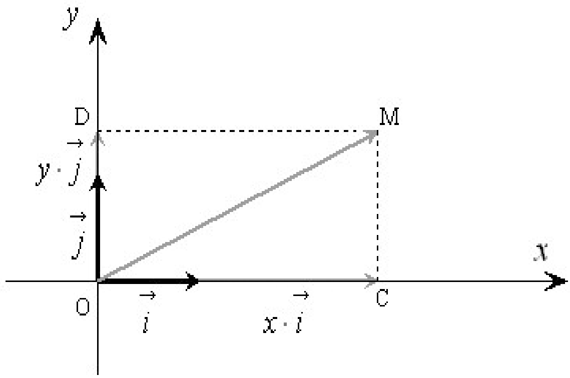


Рис. 4.6. Восстановление точки M по ее координатам.

В нашем случае $\overrightarrow{OM} = x \cdot \vec{\imath} + y \cdot \vec{\jmath}$, и по правилу параллелограмма строим радиус-вектор $\overrightarrow{OM}$, однозначно определяемый конец которого – искомая точка M.

Поскольку декартовы координаты точки на плоскости (или вектора) образуют *упорядоченную* пару чисел, то их удобно записывать в виде двухкомпонентного столбца $\left\| \begin{matrix} x \\ y \end{matrix} \right\|$, первым элементом которого является x-координата точки M, часто именуемая в литературе *абсциссой,* а вторым элементом – y-координата, называемая *ординатой.* Этот столбец, являющийся специальным способом описания точки или вектора, принято называть *координатным представлением* точки (или вектора). Договоримся также, что можно использовать сокращенное обозначение координатного представления вектора или точки в виде $\left\| \vec{a} \right\|$ или $\left\| \overrightarrow{OM} \right\|$.

В заключение сделаем два важных замечания. Во-первых, помимо декартовой системы координат возможно использование иных способов описания геометрических объектов. Например положение точки на поверхности Земли или на небесной сфере удобнее задавать в так называемой *сферической системе координат* с помощью двух упорядоченных чисел, являющихся угловыми отклонениями от заранее выбранных плоскостей (*широта* и *долгота* точки.) Для наших дальнейших целей будет достаточно декартовой системы координат, поэтому, употребление термина «система координат» будет подразумевать именно прямоугольную декартову систему.

Во-вторых, использованный способ построения декартовой системы координат не является единственно возможным. В качестве базисных можно выбирать так же неравные по длине и/или неортогональные векторы. Существенна лишь возможность построения, и притом единственного, координатного представления для каждой точки плоскости. В курсе аналитической геометрии доказывается, что необходимым и достаточным условием такой возможности является линейная независимость базисных векторов. Заметим, что линейная независимость базисных векторов $\vec{\imath}$ и $\vec{\jmath}$, использованных нами для построения прямоугольной декартовой системы координат, показана в примере 4.2.2.

4.3.2. Операции с векторами в координатном представлении

Использование декартовой системы позволяет при помощи координат (то есть аналитически) описывать векторы и точки, избегая интуитивных или геометрически наглядных представлений. При этом, необходимо выяснить, как будут выполняться операции сравнения, сложения и умножения числа на вектор в координатном представлении. Ответ на этот вопрос получить несложно, если вспомнить, что координатное представление вектора – двухкомпонентный столбец, является частным случаем матриц, для которых также были введены операции сравнения, сложения и умножения числа на матрицу.

Рассмотрим последовательно все три операции с векторами.

1) *Сравнение векторов.* Пусть $\left\|\vec{a}_1\right\| = \left\|\begin{matrix} x_1 \\ y_1 \end{matrix}\right\|$ и $\left\|\vec{a}_2\right\| = \left\|\begin{matrix} x_2 \\ y_2 \end{matrix}\right\|$, тогда

$$\vec{a}_1 = \vec{a}_2 \quad \Longleftrightarrow \quad \left\|\vec{a}_1\right\| = \left\|\vec{a}_2\right\| \quad \Longleftrightarrow \quad \left\|\begin{matrix} x_1 \\ y_1 \end{matrix}\right\| = \left\|\begin{matrix} x_2 \\ y_2 \end{matrix}\right\|.$$

В самом деле, равенство $\overrightarrow{a}_1 = \overrightarrow{a}_2$ означает, что

$$x_1 \cdot \overrightarrow{\imath} + y_1 \cdot \overrightarrow{\jmath} = x_2 \cdot \overrightarrow{\imath} + y_2 \cdot \overrightarrow{\jmath}$$

или

$$(x_1 - x_2) \cdot \overrightarrow{\imath} + (y_1 - y_2) \cdot \overrightarrow{\jmath} = \overrightarrow{o}\ .$$

В силу свойств векторов $\overrightarrow{\imath}$ и $\overrightarrow{\jmath}$ (а именно, их линейной независимости) последнее равенство возможно лишь в случае, когда числа $x_1 - x_2$ и $y_1 - y_2$ равны нулю одновременно. Поэтому справедливо

$$\begin{cases} x_1 - y_1 = 0 \\ x_2 - y_2 = 0 \end{cases} \iff \left\| \begin{matrix} x_1 - y_1 \\ x_2 - y_2 \end{matrix} \right\| = \left\| \begin{matrix} 0 \\ 0 \end{matrix} \right\| \iff \left\| \begin{matrix} x_1 \\ y_1 \end{matrix} \right\| = \left\| \begin{matrix} x_2 \\ y_2 \end{matrix} \right\| .$$

2) *Сложение векторов.* Координатное представление суммы векторов $\overrightarrow{a}_1 + \overrightarrow{a}_2$ будет иметь вид

$$\left\| \overrightarrow{a}_1 + \overrightarrow{a}_2 \right\| = \left\| \overrightarrow{a}_1 \right\| + \left\| \overrightarrow{a}_2 \right\| .$$

Действительно, если $\left\| \overrightarrow{a}_1 \right\| = \left\| \begin{matrix} x_1 \\ y_1 \end{matrix} \right\|$ и $\left\| \overrightarrow{a}_2 \right\| = \left\| \begin{matrix} x_2 \\ y_2 \end{matrix} \right\|$, то

$$\overrightarrow{a}_1 = x_1 \cdot \overrightarrow{\imath} + y_1 \cdot \overrightarrow{\jmath} \text{ и } \overrightarrow{a}_2 = x_2 \cdot \overrightarrow{\imath} + y_2 \cdot \overrightarrow{\jmath}\ .$$

Но тогда будут верными следующие равенства

$$\left\| \overrightarrow{a}_1 + \overrightarrow{a}_2 \right\| = \left\| (x_1 \cdot \overrightarrow{\imath} + y_1 \cdot \overrightarrow{\jmath}) + (x_2 \cdot \overrightarrow{\imath} + y_2 \cdot \overrightarrow{\jmath}) \right\| =$$

$$= \left\| (x_1 + x_2) \cdot \overrightarrow{\imath} + (y_1 + y_2) \cdot \overrightarrow{\jmath}) \right\| =$$

$$= \left\| \begin{matrix} x_1 + x_2 \\ y_1 + y_2 \end{matrix} \right\| = \left\| \begin{matrix} x_1 \\ y_1 \end{matrix} \right\| + \left\| \begin{matrix} x_2 \\ y_2 \end{matrix} \right\| = \left\| \overrightarrow{a}_1 \right\| + \left\| \overrightarrow{a}_2 \right\| .$$

3) *Умножение числа на вектор.* Рассуждения, аналогичные предыдущим, приводят к равенству

$$\left\| \lambda \cdot \overrightarrow{a} \right\| = \lambda \cdot \left\| \overrightarrow{a} \right\| .$$

В итоге мы приходим к важному заключению, что

Операции с векторами в координатном представлении выполняются так же, как соответствующие операции с матрицами.

Исходя из свойств перпендикуляров, опущенных из точки на прямую, не трудно видеть, что $OB = OA + AB = a_x + b_x$ и $OD = OC + CD = a_y + b_y$. Операцию сложения векторов в координатной форме иллюстрирует рис.4.7.

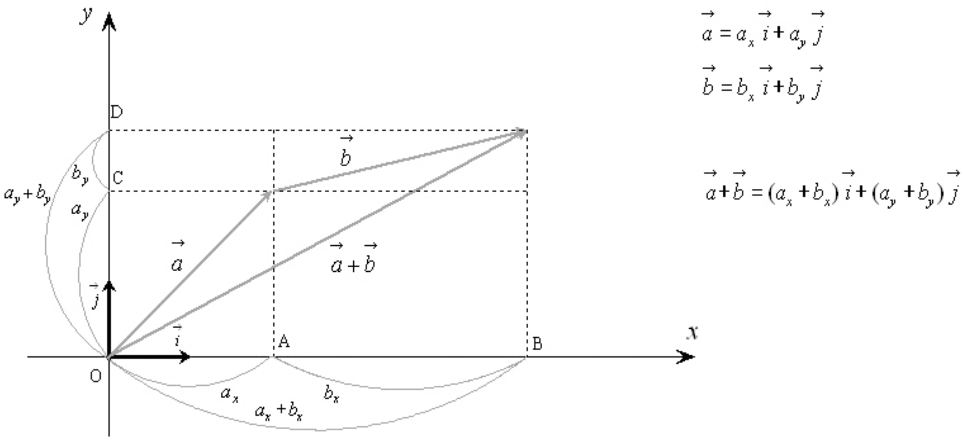


Рис. 4.7. Суммирование векторов в координатной форме.

Отметим, что для радиуса-вектора $\vec{r}$ произвольной точки M, имеющей координатное разложение

$$\vec{r} = x \cdot \vec{\imath} + y \cdot \vec{\jmath},$$

будут справедливы равенства

$$\|\vec{r}\| = x \cdot \left\|\vec{\imath}\right\| + y \cdot \left\|\vec{\jmath}\right\| \quad \text{или} \quad \left\| \begin{matrix} x \\ y \end{matrix} \right\| = x \cdot \left\| \begin{matrix} 1 \\ 0 \end{matrix} \right\| + y \cdot \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\|,$$

поскольку очевидно, что

$$\left\|\vec{\imath}\right\| = \left\|1 \cdot \vec{\imath} + 0 \cdot \vec{\jmath}\right\| = \left\| \begin{matrix} 1 \\ 0 \end{matrix} \right\| \quad \text{и} \quad \left\|\vec{\jmath}\right\| = \left\|0 \cdot \vec{\imath} + 1 \cdot \vec{\jmath}\right\| = \left\| \begin{matrix} 0 \\ 1 \end{matrix} \right\|.$$

Координатное представление векторов позволяет исследовать геометрические объекты в аналитической форме без использования каких-либо наглядных или интуитивно очевидных представлений, поскольку в этом случае, например, *прямую на плоскости* (геометрический объект, состоящий из точек и обладающий набором соответствующих свойств) можно рассматривать и соответственно анализировать как *совокупность радиусов-векторов* этих точек.

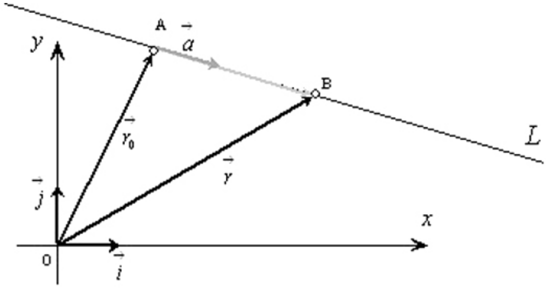


Рис. 4.8. Векторная форма описания прямой на плоскости.

Действительно, пусть некоторая прямая L (см. рис. 4.8) проходит через фиксированную точку A с известным радиусом-вектором $\vec{r}_0$ в направлении, задаваемом ненулевым вектором $\vec{a}$. Тогда $\vec{r}$ – радиус-вектор *произвольной* точки B этой прямой, можно выразить через векторы $\vec{r}_0$ и $\vec{a}$ при помощи векторных операций (по правилам треугольника и умножения числа на вектор) следующим образом.

$$\vec{r} = \vec{r}_0 + \overrightarrow{AB} \qquad \overrightarrow{AB} = \lambda \cdot \vec{a}\ ,$$

где λ – коэффициент пропорциональности между векторами $\vec{a}$ и $\overrightarrow{AB}$. Это позволяет представить совокупность радиусов-векторов всех точек прямой L в виде

$$\vec{r} = \vec{r}_0 + \lambda \cdot \vec{a}\ , \quad \forall\lambda.$$

Далее, если известны координатные представления векторов

$$\vec{r} = \left\| \begin{matrix} x \\ y \end{matrix} \right\|, \quad \vec{r}_0 = \left\| \begin{matrix} x_0 \\ y_0 \end{matrix} \right\| \text{ и } \vec{a} = \left\| \begin{matrix} a_x \\ a_y \end{matrix} \right\|,$$

то прямую L можно задавать как совокупность точек плоскости, координаты $\{x, y\}$ которых удовлетворяют, последовательно получаемым одно из другого, соотношениям

$$\|\vec{r}\| = \|\vec{r}_0\| + \lambda \cdot \left\|\vec{a}\right\|, \qquad \left\| \begin{matrix} x \\ y \end{matrix} \right\| = \left\| \begin{matrix} x_0 \\ y_0 \end{matrix} \right\| + \lambda \cdot \left\| \begin{matrix} a_x \\ a_y \end{matrix} \right\|,$$

$$\left\{ \begin{array}{l} x = x_0 + \lambda \cdot a_x, \\ y = y_0 + \lambda \cdot a_y. \end{array} \right. \ \forall\lambda.$$

Наконец, из последней системы уравнений, после исключения параметра λ, получается уравнение вида $Ax+By+C=0$, где $A=a_y$, $B=-a_x$ и $C=-a_yx_0-a_xy_0$, которое также является одной из координатных форм задания прямой на плоскости. Покажите самостоятельно, что при этом числа A и B не могут быть равными нулю *одновременно.*

4.3.3. Скалярное произведение и координаты

Также как и в случае операций сравнения, сложения и умножения числа на вектор, координатный подход может быть использован и для описания операции скалярного произведения. Получим правило подсчета скалярного произведения векторов, заданных их координатными представлениями в прямоугольной декартовой системе координат Oxy. Пусть

$$\vec{a}=\alpha_1\cdot\vec{\imath}+\alpha_2\cdot\vec{\jmath} \quad \text{или, что то же самое,} \quad \left\|\vec{a}\right\|=\left\|\begin{matrix}\alpha_1\\ \alpha_2\end{matrix}\right\| \quad \text{и}$$

$$\vec{b}=\beta_1\cdot\vec{\imath}+\beta_2\cdot\vec{\jmath} \quad \text{или} \quad \left\|\vec{b}\right\|=\left\|\begin{matrix}\beta_1\\ \beta_2\end{matrix}\right\| .$$

Найдем скалярное произведение $\vec{a}$ и $\vec{b}$.

$$\left(\vec{a},\vec{b}\right)=\left(\alpha_1\cdot\vec{\imath}+\alpha_2\cdot\vec{\jmath},\ \beta_1\cdot\vec{\imath}+\beta_2\cdot\vec{\jmath}\right)=$$

$$=\alpha_1\beta_1\left(\vec{\imath},\vec{\imath}\right)+\alpha_1\beta_2\left(\vec{\imath},\vec{\jmath}\right)+\alpha_2\beta_1\left(\vec{\jmath},\vec{\imath}\right)+\alpha_2\beta_2\left(\vec{\jmath},\vec{\jmath}\right)=\alpha_1\beta_1+\alpha_2\beta_2 ,$$

поскольку, согласно определениям 4.1.6 и 4.3.1.1,

$$\left(\vec{\imath},\vec{\imath}\right)=1 , \quad \left(\vec{\imath},\vec{\jmath}\right)=0 , \quad \left(\vec{\jmath},\vec{\imath}\right)=0 , \quad \left(\vec{\jmath},\vec{\jmath}\right)=1 .$$

Таким образом, для векторов заданных в координатном представлении, скалярное произведение находится по формулам

$$\left(\vec{a},\vec{b}\right)=\alpha_1\beta_1+\alpha_2\beta_2 \tag{4.3.3.1}$$

или

$$\left(\vec{a},\vec{b}\right)=\left\|\vec{a}\right\|^{\mathrm{T}}\left\|\vec{b}\right\|=\left\|\alpha_1\ \ \alpha_2\right\|\left\|\begin{matrix}\beta_1\\ \beta_2\end{matrix}\right\| .$$

В частности, из формулы 4°(см. §4.1) получаем, что длину вектора $\vec{a}$ можно выразить через его координаты следующим образом

$$\left|\vec{a}\right| = \sqrt{\left(\vec{a}, \vec{a}\right)} = \sqrt{\alpha_1^2 + \alpha_2^2}\,. \tag{4.3.3.2}$$

Узнали знаменитую теорему *Пифагора*?

Наконец, отметим, что не только скалярное произведение можно выражать через координаты, но и координаты могут быть представлены как скалярное произведение. Действительно, если обе части равенства

$$\vec{a} = \alpha_1 \cdot \vec{\imath} + \alpha_2 \cdot \vec{\jmath}$$

скалярно умножить на вектор $\vec{\imath}$, то из

$$\left(\vec{\imath}, \vec{a}\right) = \left(\vec{\imath}, \quad \alpha_1 \cdot \vec{\imath} + \alpha_2 \cdot \vec{\jmath}\right) = \alpha_1 \left(\vec{\imath}, \vec{\imath}\right) + \alpha_2 \left(\vec{\imath}, \vec{\jmath}\right)\,,$$

в силу $\left(\vec{\imath}, \vec{\imath}\right) = 1$ и $\left(\vec{\imath}, \vec{\jmath}\right) = 0$, следует, что

$$\alpha_1 = \left(\vec{\imath}, \vec{a}\right)\,.$$

Покажите самостоятельно, что аналогично можно получить $\alpha_2 = \left(\vec{\jmath}, \vec{a}\right)$.

Пример 4.3.3.1. Найти скалярное произведение векторов $\overrightarrow{OA}$ и $\overrightarrow{OB}$, если координатные представления этих векторов соответственно равны

$$\left\|\overrightarrow{OA}\right\| = \left\|\begin{matrix} 3 \\ \sqrt{3} \end{matrix}\right\| \qquad \text{и} \qquad \left\|\overrightarrow{OB}\right\| = \left\|\begin{matrix} 1 \\ \sqrt{3} \end{matrix}\right\|\,,$$

использовав как его определение, так и координатное представление этой операции в ортонормированной системе координат (рис. 4.9).

Решение. По определению 4.1.6 для вычисления скалярного произведения векторов надо найти их длины и угол между ними. По формуле (4.3.3.2) получаем

$$\left|\overrightarrow{OA}\right| = \sqrt{3^2 + (\sqrt{3})^2} = \sqrt{9+3} = \sqrt{12} = \sqrt{4 \cdot 3} = 2\sqrt{3}\,,$$

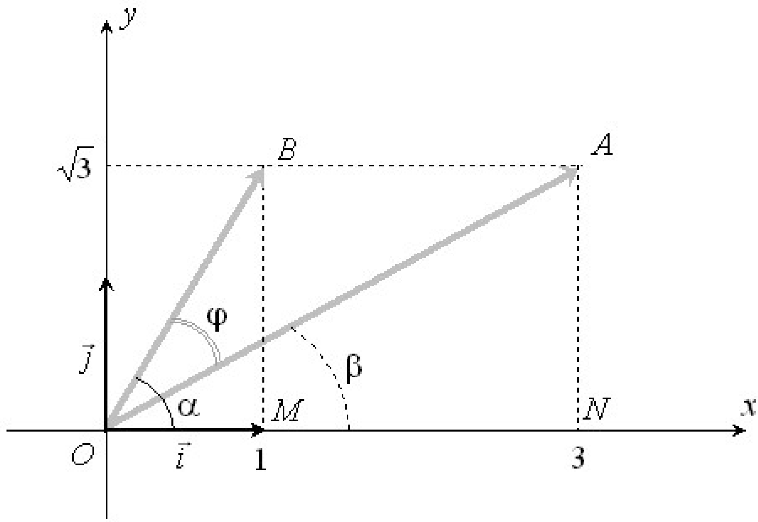


Рис. 4.9. Найти скалярное произведение векторов (пример 4.3.3.1)

$$\left|\overrightarrow{OB}\right| = \sqrt{1^2 + (\sqrt{3})^2} = \sqrt{1+3} = \sqrt{4} = 2\ .$$

Из рис. 4.9, на основании свойств прямоугольных треугольников OAN и OBM, получаем, что угол $\alpha = 60^\circ$, а угол $\beta = 30^\circ$. Тогда угол φ – между векторами $\overrightarrow{OA}$ и $\overrightarrow{OB}$, равен 30°. Окончательно, согласно определению 4.1.6, находим искомое значение скалярного произведения

$$\left(\overrightarrow{OA}, \overrightarrow{OB}\right) = \left|\overrightarrow{OA}\right| \cdot \left|\overrightarrow{OB}\right| \cdot \cos\varphi = (2\sqrt{3}) \cdot 2 \cdot \cos 30^\circ = (2\sqrt{3}) \cdot 2 \cdot \frac{\sqrt{3}}{2} = 6\ .$$

Использование формулы (4.3.3.1) позволяет получить тот же результат заметно проще: $\left(\overrightarrow{OA}, \overrightarrow{OB}\right) = 3 \cdot 1 + \sqrt{3} \cdot \sqrt{3} = 3 + 3 = 6\ .$

4.4. Координатный метод описания функций

Как уже отмечалось, координатный подход может быть применен для наглядной геометрической интерпретации разнообразных математических объектов. Для описания свойств функций особенно удобным оказывается сочетания понятия предела и координатной формы их представления, называемой *графиком функции.*

Определение 4.4.1. Графиком функции $y = f(x)$ называется множество точек на координатной плоскости Oxy, радиусы-векторы которых имеют в выбранной прямоугольной системе координат представление вида $\left\| \begin{matrix} x \\ f(x) \end{matrix} \right\|$.

Важно понимать, что, хотя график любой функции есть некоторая линия на координатной плоскости, обратное, вообще говоря, не верно. Не всякая линия на координатной плоскости является графиком какой-либо функции.

Действительно, согласно определению функции для каждого значения аргумента x, принадлежащего области определения, существует *единственное* значение y, определяемое соотношением $y = f(x)$. Но в случае, когда координаты точек некоторой линии связаны уравнением $F(x, y) = 0$, выбранному значению координаты x может соответствовать большее, чем одно число значений координаты y, удовлетворяющих этому уравнению.

Проиллюстрируем этот факт следующим примером. Прямая на координатной плоскости (как было выше показано) может задаваться уравнением вида $Ax + By + C = 0$, причем числа A и B не равны нулю одновременно. Легко видеть, это уравнение может быть записано в форме $y = f(x)$ только в случае, когда $B \neq 0$. Тогда (и только тогда!) прямая является графиком функции, и ее формульная запись (в предположении, что прямая имеет угловой коэффициент k и проходит через точку с координатами $\left\| \begin{matrix} x_0 \\ y_0 \end{matrix} \right\|$) есть $y = k(x - x_0) + y_0$. Если же $B = 0$, то прямая $Ax + By + C = 0$ параллельна оси Oy и не является графиком какой-либо функции.

Другой важной особенностью координатного метода представления функций является *зависимость этого представления от выбора системы координат.* Действительно, если, скажем, один наблюдатель выбрал в качестве начала координат точку O_1, а его коллега точку O_2, то очевидно, что одна и та же точка M для каждого из этих наблюдателей будет иметь хотя и единственные, но различные координаты. Аналогичная проблема возникает из-за произвольности выбора направления координатных осей и единиц длины вдоль них.

С одной стороны, эта особенность координатного подхода является его недостатком, поскольку наблюдателям, использующим различные системы координат, необходимо при сравнении результатов их наблюдений затрачивать дополнительные вычислительные ресурсы. Однако с другой стороны, свобода в выборе системы координат может быть использована для упрощения графического описания функции. Иными словами, можно попытаться найти такую систему координат, в которой это описание функции оказывается наиболее простым.

Поясним сказанное следующим примером. Пусть требуется построить график функции, которая в некоторой исходной прямоугольной декартовой системе координат $\{O, x, y\}$ задается формулой

$$y = \frac{2x-1}{x-1}\,. \tag{4.4.1}$$

Выберем «новую» прямоугольную систему координат $\{O', x', y'\}$, такую, что точка O' в исходной системе имеет координаты $\left\| \begin{matrix} 1 \\ 2 \end{matrix} \right\|$, а направления осей и единицы измерения длин для обеих систем координат одни и те же. Тогда очевидно, что новые $\left\| \begin{matrix} x' \\ y' \end{matrix} \right\|$ и исходные $\left\| \begin{matrix} x \\ y \end{matrix} \right\|$ координаты любой точки M (см. рис. 4.10а) будут связаны соотношениями

$$\begin{cases} x = x' + 1, \\ y = y' + 2. \end{cases} \tag{4.4.2}$$

Найдем формулу для функции (4.4.1) в новой системе координат. Для этого подставим в (4.4.1) выражения (4.4.2), получим

$$y' + 2 = \frac{2(x'+1)-1}{(x'+1)-1} = \frac{2x'+1}{x'} = 2 + \frac{1}{x'} \quad \text{или} \quad y' = \frac{1}{x'}\,.$$

А это – формула хорошо известной гиперболы. Нарисуем ее график в новой системе координат (см. рис. 4.10б). Теперь остается вернуться в исходную систему координат, удалив с координатной плоскости оси новой системы (см. рис. 4.10в).

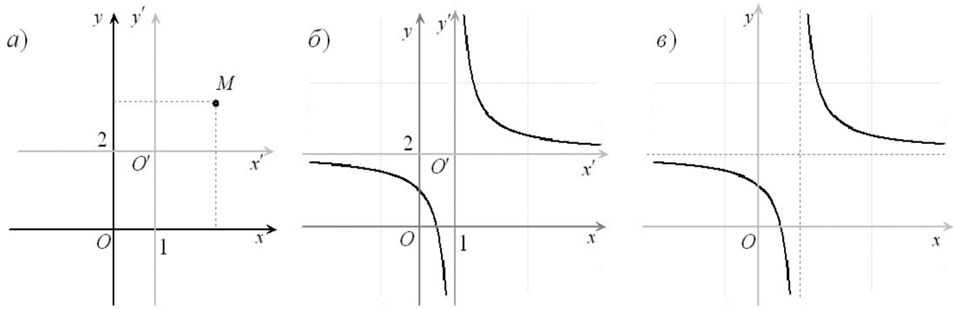


Рис. 4.10. Построение графика гиперболы (4.4.1).

В рассмотренном примере ни слова не было сказано о том, каким образом мы выбирали «новую» систему координат. Однако к этому выбору нетрудно придти, если преобразовать формулу (4.4.1) следующим образом

$$y = \frac{2x-1}{x-1} = \frac{2(x-1)+1}{x-1} = 2 + \frac{1}{x-1} \quad \text{или} \quad y - 2 = \frac{1}{x-1}\,.$$

Теперь соотношения, использованные для замены координат, становятся очевидными.

Стоит обратить внимание, что рассмотренном примере искомый график формально получается из графика стандартной гиперболы двумя последовательными параллельными переносами: на единицу вправо вдоль оси Ox и на две единицы вверх вдоль оси Oy.

Проведя аналогичные рассуждения, можно, основываясь на свойствах прямоугольной декартовой системы координат, придти к заключению, что график функции $y = Af\left(a(x-b)\right) + B$ может быть легко построен по известному графику функции $y = f(x)$. Для этого построения в большом числе практически важных случаев оказывается достаточно знать вид некоторого числа основных из них (см. примеры в §1.1) – таких, например, как

$$y = x^a,\ y = a^x,\ y = \log_a x,\ y = \sin x,\ y = \cos x,\ y = \operatorname{tg} x,$$

$$y = \arcsin x,\ y = \arccos x,\ y = \operatorname{arctg} x$$

и использовать правила преобразования графиков при *переносе начала координат* и *деформации вдоль осей.*

Эти правила, описывающим геометрический смысл каждого из параметров A, B, a и b, формулируются следующим образом

1°. График функции $y = f(x) + B$ получается из графика функции $y = f(x)$ параллельным переносом вдоль оси Oy на $|B|$ единиц вверх, если $B > 0$ и на $|B|$ единиц вниз, если $B < 0$.

2°. График функции $y = f(x - a)$ получается из графика функции $y = f(x)$ параллельным переносом вдоль оси Ox на $|a|$ единиц вправо, если $a > 0$ и на $|a|$ единиц влево, если $a < 0$.

3°. График функции $y = Af(x)$ получается из графика функции $y = f(x)$ при $A > 0$ деформацией (растяжением или сжатием) вдоль оси Ox с коэффициентом $|A|$. Если $A < 0$, то после деформации необходимо выполнить еще зеркально-симметричное отражение относительно оси Ox.

4°. График функции $y = f(ax)$ получается из графика функции $y = f(x)$ при $a > 0$ деформацией (растяжением или сжатием) вдоль оси Oy с коэффициентом $|\frac{1}{a}|$. Если $a < 0$, то после деформации необходимо выполнить еще зеркально-симметричное отражение относительно оси Oy.

Если при построении графика возникает необходимость использования комбинации этих правил, то сначала следует выполнять преобразования 3°–4°, а затем – преобразования 1°–2°.

Рассмотрим в качестве примера некоторые возможные преобразования графика функции $y = x^2$, который показан на рис. 4.11. Обратите внимание, что на рис.4.12–4.21 исходный график показан серым цветом, а преобразованный – черным. Кроме того, для большей наглядности масштабы осей Ox и Oy на некоторых из этих рисунков выбраны различными.

На рис. 4.12 показано преобразование графика $y = x^2$ в график $y = x^2 + 3$, выполняемое сдвигом на 3 единицы вверх по оси Oy.

На рис. 4.13 показано преобразование графика $y = x^2$ в график $y = (x - 2)^2$, выполняемое сдвигом на 2 единицы вправо по оси Ox. Рис. 4.14 иллюстрирует совместное выполнение двух последних преобразований, позволяющих построить график функции $y = (x - 2)^2 + 3$.

Преобразование графика $y = x^2$ в график $y = \frac{1}{2}x^2$, показанное на рис. 4.15, сводится к сжатию исходного графика к оси Ox с коэффициентом $\frac{1}{2}$.

Иногда требуемое преобразование графика может быть выполнено разными способами. Например, преобразование графика $y = x^2$ в

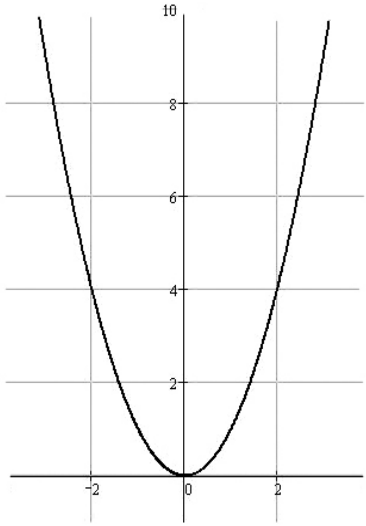


Рис. 4.11. График квадратной параболы $y = x^2$.

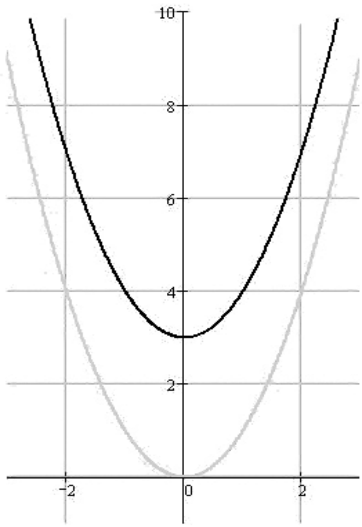


Рис. 4.12. Преобразование графика $y = x^2$ в график $y = x^2 + 3$.

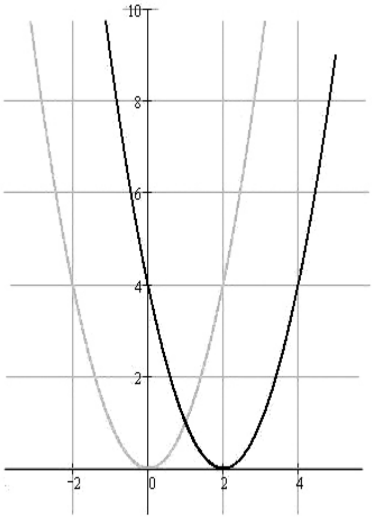


Рис. 4.13. Преобразование графика $y = x^2$ в график $y = (x-2)^2$.

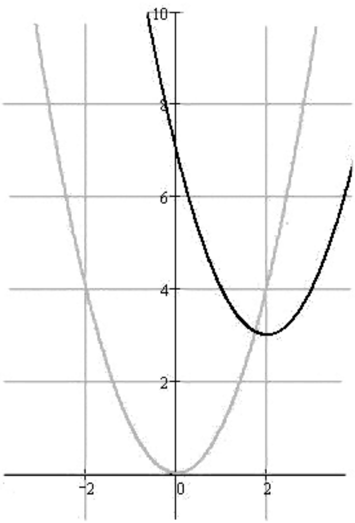


Рис. 4.14. Преобразование графика $y = x^2$ в график $y = (x-2)^2 + 3$.

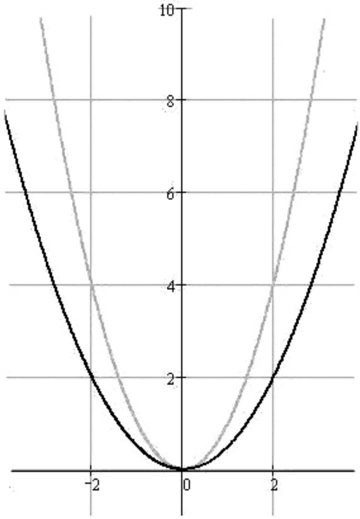


Рис. 4.15. Преобразование графика $y = x^2$ в график $y = \frac{1}{2}x^2$.

график $y = (3x)^2$ можно выполнить или сжатием исходного графика к оси Oy с коэффициентом $\frac{1}{3}$ или растяжением вдоль оси Oy с коэффициентом 9, поскольку функцию $y = (3x)^2$ можно также задать и формулой $y = 9x^2$. Результат этих преобразований показан на рис. 4.16.

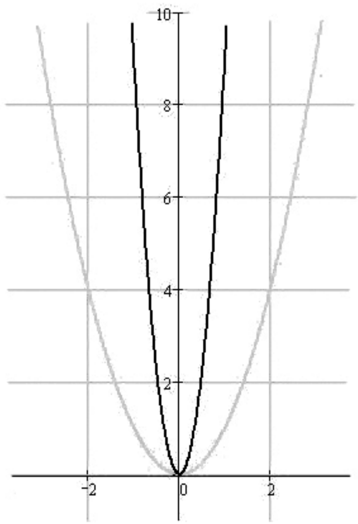


Рис. 4.16. Преобразование графика $y = x^2$ в график $y = (3x)^2$.

Преобразование графика $y = x^2$ в график $y = -3x^2$ выполняется двумя последовательными преобразованиями: растяжением исходного графика вдоль оси Oy с коэффициентом 3 и последующим зеркально-симметричным отражением относительно оси Ox. Итоговый график показан на рис. 4.17.

Рис. 4.18 представляет результат преобразования графика $y = x^2$ в график $y = -3x^2 + 5$, который получается из предыдущего дополнительным сдвигом на 5 единиц вверх по оси Oy.

На рис. 4.19 показан вид графика $y = 2(x - 3)^2 + 1$, который получается из графика $y = 2x^2$ двумя сдвигами: вдоль оси Ox вправо на 3 единицы и сдвигом вверх на одну единицу по оси Oy. Аналогично (см. рис. 4.20) график $y = -(x - 2)^2 + 3$ получается из графика $y = -x^2$ также двумя сдвигами: вдоль оси Ox вправо на 2 единицы и сдвигом

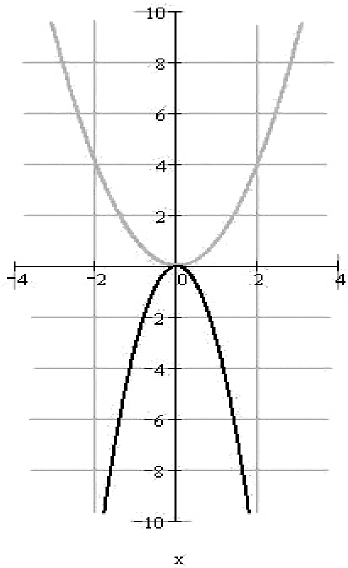


Рис. 4.17. Преобразование графика $y = x^2$ в график $y = -3x^2$.

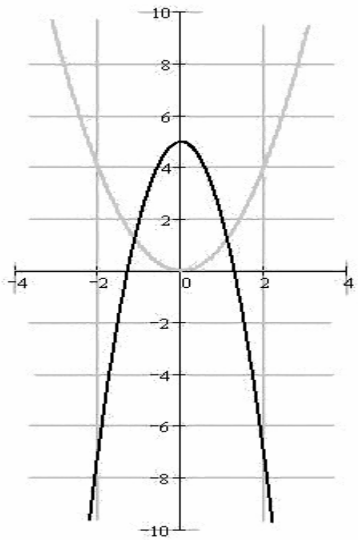


Рис. 4.18. Преобразование графика $y = x^2$ в график $y = -3x^2 + 5$.

вверх на три единицы по оси Oy.

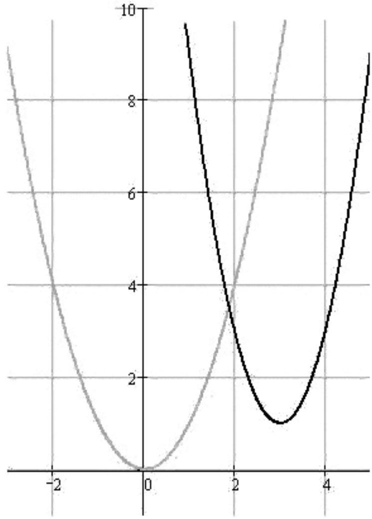


Рис. 4.19. Преобразование графика $y = x^2$ в график $y = 2(x-3)^2 + 1$.

Наконец, для преобразования графика $y = x^2$ в график $y = 3x^2 + 4x + 2$ предварительно следует привести запись функции $y = 3x^2 + 4x + 2$ методом выделения полного квадрата к виду $y = 3(x + \dfrac{2}{3})^2 + \dfrac{2}{3}$. Результат выполнения необходимых преобразований показан на рис. 4.21 .

Достаточно часто оказывается удобно описывать функцию, указывая *локальные свойства,* то есть особенности ее поведения, имеющие место для некоторого (возможно малого) подмножества области определения. Основными локальными свойствами функций являются *экстремальность, выпуклость* и *асимптотичность.*

Вначале напомним определение *экстремальной точки* функции, то есть точки, в которой функция имеет максимум или минимум.

Определение 4.4.2. Будем говорить, что функция $y = f(x)$ имеет *локальный максимум* в некоторой точке x^*, если существует Ω – некоторая окрестность этой точки, такая, что для всех $x \in \Omega$ выполнено неравенство $f(x) \le f(x^*)$. Если же $f(x) \ge f(x^*)$, $\quad \forall x \in \Omega$, то

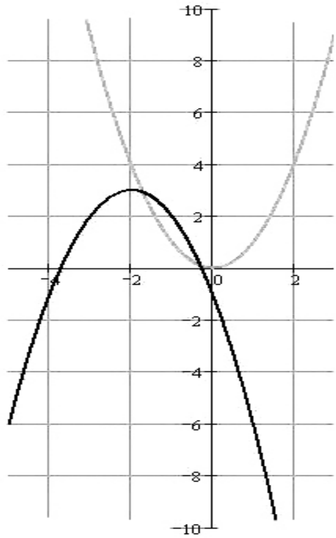


Рис. 4.20. Преобразование графика $y = x^2$ в график $y = -(x-2)^2 + 3$.

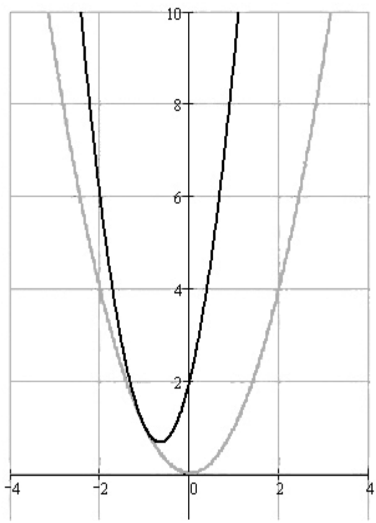


Рис. 4.21. Преобразование графика $y = x^2$ в график $y = 3x^2 + 4x + 2$.

точку x^* называют *локальным минимумом.*

Теперь введем понятие локальной выпуклости функции.

Определение 4.4.3. Функция $y = f(x)$ называется *выпуклой вниз* на некотором множестве Ω в области ее определения, если для любой пары чисел $a, b \in \Omega$ выполнено неравенство

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(a)+f(b)}{2}$$

и называется *выпуклой вверх,* если для любой пары чисел $a, b \in \Omega$

$$f\left(\frac{a+b}{2}\right) \geq \frac{f(a)+f(b)}{2}.$$

Точка, в которой функция меняет направление выпуклости на противоположное, часто именуется как *точка перегиба.*

Геометрический смысл свойства выпуклости иллюстрирует рисунок 4.22. Для функции выпуклой вниз точка C^* – середина хорды AB графика функции $y = f(x)$, оказывается выше точки C, изображающей значение этой функции при $x = \dfrac{a+b}{2}$. В то время как для функции выпуклой вверх, точка C расположена выше точки C^*.

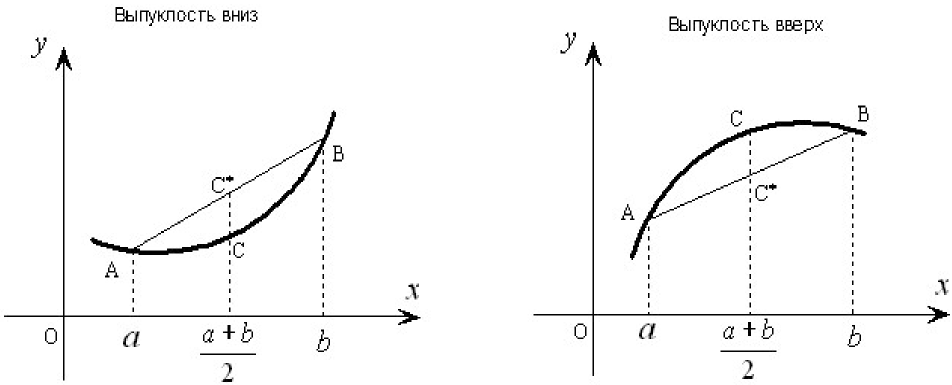


Рис. 4.22. Выпуклость функций

Достаточно часто оказывается, что в окрестности некоторой точки две разные функции «ведут себя примерно одинаково» и, говоря о локальных свойствах одной из них, можно ссылаться на другую. Например, функция $y = \sin x$ в малой окрестности точки $x_0 = 0$ ведет себя примерно так же как и функция $y = x$. При этом разница между их значениями оказывается тем меньше, чем ближе точка x к нулю. Такое совпадение свойств очевидно обусловлено «первым замечательным пределом»

$$\lim_{x\to 0} \frac{\sin x}{x} = 1 ,$$

и потому естественно говорить об эквивалентности поведения пары функций $y = f(x)$ и $y = g(x)$ в окрестности точки x_0, если

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 1 .$$

Отметим, что, говоря об эквивалентном поведении функций, часто употребляются термины «асимптотическая близость» или «асимптотическое совпадение».

В частном случае, когда поведение функции мало отличается от поведения прямой, то есть когда график данной функции неограниченно приближается к этой при стремлении аргумента к некоторому пределу, такую прямую принято называть *асимптотой.* Асимптоты бывают трех типов: вертикальные, горизонтальные и наклонные.

Приведем соответствующие определения.

Определение 4.4.4. Функция $y = f(x)$ имеет в конечной точке x_0 *вертикальную асимптоту,* если $\lim\limits_{x\to x_0} f(x) = \infty$.

Определение 4.4.5. Функция $y = f(x)$ имеет *горизонтальную асимптоту,* если $\lim\limits_{x\to\infty} f(x) = a$, где a – некоторое конечное число.

Пример 4.4.1. Функция $y = \dfrac{3x-1}{x+2}$ имеет очевидно вертикальную асимптоту при $x_0 = -2$ и горизонтальную асимптоту с $a = 3$, поскольку

$$\lim_{x\to\infty} \frac{3x-1}{x+2} = \lim_{x\to\infty} \frac{3 - \dfrac{1}{x}}{1 + \dfrac{2}{x}} = 3 ,$$

что иллюстрирует рис. 4.22. Заметим, что для большей наглядности асимптоты принято изображать *штриховыми линиями* в той же системе координат, что и графики функций.

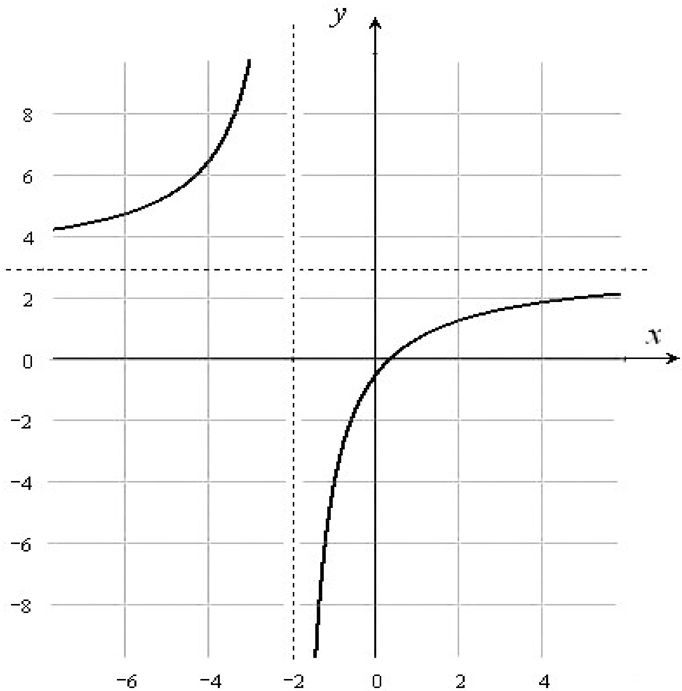


Рис. 4.23. Горизонтальная и вертикальная асимптоты функции $y = \dfrac{3x-1}{x+2}$.

Определение 4.4.6. Функция $y = f(x)$ имеет *наклонную асимптоту* вида $y = ax + b$, если

$$\begin{cases} \lim\limits_{x\to\infty} \dfrac{f(x)}{x} = a \ , \\ \lim\limits_{x\to\infty} f(x) - ax = b \ , \end{cases}$$

где a и b – некоторые конечные числа.

Пример 4.4.2. У функции $y = \dfrac{x^3}{(x-2)^2}$ имеется вертикальная асимптота при $x_0 = 2$ и наклонная асимптота вида $y = x + 4$, в силу

$$\begin{cases} \lim\limits_{x\to\infty} \dfrac{x^3}{(x-2)^2 x} = 1 \ , \\ \lim\limits_{x\to\infty} \left(\dfrac{x^3}{(x-2)^2} - 1 \cdot x \right) = \lim\limits_{x\to\infty} \dfrac{x^3 - (x-2)^2 x}{(x-2)^2} = \\ \qquad = \lim\limits_{x\to\infty} \dfrac{4x^2 - 4x}{(x-2)^2} = \lim\limits_{x\to\infty} \dfrac{4 - \dfrac{4}{x}}{\left(1 - \dfrac{2}{x}\right)^2} = 4 \ . \end{cases}$$

Эскиз графика показан на рис. 4.23.

Иногда функции имеют различные асимптоты при $x \to +\infty$ и $x \to -\infty$. Примером может служить функция $y = \operatorname{arctg} x$, график которой (показанный на рис. 1.9) при $x \to +\infty$ имеет горизонтальную асимптоту с $a = \dfrac{\pi}{2}$, а при $x \to -\infty$ – горизонтальную асимптоту с $a = -\dfrac{\pi}{2}$.

В заключение отметим, что методы исследования функций на экстремальность, выпуклость и эквивалентность будут рассмотрены в §5.4.

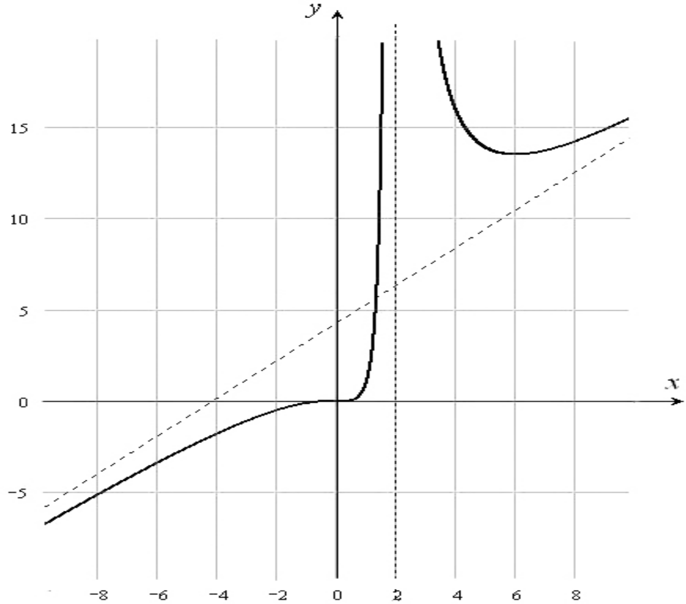


Рис. 4.24. Вертикальная и наклонная асимптоты функции $y = \dfrac{x^3}{(x-2)^2}$.

4.5. Пространственная прямоугольная система координат

Рассмотренный в §4.3.1 метод координатного описания геометрических объектов на плоскости, может быть использован для решения аналогичных задач и в пространстве. Приведем его краткое описание на примере пространственной декартовой прямоугольной системы координат.

Выберем в пространстве некоторую фиксированную точку O, которую будем называть *началом координат.* Тогда (как и в плоском случае) любая точка M (см. рис.4.25) может быть задана ее радиусом-вектором $\overrightarrow{OM}$, то есть направленным отрезком, начало которого находится в точке O, а конец – в точке M.

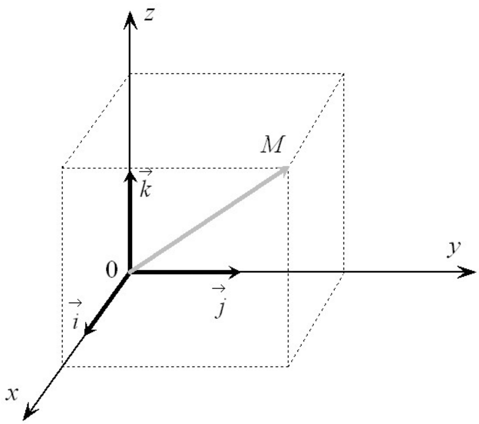


Рис. 4.25. Определение координат точки M в пространстве.

Выберем также три упорядоченных (пронумерованных) вектора $\vec{\imath}$, $\vec{\jmath}$ и $\vec{k}$ таких, что

1) каждый из них *единичной* длины;
2) они взаимно *перпендикулярны*;
3) их начала находятся в точке O.

Тогда *существует единственная* тройка чисел x, y и z таких, что

$$\overrightarrow{OM} = x \cdot \vec{\imath} + y \cdot \vec{\jmath} + z \cdot \vec{k} .$$

Данная формула называется *координатным разложением* вектора в пространстве.

Определение 4.5.1. Совокупность точки O и тройки базисных векторов (ортов) $\vec{\imath}$, $\vec{\jmath}$ и $\vec{k}$, называется *пространственной декартовой прямоугольной системой координат,* а числа x, y и z называются *декартовыми прямоугольными координатами* точки M в системе координат $\left\{O, \vec{\imath}, \vec{\jmath}, \vec{k}\right\}$.

Пространственные декартовы координаты точки (или ее радиуса-вектора) образуют упорядоченную тройку чисел, и потому их удобно записывать в виде трехкомпонентного столбца $\left\|\begin{matrix} x \\ y \\ z \end{matrix}\right\|$, первым элементом которого является x-координата точки M, называемая *абсциссой,* вторым элементом – y-координата, называемая *ординатой.* и третьим элементом – z-координата, называемая *аппликатой.* Этот столбец, по аналогии с плоским случаем, будем называть *координатным представлением* точки или ее радиуса-вектора и использовать сокращенное обозначение координатного представления вектора или точки в виде $\left\|\vec{a}\right\|$ или $\left\|\overrightarrow{OM}\right\|$.

Для пространственного случая будут справедливы все утверждения, приведенные в §4.3.2 и §4.3.3, хотя формула для выражения скалярного произведения в координатах выглядит несколько иначе, чем для векторов на плоскости. Пусть координатные разложения и представления векторов $\vec{a}$ и $\vec{b}$ соответственно имеют вид

$$\vec{a} = \alpha_1 \cdot \vec{\imath} + \alpha_2 \cdot \vec{\jmath} + \alpha_3 \cdot \vec{k}\ ; \quad \left\|\vec{a}\right\| = \left\|\begin{matrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{matrix}\right\| \quad \text{и}$$

$$\vec{b} = \beta_1 \cdot \vec{\imath} + \beta_2 \cdot \vec{\jmath} + \beta_3 \cdot \vec{k}\ ; \quad \left\|\vec{b}\right\| = \left\|\begin{matrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{matrix}\right\| ,$$

тогда скалярное произведение $\vec{a}$ и $\vec{b}$ будет равно

$$\left(\vec{a}, \vec{b}\right) = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3\ . \tag{4.5.1}$$

или

$$\left(\vec{a}, \vec{b}\right) = \left\|\vec{a}\right\|^{\mathrm{T}} \left\|\vec{b}\right\| = \left\|\alpha_1 \ \ \alpha_2 \ \ \alpha_3\right\| \left\|\begin{matrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{matrix}\right\| .$$

Как и плоском случае, из формулы 4°(см. §4.1) вытекает, что длину вектора $\overrightarrow{a}$ можно выразить через его координаты следующим образом

$$\left|\overrightarrow{a}\right| = \sqrt{\left(\overrightarrow{a}\,,\,\overrightarrow{a}\right)} = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}\,. \tag{4.5.2}$$

Эта формула является координатной версией теоремы *Пифагора* в пространстве.

В свою очередь, координаты вектора $\overrightarrow{a}$ выражаются через скалярное произведение по формулам

$$\alpha_1 = \left(\overrightarrow{\imath}\,,\,\overrightarrow{a}\right)\;;$$

$$\alpha_2 = \left(\overrightarrow{\jmath}\,,\,\overrightarrow{a}\right)\;;$$

$$\alpha_3 = \left(\overrightarrow{k}\,,\,\overrightarrow{a}\right)\;.$$

Пример 4.5.1. Для векторов $\overrightarrow{a} = 4\cdot\overrightarrow{\imath} - \overrightarrow{\jmath} + 3\cdot\overrightarrow{k}$ и $\overrightarrow{b} = 3\cdot\overrightarrow{\jmath} + \overrightarrow{k}$ найти их длины и скалярное произведение.

Решение. Координатные представления данных векторов будут иметь вид

$$\left\|\overrightarrow{a}\right\| = \left\|\begin{array}{r} 4 \\ -1 \\ 3 \end{array}\right\| \quad \text{и} \quad \left\|\overrightarrow{b}\right\| = \left\|\begin{array}{c} 0 \\ 3 \\ 1 \end{array}\right\|\,.$$

Тогда их длины находятся по формуле (4.5.2)

$$\left|\overrightarrow{a}\right| = \sqrt{4^2 + (-1)^2 + 3^2} = \sqrt{16 + 1 + 9} = \sqrt{26}\,,$$

$$\left|\overrightarrow{b}\right| = \sqrt{0^2 + 3^2 + 1^2} = \sqrt{0 + 9 + 1} = \sqrt{10}\,.$$

Скалярное произведение данных векторов определяется по формуле (4.5.1)

$$\left(\overrightarrow{a}\,,\overrightarrow{b}\right) = 4\cdot 0 + (-1)\cdot 3 + 3\cdot 1 = 0 + (-3) + 3 = 0\,.$$

Убедитесь самостоятельно, что на основании полученного решения можно утверждать, что векторы $\overrightarrow{a}$ и $\overrightarrow{b}$ образуют в пространстве *прямой угол.*

Одной из возможных областей применения пространственной декартовой системы координат является графическое представление функций, зависящих от *двух переменных*, например, $z = f(x, y)$.

Определение 4.5.2. Графиком функции $z = f(x, y)$ называется множество точек в пространстве, радиусы-векторы которых в прямоугольной системе координат $Oxyz$ имеют координатное представление вида $\left\| \begin{matrix} x \\ y \\ f(x,y) \end{matrix} \right\|$.

Графиком функции двух переменных, как правило, является некоторая поверхность в пространстве, примером которой может служить приведенное на рис.4.26 изображение так называемой *плотности вероятности системы двух случайных величин с нормальным законом распределения.*

Показанная на этом рисунке поверхность конкретно есть график функции

$$z = e^{-(x^2 - xy + y^2)} ,$$

общий вид которой будет рассмотрен в §8.3.3.

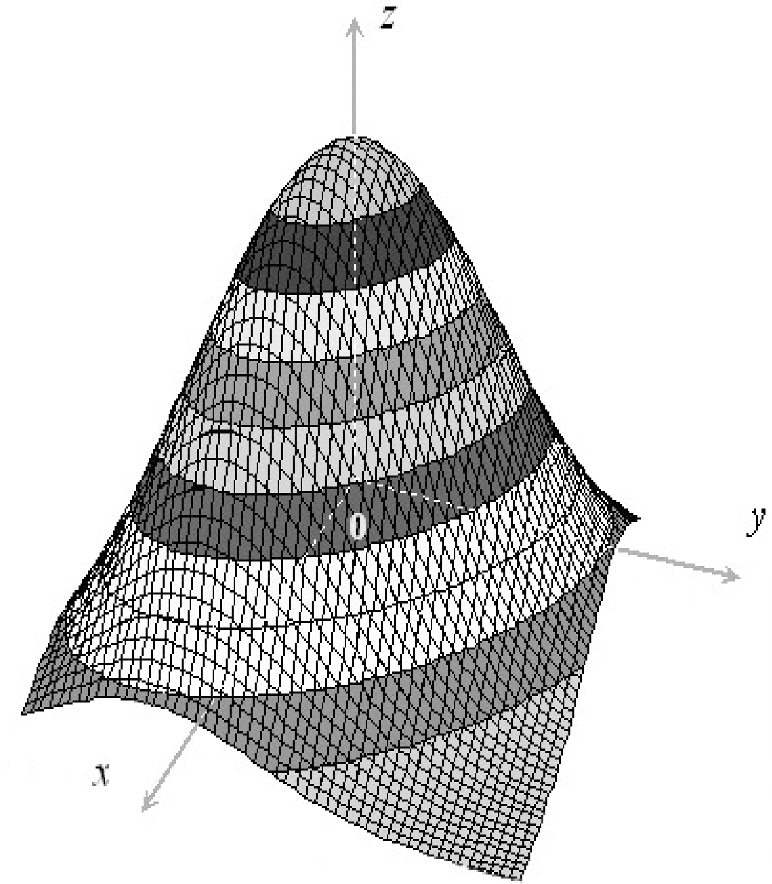


Рис. 4.26. Пример графического представления функции двух переменных.

Глава 5.

Производные и дифференциалы

5.1. Производная функции в точке

Значение функции и ее предел суть локальные числовые характеристики, позволяющие количественно описывать функцию как в некоторой точке, так и в малой ее окрестности. Однако этих характеристик оказывается недостаточно, когда требуется оценить не только сами значения функции, но и относительную скорость их изменения. Для такой оценки используется специальная количественная характеристика функции – *производная функции в точке.* Дадим ее определение.

Определение 5.1.1. *Производной функции в точке* называется число, равное пределу отношения величины приращения значения функции к величине приращения ее аргумента, когда приращение аргумента стремится к нулю.

Символически это определение означает: для функции $y = f(x)$ ее производная в точке x_0, обозначаемая как $f'(x_0)$ или $y'_x(x_0)$, равна

$$f'(x_0) = \lim_{t\to 0} \frac{f(x_0+t) - f(x_0)}{t}\,. \tag{5.1.1}$$

Действительно, если значение аргумента было x_0 , а стало $x_0 + t$, то его приращение очевидно равно $(x_0 + t) - x_0 = t$. Аналогично, если значение функции было $f(x_0)$, а стало $f(x_0+t)$, то ее соответствующее приращение составляет $f(x_0 + t) - f(x_0)$.

Из данного определения следует, что функция $y = f(x)$ должна иметь значения в некоторой окрестности точки x_0 , а также быть *непрерывной* в точке x_0. Последнее условие необходимо (но не достаточно!) для существования производной функции в точке, поскольку лишь для непрерывной функции предел приращения значения функции равен нулю при стремлении к нулю приращения аргумента. Однако, даже для непрерывной функции предел (5.1.1) является неопределенностью вида «$\frac{0}{0}$», то есть заключение о существовании (или не существовании) производной в точке можно делать лишь после «раскрытия» этой неопределенности.

Поясним определение 5.1.1 следующими примерами.

Пример 5.1.1. Найти производную функции $y = x^3$ в точке $x_0 = 2$.

Вначале решим эту задачу для произвольной фиксированной точки x_0 . Пусть приращение аргумента в точке x_0 равно t, найдем соответствующее приращение значения данной функции, использовав формулу из п.2° §1.1,

$$f(x_0+t)-f(x_0) = (x_0+t)^3-x_0^3 = \left(x_0^3 + 3x_0^2t + 3x_0t^2 + t^3\right)-x_0^3 = 3x_0^2t+3x_0t^2+t^3 \ .$$

Тогда

$$f'(x_0) = \lim_{t\to 0} \frac{3x_0^2t + 3x_0t^2 + t^3}{t} = \lim_{t\to 0} \left(3x_0^2 + 3x_0t + t^2\right) = 3x_0^2 \ .$$

Подставив в полученное выражение $x_0 = 2$, найдем, что искомое значение $y'_x(2)$ – производной для функции $y = x^3$ в точке 2, равно 12.

Пример 5.1.2. Найти производную функции $y = |x|$ в точке $x_0 = 0$.

Для данной функции, в силу $x_0 = 0$

$$f'(x_0) = \lim_{t\to 0} \frac{|x_0 + t| - |x_0|}{t} = \lim_{t\to 0} \frac{|t|}{t} = \lim_{t\to 0} \left[\begin{array}{rl} 1 \ , & \text{если } t > 0, \\ -1 \ , & \text{если } t < 0. \end{array} \right. \qquad (5.1.2)$$

Откуда можно сделать заключение, что у рассматриваемой функции нет производной в нуле, так как можно указать две различные числовые последовательности, например, $\left\{t_n = \frac{1}{n}\right\} \underset{n\to\infty}{\longrightarrow} 0$ и

$\left\{\tau_n = -\dfrac{1}{n}\right\} \underset{n\to\infty}{\longrightarrow} 0$ такие, что $f'\left(\dfrac{1}{n}\right) \underset{n\to\infty}{\longrightarrow} 1$ и $f'\left(-\dfrac{1}{n}\right) \underset{n\to\infty}{\longrightarrow} -1$, то есть предел (5.1.2) не существует.

Завершая обсуждение определения 5.1.1 отметим, что в математических текстах используются различные способы обозначения производной функции в точке. Помимо использованных выше, к наиболее часто встречающимся обозначениям относятся

$$y'_x(x)|_{x=x_0} \ ; \qquad \left.\frac{dy}{dx}\right|_{x=x_0} \ ; \quad f'(x)|_{x=x_0} \ .$$

Понятие производной функции в точке допускает *геометрическую интерпретацию*, смысл которой лучше всего иллюстрирует задача построения касательной к графику функции в некоторой его точке.

Допустим, требуется провести касательную к графику функции $y = f(x)$ в точке A с координатами $\left\|\begin{matrix} x_0 \\ y_0 \end{matrix}\right\|$. Выберем на графике другую точку B , имеющую координатное представление $\left\|\begin{matrix} x_1 \\ y_1 \end{matrix}\right\|$. Поскольку обе эти точки (в силу определения 4.4.1) лежат на графике, то справедливы равенства $y_0 = f(x_0)$ и $y_1 = f(x_1)$. Проведем через выбранные точки секущую AB , (см. рис. 5.1.) Нетрудно видеть, что уравнение прямой, проходящей через эти точки, имеет вид $y = k(x-x_0)+y_0$, где значение *углового коэффициента* $k = \dfrac{y_1 - y_0}{x_1 - x_0} = \operatorname{tg}\alpha$.

Начнем теперь, «скользя» по графику, приближать точку B к точке A . Тогда $x_1 \longrightarrow x_0$ и, значит, $t = x_1 - x_0 \longrightarrow 0$. В пределе секущая станет искомой касательной, значение углового коэффициента которой равно

$$k = \lim_{x\to x_0} \frac{y(x) - y(x_0)}{x - x_0} = \lim_{t\to 0} \frac{f(x_0+t) - f(x_0)}{t} = f'(x_0) \ ,$$

где $t = x - x_0$.

Таким образом, мы приходим к заключению, что *значение производной функции в точке равно угловому коэффициенту касательной в этой точке к графику функции.*

В заключение отметим, что из геометрически очевидной горизонтальности касательной к графику функции в ее экстремальных точках, из вышеприведенных рассуждений следует необходимость *равенства нулю значения производной функции в этих точках* (если, разумеется, эта производная существует).

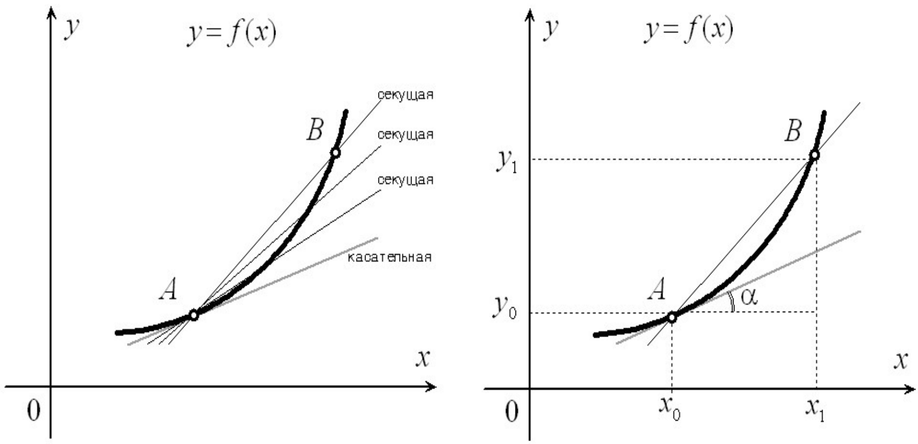


Рис. 5.1. Геометрический смысл производной функции в точке.

5.2. Производная функция

Поиск значений производных функций в точках (непосредственно по определению 5.1.1) в большинстве случаев является сложной задачей, требующей для своего решения значительных затрат вычислительных усилий. На практике оказывается гораздо удобнее применять иной подход, основанный на следующих соображениях.

Используемый в определении 5.1.1 предельный переход осуществляется по вспомогательной переменной t, в то время как, значение x_0 – аргумента функции $f(x)$, для которой ищется производная в точке, является фиксированным числовым параметром. Если изменить x_0, то значение предела (5.1.1), вообще говоря, изменится. Однако для каждого конкретного x_0 оно *одно,* поскольку если предел существует, то он единственен.

Поэтому определение 5.1.1 можно рассматривать как правило, по которому каждому значению x_0, принадлежащему некоторому подмножеству области определения функции $y = f(x)$, ставится в соответствие единственное число $f'(x_0)$, то есть можно сказать, что таким образом задана некоторая новая *функция,* значение которой в точке x_0 равно $f'(x_0)$.

Эту функцию принято называть *производной функцией* от $y = f(x)$ и обозначать как $f'(x)$, операцию же поиска $f'(x)$ называют *дифференцированием.* В случае, когда для $f(x)$ существует $f'(x)$, говорят также, что функция $f(x)$ *дифференцируемая.* С другой стороны, функцию $y = F(x)$ такую, что $F'(x) = f(x)$, называют *первообразной* для $y = f(x)$. Операция ее нахождения называется *интегрированием.*

Например, воспользовавшись решением задачи 5.1.1, можно утверждать, что функция производная от $y = x^3$ есть $y = 3x^2$, поскольку использованный нами метод раскрытия неопределенности годится для *любого* x_0. Теперь для нахождения значения производной функции $y = x^3$ в любой точке x_0 достаточно подставить вместо x значение x_0 в формулу $3x^2$.

Производные функции принято также обозначать как y'_x ; $\dfrac{dy}{dx}$; $y'(x)$. В тех случаях, когда функция зависит более чем от одной переменной, идентификатор переменной, по которой берется производная, указывается явно в виде нижнего индекса. Например, для функции $f(x, p)$ – зависящей от x и p, запись $f'_x(x, p)$ означает производную по переменной x, в то время как p считается фиксированным параметром.

Стоит также отметить, что в русскоязычных математических текстах понятия «производная функции в точке» и «производная функция» часто обозначаются одним и тем же словом – «производная», полагая, что из контекста понятно, о чем идет речь. В английском языке, например, этой проблемы не существует, поскольку производная функции в точке это – derivation, а производная функция – derivative.

Очевидно, что использовать производные функции для нахождения значений производной функции в точке гораздо удобнее, чем пользоваться для этой цели определением 5.1.1. Но тогда возникает вопрос: как находить производные функции?

Ответ следующий – надо использовать:

1) таблицу производных для некоторого небольшого набора элементарных функций, полученных непосредственно при помощи определения 5.1.1, и
2) правила дифференцирования, которые позволяют выражать производные одних функций через производные других. Обоснование их справедливости, основанное на использовании определения 5.1.1 и свойств пределов функций, можно найти в полном курсе математического анализа.

$f(x)$	$f'(x)$
x^a	ax^{a-1}
e^x	e^x
$a^x, \quad a>0, a\neq 1$	$a^x \ln a$
$\ln\lvert x\rvert$	$\frac{1}{x}$
$\log_a \lvert x\rvert, \quad a>0, a\neq 1$	$\frac{1}{x\ln a}$
$\sin x$	$\cos x$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\operatorname{arctg} x$	$\frac{1}{1+x^2}$

(Таблица 5.2.1)

Первая из этих таблиц имеет вид (5.2.1), но поскольку очевидно, что этой таблицы недостаточно, чтобы получать производные *любых* функций, задаваемых формулами, то следует использовать также и таблицу, описывающую *правила выражения производных одних функций, через производные других,*

1°	$(\ f(x)+g(x)\ )' = f'(x)+g'(x)$
2°	$(\ C\cdot f(x)\ )' = C\cdot f'(x)$, где C – const
3°	$(\ f(x)\cdot g(x)\ )' = f'(x)\cdot g(x) + f(x)\cdot g'(x)$
4°	$(\ f(g(x))\ )'_x = f'_g(g)\cdot g'_x(x)$

(Таблица 5.2.2)

Проиллюстрируем использование таблиц 5.2.1 и 5.2.2 на следующих примерах.

Пример 5.2.1. Пусть требуется найти производные функции от

$$y = \frac{\sqrt[3]{x^2}+\sqrt{x}+1}{x}, \quad y = \frac{\sin x}{x} \quad \text{и} \quad y = e^{\arccos x}.$$

Решение.

1) Сравнение правил 1° и 3° таблицы 5.2.2 убеждает, что легче дифференцировать сумму функций, чем произведение, поэтому вначале выполним почленное деление, то есть будем искать производную функции

$$\left(\frac{\sqrt[3]{x^2}+\sqrt{x}+1}{x}\right)' = \left(x^{-\frac{1}{3}} + x^{-\frac{1}{2}} + x^{-1}\right)' =$$

а, используя первые формулы таблиц 5.2.1 и 5.2.2, получаем

$$= -\frac{1}{3}\cdot x^{-\frac{4}{3}} - \frac{1}{2}\cdot x^{-\frac{3}{2}} - x^{-2} = -\frac{1}{3x\sqrt[3]{x}} - \frac{1}{2x\sqrt{x}} - \frac{1}{x^2}.$$

2) В таблице 5.2.2 нет формулы для дифференцирования дроби (хотя во многих учебниках Вы ее можете встретить), поэтому сначала преобразуем данную функцию в произведение, и лишь

потом применим правило 3° таблицы 5.2.2

$$\left(\frac{\sin x}{x}\right)' = \big((\sin x)\cdot(x^{-1})\big)' = (\sin x)'\cdot(x^{-1}) + (\sin x)\cdot(x^{-1})' =$$

$$= (\cos x)\cdot x^{-1} + (\sin x)\cdot(-x^{-2}) = \frac{\cos x}{x} - \frac{\sin x}{x^2} = \frac{x\cos x - \sin x}{x^2}.$$

3) Вначале заметим, что (см. §1.1 п.7°)

$$g(x) = \arccos x = \frac{\pi}{2} - \arcsin x\ ,$$

и потому, согласно первому правилу таблицы 5.2.2 и предпоследней строке таблицы 5.2.1,

$$g'_x(x) = -\frac{1}{\sqrt{1-x^2}}.$$

Тогда,

$$y'_x = (e^{\arccos x})'_x = \left(e^{g(x)}\right)'_x = (e^g)'_g \cdot g'_x(x) =$$

$$= e^g \cdot g'_x(x) = e^g \cdot \left(-\frac{1}{\sqrt{1-x^2}}\right) = -\frac{e^{\arccos x}}{\sqrt{1-x^2}}.$$

В заключение данного параграфа отметим, что в процессе исследования свойств функции иногда оказывается необходимым использование характеристики, показывающей *скорость изменения величины производной функции* в окрестности точки x_0. Эта характеристика называется *производной второго порядка* или, просто, *второй производной.* Ее принято обозначать как

$$y >>_{x=x_0}\ ; \quad \left.\frac{d^2y}{dx^2}\right|_{x=x_0}\ ; \quad y >> (x)|_{x=x_0}\ ,$$

а поиск значения этой числовой характеристики сводится к вычислению предела вида

$$f >> (x_0) = \lim_{t\to 0}\frac{f'(x_0+t) - f'(x_0)}{t}. \tag{5.2.1}$$

Поскольку вторые производные функции в точке (как пределы функции) определены однозначно, то можно задать новую функцию, значениями которой являются числа, получаемые по формуле (5.2.1). Эту функцию (производную от производной) называют *второй производной функцией от функции* $y = f(x)$. Для ее нахождения следует использовать те же правила, что и для первой производной функции. Например, если $y = \ln|x|$, то, согласно таблице 5.2.1, $y' = \frac{1}{x}$, а, в свою очередь, производная функция от $\frac{1}{x} = x^{-1}$ по той же таблице 5.2.1 равна $(-1)x^{-2} = -\frac{1}{x^2}$. То есть $(\ln|x|) >>= -\frac{1}{x^2}$.

Геометрический смысл второй производной будет разъяснен в §5.4.

5.3. Формула Тейлора. Дифференциалы

Рассмотрим теперь методы исследования функции в малой окрестности некоторой точки, основанные на использовании значений ее производных.

Пусть точка x_0 для функции $y = f(x)$ – внутренняя точка области определения, то есть существует окрестность x_0 целиком содержащаяся в области определения этой функции. И пусть в этой точке функция $y = f(x)$ имеет *непрерывные производные первого и второго порядка.*

В ситуации, когда полное аналитическое описание (то есть в виде формулы) функции $y = f(x)$ слишком сложно или же неизвестно вовсе, а нас интересует ее поведение лишь в относительно небольшой окрестности точки x_0, представляется целесообразным использование приближенного описания этой функции в виде линейной комбинации степенных функций только первого и второго порядков

$$y = a + b(x - x_0) + c(x - x_0)^2 + r \ , \tag{5.3.1}$$

где a, b и c – некоторые константы, а r – остаточная функция, значением которой можно пренебречь при малых величинах $|x - x_0|$.

Поскольку эта аппроксимация будет «простой» при любых значениях параметров a, b и c, то имеет смысл выбирать их значения так, чтобы погрешность, возникающая при игнорировании (отбрасывании) остатка r, была минимальной. Для достижения этой цели поступим следующим образом.

Потребуем прежде всего, чтобы значение остаточной функции было равно нулю при $x = x_0$. Если записать соотношение (5.3.1) в виде

$$f(x) = a + b(x - x_0) + c(x - x_0)^2 + r \ , \tag{5.3.2}$$

то при $x = x_0$ оно превратится в равенство $f(x_0) = a + r$, из которого следует, что для обеспечения в этой точке условия $r = 0$ нужно взять $a = f(x_0)$.

Далее, потребуем, чтобы первая производная остаточного члена также равнялась бы нулю при $x = x_0$. Если функции равны, то очевидно равны и их производные, потому будет верным равенство, получаемое из (5.3.2) почленным его дифференцированием

$$f'(x) = b + 2c(x - x_0) + r' \ . \tag{5.3.3}$$

Откуда при $x = x_0$ получаем, что $r' = 0$ обеспечивается при $b = f'(x_0)$.

Наконец потребуем, чтобы и $r >>= 0$ при $x = x_0$. Дифференцируя обе части равенства (5.3.3), получим

$$f >> (x) = 2c + r >> \ ,$$

из которого при $x = x_0$ следует, что для обеспечения $r >>= 0$ необходимо, чтобы $c = \frac{1}{2} f >> (x_0)$.

В итоге приходим к заключению, что «наилучшая» аппроксимация исходной функции $y = f(x)$, носящая название *формулы Тейлора,* имеет вид

$$y = f(x_0) + f'(x_0)(x - x_0) + \frac{f >> (x_0)}{2}(x - x_0)^2 + r \ . \tag{5.3.4}$$

Для упрощения описания и использования данной аппроксимационной процедуры, произведение $f'(x_0)(x - x_0)$ принято называть *первым дифференциалом* функции $y = f(x)$ в точке x_0, и обозначать dy, df или $df(x_0)$, а произведение $f >> (x_0)(x - x_0)^2$ – *вторым дифференциалом,* для которого используются обозначения d^2y, d^2f или $d^2f(x_0)$.

Кроме того, по определению будем считать, что дифференциал независимой переменной равен ее приращению, то есть $dx = \Delta x = x - x_0$. В этом случае формула (5.3.4) записывается в виде

$$y = f(x_0) + f'(x_0)dx + \frac{f >> (x_0)}{2}(dx)^2 + r \tag{5.3.5}$$

или

$$y = f(x_0) + df + \frac{1}{2}d^2 f + r\ .$$

Откуда следует, что значение первого дифференциала дает оценку приращения этой функции в окрестности точки x_0 с погрешностью порядка малости $(x - x_0)^2$.

Свойства дифференциалов определяются свойствами производных и основные из них приведены в таблице 5.3.1.

1°	$d\,(f(x) + g(x)) = df + dg$
2°	$d\,(C \cdot f(x)) = C \cdot df$, где C – const
3°	$d\,(f(x) \cdot g(x)) = df \cdot g(x) + f(x) \cdot dg$
4°	$df(g(x)) = f'_g\ dg = f'_g(g(x)) \cdot g'_x\ dx$

(Таблица 5.3.1)

Для дальнейшего также важно понимать, что дифференциал функции одной переменной $y = f(x)$ является функцией, зависящей от *двух* переменных: от x и от x_0 . Или, что чаще встречается на практике, от x_0 и от dx , где $dx = x - x_0$.

5.4. Исследование свойств функций

Теперь используем построенную аппроксимацию, чтобы выяснить при каких условиях функция $y = f(x)$ имеет при $x = x_0$ *локальный экстремум,* то есть *максимум или минимум.*

Из формулы (5.3.4) следует, что в случае $f'(x_0) \neq 0$ экстремума быть не может. Действительно, выберем $x = x_0 + t$ так, чтобы слагаемое

$$f'(x_0)(x - x_0) = f'(x_0) \cdot t > 0.$$

Тогда $f(x_0 + t) > f(x_0)$, но очевидно будет верным также и $f(x_0 - t) < f(x_0)$. То есть в этом случае экстремума нет и мы получаем, что равенство $f'(x_0) = 0$ является *необходимым условием экстремума* при $x = x_0$.

Пусть теперь $f'(x_0) = 0$. Это означает, что аппроксимация функции $y = f(x)$ в малой окрестности точки x_0 имеет вид

$$y = f(x_0) + \frac{f >> (x_0)}{2}(x - x_0)^2 + r \ . \tag{5.4.1}$$

Поскольку множитель $(x - x_0)^2$ неотрицателен, то для любых x в малой окрестности точки x_0 $f(x) > f(x_0)$, если $f >> (x_0) > 0$, и $f(x) < f(x_0)$, если $f >> (x_0) < 0$. Откуда заключаем, что $f >> (x_0) > 0$ – *достаточное условие локального минимума* при $x = x_0$, а $f >> (x_0) < 0$ – *достаточное условие локального максимума.*

Исследуя функции на экстремальность, нужно иметь в виду, что при одновременном выполнении условий $f'(x_0) = 0$ и $f >> (x_0) = 0$ функция $y = f(x)$ в точке x_0 может как иметь экстремум, так и не иметь его. Проверьте самостоятельно, что например у функций $y = x^3$ и $y = x^4$ при $x = 0$ производные первого и второго порядков равны нулю, однако в этой точке вторая из них имеет минимум, а у первой экстремума нет.

Отметим еще одно важное свойство аппроксимации (5.3.4). Если в этой формуле остаточное слагаемое $r(x, x_0)$ рассматривать как функцию, зависящую от значений x и x_0, то будет верно следующее равенство

$$\lim_{x \to x_0} \frac{r(x, x_0)}{(x - x_0)^2} = 0 \ ,$$

которое означает, что при стремлении x к x_0 величина остаточного слагаемого $r(x, x_0)$ убывает быстрее, чем $(x - x_0)^2$.

По схеме, аналогичной исследованию функций на экстремум, можно изучать и другие их локальные свойства.

Выясним, например, геометрический смысл второй производной функции в точке, используя формулы (5.1.1) и (5.1.2) для приближенной оценки значений производных функций при малых по модулю t.

$$f >> (x_0) \approx \frac{f'(x_0 + t) - f'(x_0)}{t} \approx$$

$$\approx \frac{\dfrac{f(x_0 + 2t) - f(x_0 + t)}{t} - \dfrac{f(x_0 + t) - f(x_0)}{t}}{t} =$$

$$= \frac{f(x_0+2t)-2f(x_0+t)+f(x_0)}{t^2}\,.$$

Принимая во внимание, что (в силу определения 4.4.2) неравенство

$$f(x_0+t) \le \frac{f(x_0+2t)+f(x_0)}{2}$$

означает *выпуклость вниз* функции $y = f(x)$ в малой окрестности точки $x_0 + t$, можно придти к заключению, что условие $f >> (x_0) > 0$ является для этого достаточным. Совершенно аналогично получаем, что выполнения неравенства $f >> (x_0) < 0$ достаточно, чтобы гарантировать *выпуклость вверх* функции $y = f(x)$ в малой окрестности точки x_0. Таким образом, в случае когда пределы (5.1.1) и (5.2.1) существуют, можно делать заключение о направлении локальной выпуклости исследуемой функции по знаку ее второй производной.

В заключение, продемонстрируем применение локальных аппроксимаций для нахождения пределов функций в случае возникновения неопределенности вида «$\frac{0}{0}$».

Пусть функции $f(x)$ и $g(x)$ имеют непрерывные первые производные в точке $x = x_0$. И пусть, кроме того, $f(x_0) = g(x_0) = 0$, и $g'(x) \ne 0$. Тогда оказывается справедливым *правило Лопиталя*

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}\,.$$

Действительно, для данных функций в малой окрестности точки x_0 будут справедливы следующие аппроксимации

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + r\ , \qquad g(x) = g(x_0) + g'(x_0)(x-x_0) + q\ ,$$

причем для остаточных слагаемых r и q будут выполняться предельные равенства

$$\lim_{x\to x_0} \frac{r}{x-x_0} = 0\ , \qquad \lim_{x\to x_0} \frac{q}{x-x_0} = 0\ .$$

Тогда, приняв во внимание, что $f(x_0) = g(x_0) = 0$, получим

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x_0)(x-x_0)+r}{g'(x_0)(x-x_0)+q} = \lim_{x\to x_0} \frac{f'(x_0) + \dfrac{r}{x-x_0}}{g'(x_0) + \dfrac{q}{x-x_0}} =$$

$$= \frac{f'(x_0) + \lim\limits_{x \to x_0} \dfrac{r}{x - x_0}}{g'(x_0) + \lim\limits_{x \to x_0} \dfrac{q}{x - x_0}} = \frac{f'(x_0)}{g'(x_0)} \,.$$

Проиллюстрируем применение правила Лопиталя на следующем примере.

Пример 5.4.1. Найти $\lim\limits_{x \to 3} \dfrac{3^x - 27}{x - 3}$.

Решение. В данном случае мы имеем неопределенность вида «$\frac{0}{0}$», а, поскольку $(3^x - 27)' = 3^x \ln 3$ и $(x - 3)' = 1 \neq 0$, то согласно правилу Лопиталя

$$\lim_{x \to 3} \frac{3^x - 27}{x - 3} = \lim_{x \to 3} \frac{3^x \ln 3}{1} = 27 \ln 3 \,.$$

5.5. Построение графиков функций

В процессе построения графика функции $y = f(x)$ целесообразно (хотя и вовсе необязательно!) придерживаться следующей последовательности действий.

1°. Нахождение области определения.

2°. Поиск области значений.

3°. Исследование вопроса: обладает ли функция $y = f(x)$ свойствами четности и периодичности.

4°. Исследование асимптотического поведения данной функции (то есть выяснение характера ее изменения при $x \to \pm\infty$, а также поиск вертикальных асимптот графика функции.)

5°. Поиск промежутков знакопостоянства, нулей функции, точек пересечения ее графика с осями координат и точек разрыва.

6°. Определение промежутков монотонности, точек локальных экстремумов (максимумов и минимумов) (см. §5.4).

7°. Установление промежутков выпуклости (вверх или вниз) и поиск точек перегиба графика (см. §5.4).

8°. Сведение полученной информации в итоговую таблицу и построение эскиза графика функции.

Продемонстрируем применение данной схемы на примере построения графика функции $y = x^3 - 2x + 1$.

1°. *Область определения*: все операции, использованные для записи формулы функции, выполнимы для любых действительных x. Значит $X : x \in (-\infty, +\infty)$.

2°. *Область значений*: в силу того, что уравнение $x^3 - 2x + 1 = p$ имеет вещественные решения для любого вещественного значения параметра p, (факт, обоснование которого выходит за рамки данного курса) можно утверждать, что областью значений является множество всех вещественных чисел.

3°. Данная функция $y = f(x)$ свойствами *четности, нечетности и периодичности* не обладает.

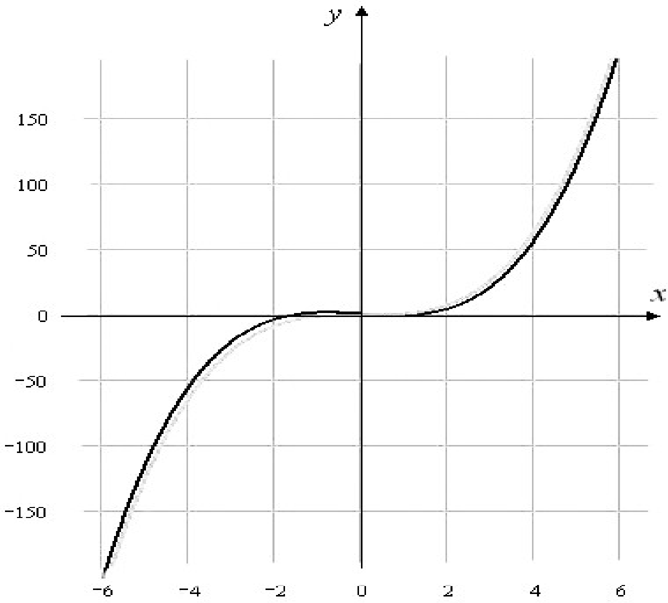


Рис. 5.2. Графики функций $y = x^3 - 2x + 1$ (черный) и $y = x^3$ (серый)

4°. Поскольку

$$\lim_{x\to\infty} \frac{x^3 - 2x + 1}{x^3} = \lim_{x\to\infty} \left(1 - \frac{2}{x^2} + \frac{1}{x^3}\right) = 1 \ ,$$

асимптотически (то есть при больших по модулю значениях x) данная функция ведет себя подобно функции $y = x^3$. См. рис. 5.2.

5°. *Промежутки знакопостоянства*: преобразуем формулу функции

$$y = x^3 - 2x + 1 = (x^3 - x) - (x - 1) = (x - 1)(x^2 + x - 1) =$$

$$= (x - x_1)(x - x_2)(x - 1) \ ,$$

где

$$x_1 = -\frac{1}{2} - \frac{\sqrt{5}}{2} \approx -1.6 \qquad x_2 = -\frac{1}{2} + \frac{\sqrt{5}}{2} \approx 0.6 \ .$$

Используя «метод интервалов», приходим к заключению, что

$$y \geq 0, \text{ если } \left[\begin{array}{c} x_1 \leq x \leq x_2 \\ x \geq 1 \end{array} \right. \qquad \text{и} \qquad y < 0, \text{ если } \left[\begin{array}{r} x < x_1 \\ x_2 < x < 1 \ . \end{array} \right.$$

6°. *Промежутки монотонности*: согласно формуле (5.3.4) значение функции возрастает в окрестности точки, где производная положительна, и соответственно убывает там, где производная отрицательна. Поэтому для определения промежутков возрастания или убывания значений функции необходимо найти ее производную функцию и исследовать ее «на знак». В нашем случае $y' = 3x^2 - 2$, поэтому можно утверждать, что

$$y' \geq 0, \text{ если } -\sqrt{\frac{2}{3}} \leq x \text{ , либо } x \geq \sqrt{\frac{2}{3}}$$

и

$$y' < 0, \text{ если } -\sqrt{\frac{2}{3}} < x < \sqrt{\frac{2}{3}} \, .$$

7°. *Направление выпуклости*: точки $x_3 = -\sqrt{\dfrac{2}{3}}$ и $x_4 = \sqrt{\dfrac{2}{3}}$ – стационарные, в них первая производная обращается в нуль. Факт наличия в этих точках экстремума можно установить по знаку второй производной, которая в рассматриваемом случае равна $y >>= 6x$. Это означает, что при положительных x график функции имеет выпуклость вниз, при отрицательных – выпуклость вверх, а при $x = 0$ направление выпуклости меняется на противоположное, то есть это – *точка перегиба* графика функции. Кроме того, отсюда следует, что в точке $x_3 = -\sqrt{\dfrac{2}{3}}$

функция имеет локальный максимум, а в точке $x_4 = \sqrt{\dfrac{2}{3}}$ – локальный минимум.

8°. *Сводная таблица свойств функции*:

x	$-\infty$	$-$	x_1	$-$	x_3	$-$	0	$+$	x_2	$+$	x_4	$+$	1	$+$
y	$-\infty$	$-$	0	$+$	$+$	$+$	1	$+$	0	$-$	$-$	$-$	0	$+$
y'	$+\infty$	$\nearrow$	$\nearrow$	$\nearrow$	0	$\searrow$	$\searrow$	$\searrow$	$\searrow$	$\searrow$	0	$\nearrow$	$\nearrow$	$\nearrow$
$y >>$	$-\infty$	$\cap$	$\cap$	$\cap$	$\cap$	$\cap$	0	$\cup$	$\cup$	$\cup$	$\cup$	$\cup$	$\cup$	$\cup$
					max		пер.				min			

Детализированный вид графика показан на рис. 5.3.

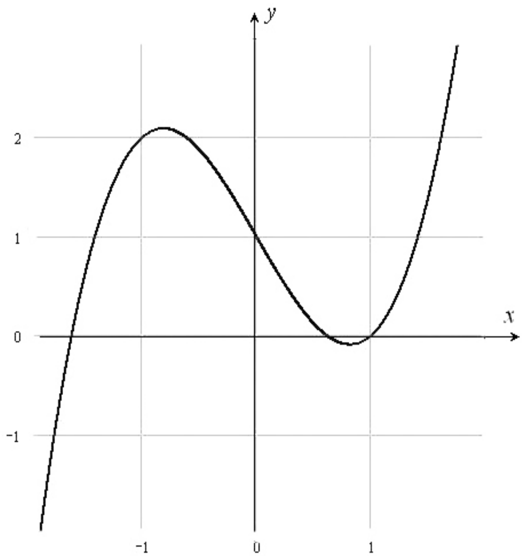


Рис. 5.3. График функции $y = x^3 - 2x + 1$

Исследуем функцию $y = (2x+5)e^x$ и построим ее график.

1°. *Область определения*: поскольку операции, использованные для записи формулы функции, выполнимы для любых действительных x, то $X : x \in (-\infty, +\infty)$.

2°. *Область значений*: очевидно, что при $x \geq 0$ значение y неограничено велико. Если $x \leq 0$, то в силу $\lim\limits_{x \to -\infty} (2x+5)e^x = 0$ и непрерывности данной функции для любых x, приходим к заключению о ее «ограниченности снизу». При этом нижнюю границу области значений удобнее будет найти из условия стационарности, поскольку исследуемая функция не только непрерывна, но и имеет производную для любого значения x.

3°. Функция $y = (2x+5)e^x$ не обладает свойствами *четности, нечетности или периодичности.*

4°. Поскольку $\lim\limits_{x \to -\infty} (2x+5)e^x = 0$, то график данной функции асимптотически приближается к оси Ox при x стремящемся к $-\infty$.

5°. *Промежутки знакопостоянства*: поскольку $e^x > 0 \; \forall x$, то $y = 0$ при $x = -\dfrac{5}{2}$ и

$$y < 0, \text{ если } -\infty < x < -\frac{5}{2} \quad \text{и} \quad y > 0, \text{если } -\frac{5}{2} < x < +\infty$$

6°. *Промежутки монотонности*: в рассматриваемом случае

$$y' = 2e^x + (2x+5)e^x = (2x+7)e^x \ ,$$

поэтому можно утверждать, что

$$y' = 0, \text{ если } x = -\frac{7}{2}$$

и

$$y' < 0, \text{ если } -\infty < x \leq -\frac{7}{2} \quad \text{и} \quad y' > 0, \text{ если } -\frac{7}{2} < x < +\infty \ .$$

7°. *Направление выпуклости*: можно установить по знаку второй производной, которая в рассматриваемом случае равна $y >>= (2x + 9)e^x$. То есть при $-\infty < x < -\dfrac{9}{2}$ график функции имеет выпуклость вверх, а при $-\dfrac{9}{2} < x < +\infty$ – выпуклость вниз. Отсюда следует, что в точке $x = -\dfrac{7}{2}$ функция имеет локальный минимум, а в точке $x = -\dfrac{9}{2}$ – точку перегиба графика функции.

8°. *Сводная таблица свойств функции*:

x	$-\infty$	−	$-\dfrac{9}{2}$	−	$-\dfrac{7}{2}$	−	$-\dfrac{5}{2}$	+	0	+	$+\infty$
y	–0	−	−	−	−	−	0	+	5	+	$+\infty$
y'	–0	↘	↘	↘	0	↗	↗	↗	↗	↗	$+\infty$
$y >>$	–0	$\cap$	0	$\cup$	$\cup$	$\cup$	$\cup$	$\cup$	$\cup$	$\cup$	$+\infty$
			перегиб		min						

Вид графика функции $y = (2x+5)e^x$ показан на рис. 5.4.

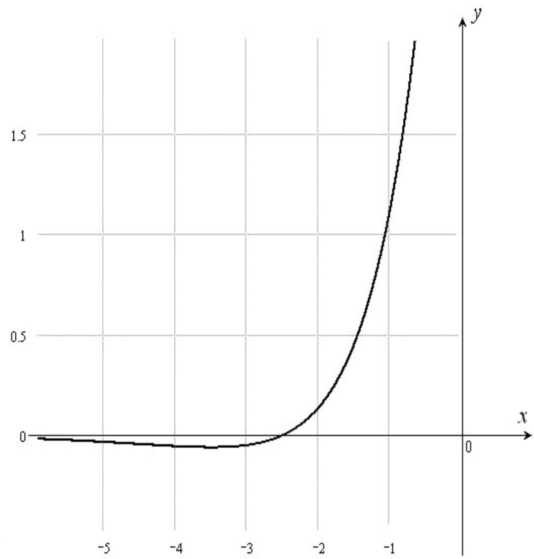


Рис. 5.4. График функции $y = (2x+5)e^x$

В заключение выполним исследование функции $y = \dfrac{x^3+6x-2}{x^2}$.

1°. *Область определения*: любые действительные $x \neq 0$.

2°. *Область значений*: любые действительные y.

3°. Данная функция не обладает свойствами *четности, нечетности или периодичности.*

4°. График функции имеет вертикальную асимптоту при $x \to 0$, поскольку $\lim\limits_{x\to 0} \dfrac{x^3+6x-2}{x^2} = -\infty$. Кроме того, переписав формулу задающую исследуемую функцию в виде $y = x + \dfrac{6}{x} - \dfrac{2}{x^2}$, можно заметить, что эта функция имеет наклонную асимптоту вида $y = x$.

5°. *Промежутки знакопостоянства*: из формулы $y = x + \dfrac{6}{x} - \dfrac{2}{x^2}$, очевидно, что $y < 0$ при любых $x < 0$. В области $x > 0$ знак y совпадает со знаком трехчлена x^3+6x-2, который отрицателен при $x < \sqrt[3]{4} - \sqrt[3]{2}$ и положителен при $x > \sqrt[3]{4} - \sqrt[3]{2}$.

Это утверждение явно не очевидно, и потому приведем его обоснование, использовав «школьные методы», хотя проще было бы применить какой-нибудь компьютерный инструмент, например, MathCAD или MAPPLE.

Итак, надо найти вещественные решения уравнения $x^3+6x-2 = 0$. Будем искать его корни в виде суммы двух новых неизвестных $x = u + v$. Подставив их в уравнение, получим

$$(u+v)^3+6(u+v)-2 = 0 \quad \Longrightarrow \quad \big(u^3+v^3+3uv(u+v)\big)+6(u+v)-2 = 0$$

$$\Longrightarrow \quad \big(u^3+v^3-2\big)+(u+v)(3uv+6) = 0 \ .$$

Откуда получаем: для того чтобы x было решением данного уравнения необходимо, чтобы значения неизвестных u и v удовлетворяли системе уравнений

$$\begin{cases} u^3+v^3 = 2, \\ uv = -2, \end{cases}$$

которая в свою очередь равносильна системе

$$\begin{cases} v = -\dfrac{2}{u}, \\ u^3 - \dfrac{8}{u^3} = 2. \end{cases}$$

Введем новую неизвестную $t = u^3$, тогда второе уравнение последней системы сводится к квадратному и легко решается:

$$t - \frac{8}{t} - 2 = 0 \quad \Longrightarrow \quad t^2-2t-8 = 0 \quad \Longrightarrow \quad \left[\begin{array}{l} t_1 = 4, \\ t_2 = -2 \end{array} \right. \quad \Longrightarrow$$

$$\Longrightarrow \quad \left[\begin{array}{l} u_1 = \sqrt[3]{4}, \\ u_2 = -\sqrt[3]{2} \end{array}\right. \quad \Longrightarrow \quad \left[\begin{array}{l} v_1 = -\sqrt[3]{2}, \\ v_2 = \sqrt[3]{4} \end{array}\right. \quad \Longrightarrow \quad x_3 = u + v = \sqrt[3]{4} - \sqrt[3]{2}\ .$$

6°. *Промежутки монотонности*: для данной функции

$$y' = \frac{x^3 - 6x + 4}{x^3} = \frac{(x-2)(x^2+2x-2)}{x^3} = \frac{(x-2)(x+1-\sqrt{3})(x+1+\sqrt{3})}{x^3}\ ,$$

поэтому можно утверждать, что

$$y' = 0, \text{ если } \left[\begin{array}{l} x = 2, \\ x = -1+\sqrt{3}, \\ x = -1-\sqrt{3} \end{array}\right.$$

и, согласно «методу интервалов»,

$$y' > 0, \text{ если } x \in (-\infty, -1-\sqrt{3}) \bigcup (0, -1+\sqrt{3}) \bigcup (2, +\infty)$$

$$\text{и} \quad y' < 0, \text{ если } (-1-\sqrt{3}, 0) \bigcup (-1+\sqrt{3}, 2)\ .$$

7°. *Направление выпуклости*: вторая производная в рассматриваемом случае равна $y >>= \dfrac{12(x-1)}{x^4}$. Поэтому при $x < 1$ график функции имеет выпуклость вверх, а при $x > 1$ – выпуклость вниз. Отсюда следует, что в точке $x = 2$ функция имеет локальный минимум, в точках $x_1 = -1-\sqrt{3}$ и $x_2 = -1+\sqrt{3}$ – локальный минимум, а в точке $x = 1$ – точку перегиба графика функции.

8°. *Сводная таблица свойств функции*:

x	$-\infty$	$-$	x_1	$-$	0	$+$	x_3	$+$	x_2	$+$	1	$+$	2	$+$
y	$-\infty$	$-$	$-$	$-$	$-\infty$	$-$	0	$+$	$+$	$+$	$+$	$+$	$+$	$+$
y'	1	$\nearrow$	0	$\searrow$	$\mp\infty$	$\nearrow$	$\nearrow$	$\nearrow$	0	$\searrow$	$\searrow$	$\searrow$	0	$\nearrow$
$y >>$	-0	$\bigcap$	$\bigcap$	$\bigcap$	$-\infty$	$\bigcap$	$\bigcap$	$\bigcap$	$\bigcap$	$\bigcap$	0	$\bigcup$	$\bigcup$	$\bigcup$
			max		в.ас.				max		пер.		min	

Вид графика функции $y = \dfrac{x^3+6x-2}{x^2}$ показан на рис. 5.5.

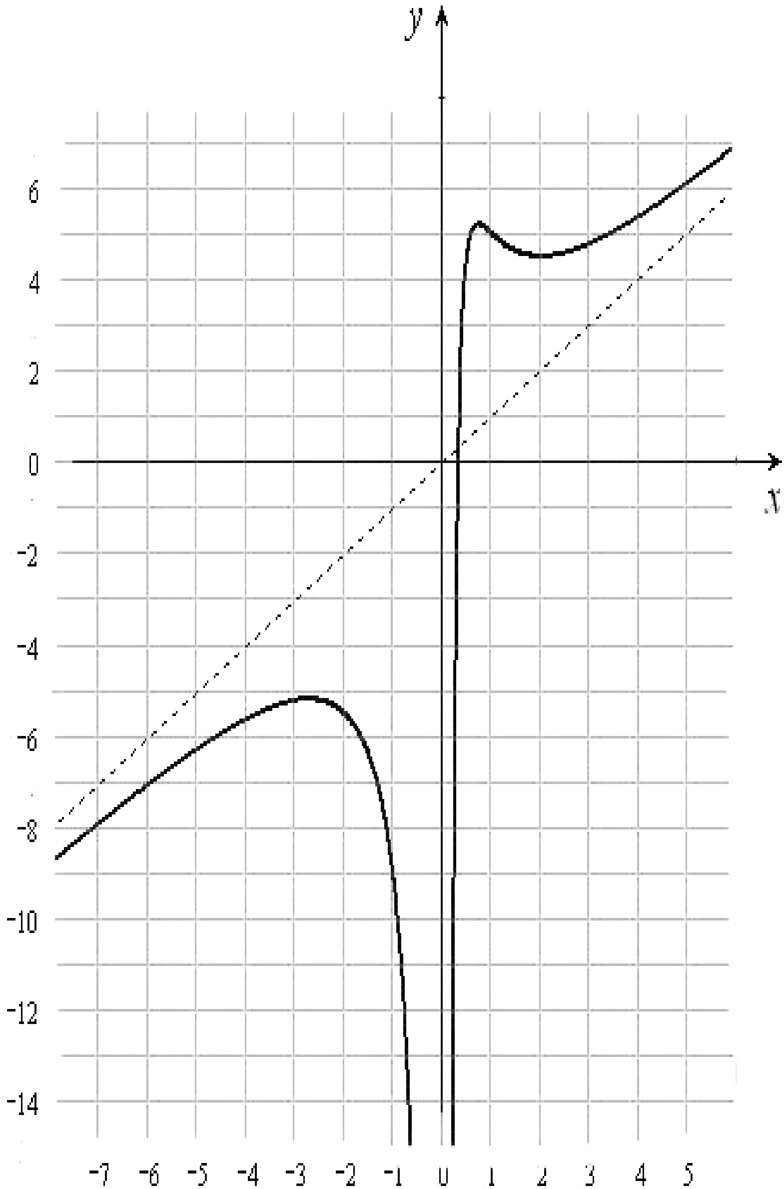


Рис. 5.5. График функции $y = \dfrac{x^3 + 6x - 2}{x^2}$

Глава 6.

Интегралы и ряды

6.1. Определенный интеграл

6.1.1. Интегрирование функций. Определенный интеграл как предел последовательности

В предыдущих главах рассматривались методы исследования функций при помощи их локальных (то есть относящихся к некоторой небольшой окрестности фиксированной точки) количественных характеристик. Однако, достаточно часто возникают задачи исследования свойств функций, относящихся к *не малым* промежуткам области определения. Их характерным примером является *интегрирование* – задача нахождения значения первообразной функции в некоторой точке b по ее значению в точке a и известной производной функции, в случае, когда a и b не находятся в малых окрестностях друг друга.

Рассмотрим решение этой задачи. Пусть заданная функция $f(x)$ является производной функций от некоторой первообразной $F(x)$, для которой известно лишь ее значение $F(a)$, и требуется найти $F(b)$ – значение неизвестной нам первообразной в точке b.

Разобьем отрезок $[a,b]$ на N равных частей точками $x_1, x_2, \ldots x_{N-1}$. Кроме того, будем считать, что $x_0 = a$ и $x_N = b$. Тогда

$$\Delta x_k = x_{k+1} - x_k = \frac{b-a}{N}$$

для любого $k = 0, 1, 2, \dots N-1$. Заметим, что числа $k = 0, 1, 2, \dots N-1$ служат не только номерами точек разбиения, но также и номерами *отрезков,* на которые мы разбили $[a, b]$.

Приращение значения функции $F(x)$ на отрезке $[x_k, x_{k+1}]$ равно

$$\Delta F_k = F(x_{k+1}) - F(x_k) \ ,$$

тогда, прибавляя и отнимая одинаковые числа, после группировки соседних слагаемых, получаем

$$F(b) - F(a) = F(x_N) - F(x_0) =$$

$$= \underbrace{F(x_N) - F(x_{N-1})}_{\Delta F_{N-1}} + \underbrace{F(x_{N-1}) - F(x_{N-2})}_{\Delta F_{N-2}} + \cdots$$

$$\cdots + \underbrace{F(x_2) - F(x_1)}_{\Delta F_1} + \underbrace{F(x_1) - F(x_0)}_{\Delta F_0} =$$

$$= \sum_{k=0}^{N-1} \Delta F_k \ .$$

Откуда искомое значение $F(b) = F(a) + \sum\limits_{k=0}^{N-1} \Delta F_k$.

Таким образом рассматриваемая задача свелась к нахождению $\sum\limits_{k=0}^{N-1} \Delta F_k$. Поскольку по условию задачи $F'(x) = f(x)$, то, заменяя приращение функции ее *дифференциалом* (см. §5.3) по формуле

$$\Delta F = F(x) - F(x_0) \approx dF(x_0) = F'(x_0) \cdot (x - x_0) = F'(x_0) \cdot \Delta x,$$

получим приближенное равенство

$$\Delta F_k \approx dF(x_k) = F'(x_k)(x_{k+1} - x_k) = f(x_k)(x_{k+1} - x_k) = f(x_k)\Delta x_k,$$

в силу которого

$$F(b) \approx F(a) + \sum_{k=0}^{N-1} f(x_k)\Delta x_k \ .$$

Погрешность аппроксимации будет тем меньше, чем меньше длина отрезка $[x_k, x_{k+1}]$, которая равна $\dfrac{b-a}{N}$ и уменьшается с ростом N. С другой стороны, при увеличении N растет число слагаемых в сумме

$$\sum_{k=0}^{N-1} f(x_k)\Delta x_k, \tag{6.1.1.1}$$

и предельный переход $N \to \infty$ приводит к необходимости раскрытия неопределенности вида «$0 \cdot \infty$». Сумму (6.1.1.1) принято называть *интегральной.*

В курсе математического анализа показывается, что для существования предела $\lim\limits_{N\to\infty} \sum\limits_{k=0}^{N-1} f(x_k)\Delta x_k$ достаточно, чтобы функция $f(x)$ была непрерывной на отрезке $[a, b]$.

Определение 6.1.1.1. Число, равное $\lim\limits_{N\to\infty} \sum\limits_{k=0}^{N-1} f(x_k)\Delta x_k$, называется *определенным интегралом функции* $f(x)$ в пределах от a до b и обозначается

$$\int\limits_a^b f(x)\,dx \quad .$$

Теперь решение задачи нахождения значения функции $F(x)$ в точке $x = b$ можно записать, при помощи определенного интеграла

$$F(b) = F(a) + \int\limits_a^b f(x)\,dx \quad ,$$

хотя на практике чаще используется иной вид записи этой формулы:

$$\int\limits_a^b f(x)\,dx = F(b) - F(a) \quad . \qquad (6.1.1.2)$$

или, для краткости

$$\int\limits_a^b f(x)\,dx = F(x)\Big|_a^b \quad . \qquad (6.1.1.3)$$

Соотношение (6.1.1.2) принято называть *формулой Ньютона-Лейбница.* Она связывает значения функции $F(x)$ в двух, вообще говоря не близких друг к другу точках через определенный интеграл от функции $f(x)$ в случае, когда $F'(x) = f(x)$, или позволяет находить значение определенного интеграла по значениям функции $F(x)$ в точках, являющимися пределами интегрирования. Наконец, стоит запомнить, что функцию $f(x)$ обычно называют *подынтегральной* *функцией,* переменную x – *переменной интегрирования,* а интервал (a, b) – *интервалом интегрирования.*

Пример 6.1.1.1 Рассмотрим в качестве иллюстрации вычисление определенного интеграла функции $f(x) = x^3$ на отрезке $[0, 1]$, то есть

$$\int_0^1 x^3\,dx \quad .$$

Легко видеть, что в этом случае $\Delta x_k = \dfrac{1}{N}$ и $x_k = \dfrac{k}{N}$. Поэтому интегральная сумма (6.1.1.1) будет иметь вид

$$\sum_{k=0}^{N-1} \left(\frac{k}{N}\right)^3 \cdot \frac{1}{N} \quad = \quad \frac{1}{N^4} \cdot \sum_{k=0}^{N-1} k^3$$

и нам нужно найти предел этого выражения при $N \to \infty$.

Согласно формуле, приведенной в конце пункта 10° §1.1,

$$\sum_{k=0}^{N-1} k^3 = \frac{(N-1)^2 N^2}{4} \,,$$

поэтому

$$\int_0^1 x^3\,dx = \lim_{N\to\infty} \frac{(N-1)^2 N^2}{4N^4} = \lim_{N\to\infty} \frac{\left(1 - \frac{1}{N}\right)^2}{4} = \frac{1}{4} \,.$$

Обратим внимание на следующие нюансы. Во-первых, значение определенного интеграла не зависит от того, как обозначается переменная интегрирования. Например, вместо x можно использовать t. Иначе говоря,

$$\int_a^b f(x)\,dx \qquad \text{и} \qquad \int_a^b f(t)\,dt \qquad \text{– это одно и то же!}$$

Во-вторых, если для функции $f(x)$ удается подобрать функцию $F(x)$, то вместо предельного перехода в интегральной сумме (6.1.1.1) при нахождении определенного интеграла гораздо удобнее пользоваться формулой Ньютона-Лейбница. Так в рассмотренном выше примере, для $f(x) = x^3$ легко находится, скажем, $F(x) = \dfrac{x^4}{4}$. Тогда

$$\int\limits_0^1 x^3\,dx = F(1) - F(0) = \left.\frac{x^4}{4}\right|_0^1 = \frac{1^4}{4} - \frac{0^4}{4} = \frac{1}{4} - 0 = \frac{1}{4}\,.$$

В заключение отметим, что dx в обозначении интеграла следует рассматривать с одной стороны как множитель подынтегрального выражения, а с другой – как дифференциал независимой переменной интегрирования x.

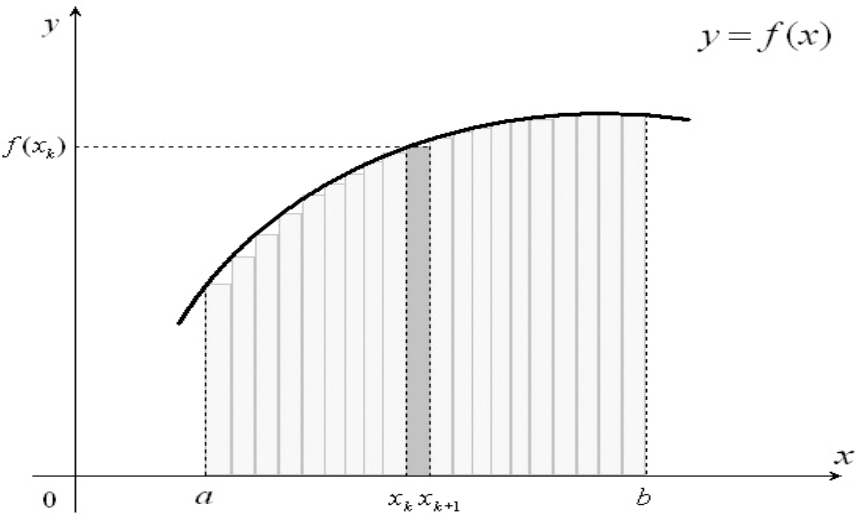


Рис. 6.1. Геометрический смысл определенного интеграла

6.1.2. Геометрический смысл и свойства определенного интеграла

Выясним теперь *геометрический смысл* определенного интеграла. На рис. 6.1 показаны график функции $y = f(x)$ и разбиение отрезка $[a, b]$ на N частей точками $x_1, x_2, \dots x_{N-1}$.

Нетрудно заметить, что для показанного на рис. 6.1, примера интегральная сумма (6.1.1.1)

$$\sum_{k=0}^{N-1} f(x_k)\Delta x_k,$$

равна площади «ступенчатой» фигуры, образованной из N прямоугольников, k−й из которых имеет длину основания $\Delta x_k = x_{k+1} - x_k$ и высоту $f(x_k)$.

При предельном переходе $N \to \infty$ эта площадь стремится к S – площади фигуры, ограниченной снизу осью Ox, слева и справа – вертикальными прямыми, проходящими через точки $x = a$ и $x = b$, и ограниченной сверху графиком функции $y = f(x)$. Поэтому в данном случае будет справедливо равенство

$$S = \int\limits_a^b f(x)\,dx \quad . \tag{6.1.2.1}$$

Важно отметить, что, поскольку площадь геометрической фигуры должна быть *неотрицательным* числом, использование приведенной геометрической интерпретации определенного интеграла для подсчета площадей следует производить *с учетом знака* подынтегральной функции $y = f(x)$.

Проиллюстрируем данное свойство определенного интеграла следующим примером. Пусть нам необходимо найти площадь фигуры, ограниченной сверху «полуволной» синусоиды и ограниченной снизу отрезком оси Ox в пределах от 0 до π (см. рис. 6.2.)

Согласно формуле (6.1.2.1) искомая площадь равна

$$S = \int\limits_0^\pi \sin x\,dx \ .$$

Вычислим этот интеграл по определению 6.1.1.1, по схеме, аналогичной использованной при решении примера 6.1.1.1. В данном случае $\Delta x_k = \dfrac{\pi}{N}$ и $x_k = \dfrac{\pi k}{N}$. Поэтому интегральная сумма (6.1.1.1) будет иметь вид

$$\sum_{k=0}^{N-1} \frac{\pi}{N} \sin\frac{\pi k}{N}$$

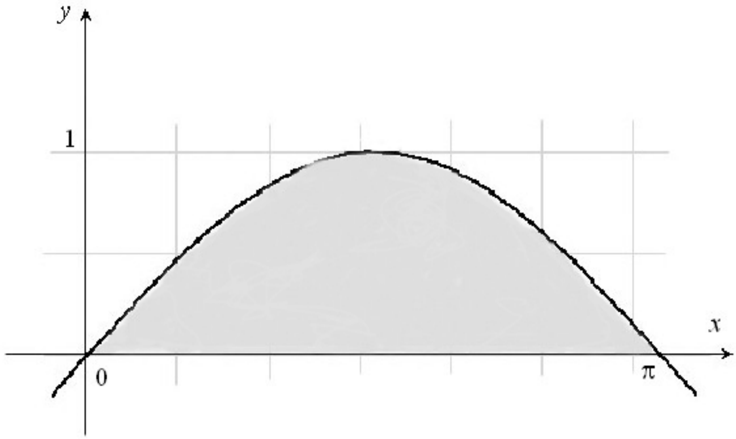


Рис. 6.2. Площадь фигуры, ограниченной «полуволной» синусоиды

и нам необходимо найти его предел при $N \to \infty$.
По формуле, приведенной в конце пункта 10° §1.1,

$$\sum_{k=0}^{N-1} \sin \frac{\pi k}{N} = \frac{\sin \dfrac{\pi(N+1)}{2N} \cdot \sin \dfrac{\pi(N-1)}{2N}}{\sin \dfrac{\pi}{2N}} \, ,$$

поэтому

$$\int\limits_0^\pi \sin x \, dx = \lim_{N\to\infty} \frac{\pi}{N} \cdot \frac{\sin \dfrac{\pi(N+1)}{2N} \cdot \sin \dfrac{\pi(N-1)}{2N}}{\sin \dfrac{\pi}{2N}} =$$

$$= \lim_{N\to\infty} 2 \cdot \left(\frac{\dfrac{\pi}{2N}}{\sin \dfrac{\pi}{2N}} \right) \cdot \sin \left(\frac{\pi}{2} \left(1 + \frac{1}{N} \right) \right) \cdot \sin \left(\frac{\pi}{2} \left(1 - \frac{1}{N} \right) \right) = 2 \, .$$

Здесь мы использовали, во-первых непрерывность функции $y = \sin x$, во-вторых равенство $\sin \dfrac{\pi}{2} = 1$ и, наконец, «первый замечательный предел», согласно которому

$$\lim_{N\to\infty} \frac{\sin\dfrac{\pi}{2N}}{\dfrac{\pi}{2N}} = 1 \ .$$

Отметим, что данная задача решается и при помощи формулы Ньютона-Лейбница (6.1.1.2), если заметить, что в соотношении $F'(x) = f(x)$ при $f(x) = \sin x$ можно использовать $F(x) = -\cos x$. Тогда

$$\int\limits_0^{\pi} \sin x\, dx = -\cos x\Big|_0^{\pi} = (-\cos\pi) - (-\cos 0) = (-(-1)) - (-1) = 2 \ .$$

Приведем теперь (без доказательства) основные свойства определенного интеграла, вытекающие из его определения.

1°. *Интегрирование суммы функций*:

$$\int\limits_a^b \Big(\ f(x) + g(x)\ \Big)\, dx = \int\limits_a^b f(x)\ dx + \int\limits_a^b g(x)\, dx \ \ .$$

2°. *Интегрирование произведения числа на функцию*:

$$\int\limits_a^b \Big(k\cdot f(x)\Big)\, dx = k\cdot \int\limits_a^b f(x)\, dx \ \ .$$

3°. *Аддитивность интегрирования, то есть интегрирование функции по объединению отрезков интегрирования*:

$$\int\limits_a^b f(x)\, dx + \int\limits_b^c f(x)\, dx \ \ = \ \ \int\limits_a^c f(x)\, dx \ \ .$$

4°. *По определению также считается, что справедливы равенства*:

$$\int\limits_a^b f(x)\, dx = -\int\limits_b^a f(x)\, dx \quad \text{и} \quad \int\limits_a^a f(x)\, dx = 0 \ .$$

6.2. Неопределенный интеграл

6.2.1. Связь первообразной и производной функций

Как уже было отмечено, в случае $f(x) = F'(x)$ функция $f(x)$ называется *производной функцией от функции* $F(x)$, а функция $F(x)$ – *первообразной функцией для функции* $f(x)$. Соотношение $f(x) = F'(x)$ выражает производную через первообразную, и естественно возникает вопрос о том, как выразить первообразную через производную. Ответ дается следующей формулой

$$F(x) = \int\limits_a^x f(t)\,dt\ , \tag{6.2.1.1}$$

где a – любое фиксированное число из области определения функции $f(x)$. Определенный интеграл в правой части равенства (6.2.1.1) называют *определенным интегралом с переменным верхним пределом.*

Убедимся, что из равенства (6.2.1.1) следует $F'(x) = f(x)$ для случая, когда функция $f(x)$ непрерывна. Действительно, для любой фиксированной точки x_0 из области определения $f(x)$ справедливы соотношения

$$F'(x_0) = \lim_{\Delta\to 0} \frac{F(x_0+\Delta) - F(x_0)}{\Delta} = \lim_{\Delta\to 0} \frac{1}{\Delta} \cdot \int\limits_{x_0}^{x_0+\Delta} f(t)\,dt =$$

$$= \lim_{\Delta\to 0} \frac{f(x_0)}{\Delta} \cdot \int\limits_{x_0}^{x_0+\Delta} 1 \cdot dt = \lim_{\Delta\to 0} f(x_0) \cdot \frac{1}{\Delta} \cdot \Delta = f(x_0)\ \ ,$$

в силу формулы Ньютона-Лейбница и по свойствам 2° и 3° для определенного интеграла. (Фактически мы допустили, что на малом промежутке $[x_0, x_0+\Delta]$ функция $f(x)$ имеет постоянное значение $f(x_0)$.) Таким образом, приходим к выводу, что каждая функция вида

$$F(x) = \int\limits_a^x f(t)\,dt$$

является первообразной для функции $f(x)$ и каждая первообразная для $f(x)$ представима в этом виде.

Попробуем теперь найти способ описания множества *всех* первообразных для непрерывной функции $f(x)$. Как мы убедились, интеграл (6.2.1.1) является (при фиксированном a) функцией от переменной x, которая также будет и одной из первообразных для $f(x)$ функций. Если изменить значение нижнего предела интегрирования, положив вместо a некоторое b, то мы получим *другую* первообразную функцию.

Обозначим эту новую первообразную как $F_1(x)$ и найдем связь между $F_1(x)$ и $F(x)$. По свойствам определенного интеграла

$$F_1(x) = \int\limits_b^x f(t)\,dt = \int\limits_b^a f(t)\,dt \ + \ \int\limits_a^x f(t)\,dt = F(x) + C\ ,$$

где $C = \int\limits_b^a f(t)\,dt$ – некоторая константа. Следовательно, все первообразные непрерывной функции $f(x)$ могут отличаться друг от друга лишь на произвольную постоянную.

Для практических целей удобно ввести специальное обозначение для всех первообразных функции $f(x)$.

Определение 6.2.1.1. Совокупность *всех* первообразных функций для некоторой функции $f(x)$ называется *неопределенным интегралом* и обозначается как $\int f(x)\,dx$.

Иными словами, в использованных нами обозначениях

$$\int f(x)\,dx \quad \equiv \quad \{\ F(x) + C\ ,\ \forall C\ \}\quad .$$

6.2.2. Свойства неопределенного интеграла. Правила интегрирования

На основании приведенных в §6.1.1 и §6.1.2 примеров можно заключить, что использование формулы Ньютона-Лейбница для подсчета значения определенного интеграла

$$\int\limits_a^b f(x)\,dx$$

гораздо удобнее, чем определение 6.1.1.1. Правда, для этого необходимо знать неопределенный интеграл или, хотя бы $F(x)$ – какую-нибудь первообразную для функции $f(x)$.

Естественно возникает вопрос: всегда ли по формуле для $f(x)$ можно построить формулу для $F(x)$? И, если можно, то как это сделать? Напомним, кстати, что по формуле $F(x)$ формула для $f(x)$ может быть построена (см. §5.2) всегда. Но в рассматриваемом случае – «увы и ах!» ответ на данный вопрос, вообще говоря, *отрицательный.* Самое лучшее, что здесь можно предложить, это запись вида

$$F(x)=\int\limits_a^x f(t)\,dt\ .$$

Примерами функций $f(x)$, первообразные для которых не выражаются через элементарные функции, служат e^{x^2}, $\dfrac{\sin x}{x}$, $\dfrac{1}{\ln x}$. Соответственно интегралы типа

$$\int e^{x^2}\,dx\ , \qquad \int \frac{\sin x}{x}dx\ , \qquad \int \frac{dx}{\ln x}$$

принято называть «неберущимися».

Вместе с тем, надо заметить, что, помимо очевидных соотношений типа

$$\int F'(x)\,dx=\Big\{\ F(x)+C\ ,\forall C\ \Big\}\ ,$$

в ряде практически важных случаев все же удается найти формулу для $F(x)$ по известной формуле для $f(x)$. Общее число этих случаев изменяется тысячами и их описание составляет содержание весьма, солидных по размерам и числу страниц, справочников. Для наших же целей вполне будет достаточно коллекции неопределенных интегралов, представленных в таблице 6.2.1. Эти интегралы в дальнейшем мы будем называть «табличными» и считать их известными.

Хотелось бы еще раз обратить внимание на существенность предположения о непрерывности функции $f(x)$. Рассмотрим подробнее четвертую формулу (отмеченную «звездочкой») в таблице 6.2.1. Можно убедиться, что приведенная формула для первообразной задает *не все* функции $F(x)$ такие, что $F'(x)=f(x)$. Например, для функции

$$F(x)=\left[\begin{array}{ll}\ln x+C_1\ , & \text{если}\ \ x>0\ ,\\ \ln(-x)+C_2\ , & \text{если}\ \ x<0\end{array}\right.$$

производные как для случая $x>0$, так и для $x<0$ представляются одной и той же формулой $\dfrac{1}{x}$, даже, если $C_1\neq C_2$.

Действительно,

$$\frac{d\,(\ln x + C_1)}{dx} = \frac{1}{x}\,,\ \text{если}\ \ x > 0\ ,$$
$$\frac{d\,(\ln(-x) + C_2)}{dx} = \frac{1}{(-x)} \cdot (-1) = \frac{1}{x}\,,\ \text{если}\ \ x < 0\ .$$

Дело в том, что функция $f(x) = \dfrac{1}{x}$ не является непрерывной: при $x_0 = 0$ она имеет точку неустранимого разрыва, что и приводит к нарушению равенства (6.2.1.1.) Впрочем, для практики эти проблемы не являются существенным ограничением. Скажем, функцию $f(x) = \dfrac{1}{x}$ можно рассматривать для случаев $x > 0$ и $x < 0$ независимо друг от друга. Тогда формула (6.2.1.1) будет выполняться, поскольку эта функция непрерывна как для всех $x > 0$, так и для любых отрицательных значений своего аргумента.

(Таблица 6.2.1a)

	$f(x)$	$\int f(x)\,dx$
1°	$x^a, a \neq -1$	$\dfrac{x^{a+1}}{a+1} + C$
2°	e^x	$e^x + C$
3°	$a^x, \quad a > 0, a \neq 1$	$\dfrac{a^x}{\ln a} + C$
4°	$\dfrac{1}{x}$	$\ln\|x\| + C^*$
5°	$\cos x$	$\sin x + C$
6°	$\sin x$	$-\cos x + C$

(Таблица 6.2.1b)

	$f(x)$	$\int f(x)\,dx$
7°	$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin x + C$
8°	$\dfrac{1}{a^2+x^2}$	$\dfrac{1}{a}\operatorname{arctg}\dfrac{x}{a}+C$, если $a \neq 0$
9°	$\dfrac{1}{\sqrt{x^2+a^2}}$	$\ln\left\lvert x+\sqrt{x^2+a^2}\right\rvert + C$, если $a > 0$
10°	$\sqrt{a^2-x^2}$	$\dfrac{x}{2}\sqrt{a^2-x^2}+\dfrac{a^2}{2}\arcsin\dfrac{x}{a}+C$, если $a > 0$
11°	$\sqrt{x^2-a^2}$	$\dfrac{x}{2}\sqrt{x^2-a^2}-\dfrac{a^2}{2}\ln\left\lvert x+\sqrt{x^2-a^2}\right\rvert + C$

Столь небольшое число «табличных» интегралов оказывается достаточным для решения значительного числа задач, поскольку в нашем распоряжении имеются также и формулы, позволяющие *выражать неопределенные интегралы одних функций через интегралы от других*. Эти формулы сведены в таблицу 6.2.2.

1°	$\int (\ f(x) + g(x)\ )\ dx = \int f(x)\,dx + \int g(x)\,dx$
2°	$\int (\ k \cdot f(x)\ )\ dx = k \cdot \int f(x)\,dx$, где k – const
3°	$\int f'(x) \cdot g(x)\,dx = f(x) \cdot g(x) - \int f(x) \cdot g'(x)\,dx$
4°	$\int f(g(x)) \cdot g'(x)\,dx = \int f(u)\,du$, где $u = g(x)$

(Таблица 6.2.2)

Данные формулы нуждаются в доказательстве, и для примера убедимся в справедливости формулы 3°, часто называемой правилом *интегрирования по частям.* Согласно пункту 3° таблицы 5.1.2 и определению интеграла

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x) \;\Rightarrow$$

$$\Rightarrow \int (f(x) \cdot g(x))'\ dx = \int (f'(x) \cdot g(x) + f(x) \cdot g'(x))\ dx \;\Rightarrow$$

$$\Rightarrow\; f(x) \cdot g(x) = \int f'(x) \cdot g(x)\,dx + \int f(x) \cdot g'(x)\,dx\ ,$$

поскольку из равенства функций следует равенство их неопределенных интегралов (но не первообразных!) Из последнего соотношения и следует формула интегрирования по частям.

$$\int f'(x) \cdot g(x)\,dx = f(x) \cdot g(x) - \int f(x) \cdot g'(x)\,dx\ .$$

Правило 4°, часто называемое *правилом замены переменной интегрирования,* также вполне очевидно, ибо по определению первого дифференциала из равенства $u = g(x)$ следует $du = g'(x)\,dx$. Заметим, что достаточно часто (особенно, если функция $g(x)$ не слишком сложна) правило 4° записывают в виде

$$\int f(g(x)) \cdot g'(x)\,dx = \int f(g(x))\,d(g(x))\ .$$

6.2.3. Примеры нахождения интегралов

Рассмотрим примеры подсчета значений определенных интегралов, основанных на следующей схеме:

A) подынтегральную функцию данного определенного интеграла приводим к виду, удобному для использования таблиц 6.2.1 и 6.2.2;

B) при помощи таблиц 6.2.1 и 6.2.2 находим неопределенный интеграл (или какую-нибудь первообразную);

C) применяем формулу Ньютона-Лейбница.

Вначале продемонстрируем как при помощи таблицы 6.2.2, зная неопределенный интеграл $\int f(x)\,dx$, можно получить формулу для $\int f(ax+b)\,dx$.

Пример 6.2.2.1. Пусть требуется найти

$$\int \cos(2x-7)\,dx \ .$$

Решение. Заметим, что в таблице 6.2.1 (формула 5°) имеется табличный интеграл

$$\int \cos x\,dx = \sin x + C \ .$$

Введем новую переменную $u = 2x - 7$, для которой (согласно таблице 5.3.1 – «свойства дифференциалов»), имеем

$$du = d(2x-7) = d(2x)+d(-7) = 2\,dx+0 = 2\,dx \qquad \Rightarrow \qquad dx = \frac{du}{2} \ .$$

Тогда искомый интеграл при помощи таблицы 6.2.2 (формула 2°) находится следующим образом

$$\int \cos(2x-7)\,dx \ = \int \cos u\,\frac{du}{2} \ = \frac{1}{2}\int \cos u\,du = \frac{1}{2}\sin u + C = \frac{1}{2}\sin(2x-7)+C.$$

Заметьте, что в ответе вместо $\dfrac{C}{2}$ написано C – это не опечатка. Произвольная константа, деленная пополам, все равно остается произвольной константой.

Проверьте самостоятельно, что аналогичным методом можно, например, получить формулы

$$\int e^{3x+1}\,dx = \frac{1}{3}e^{3x+1} + C \qquad \text{и} \qquad \int \frac{dx}{5x+2} = \frac{1}{5}\ln|5x+2| + C\,.$$

Далее заметим, что, несмотря на невозможность в общем случае записи интеграла в виде некоторой элементарной функции, имеют место случаи, когда это принципиально *всегда выполнимо*. К таким интегралам, в первую очередь, относятся интегралы от *дробно-рациональных* функций, то есть функций, представимых в виде дроби, числитель и знаменатель которых есть алгебраические многочлены, например,

$$\int \frac{x^2+4x+3}{x^3-3x^2+8x-6}\,d\,x\,.$$

Рассмотрим два основных метода интегрирования дробно-рациональных функций, в первом из которых удается разложить знаменатель на линейные множители.

Пример 6.2.2.2. Пусть требуется найти

$$\int \frac{(x-4)\,dx}{x^2-5x+6}\,.$$

Решение. Поскольку $x^2-5x+6=(x-2)(x-3)$, то попробуем представить подынтегральную функцию в виде

$$\frac{x-4}{x^2-5x+6} = \frac{A}{x-2} + \frac{B}{x-3}\,,$$

где A и B – некоторые числа, значения которых найдем из следующей цепочки равенств.

$$\frac{x-4}{x^2-5x+6} = \frac{A}{x-2} + \frac{B}{x-3} = \frac{A(x-3)-B(x-2)}{(x-2)(x-3)} = \frac{(A+B)x-(3A+2B)}{x^2-5x+6}\,.$$

Сравнивая начальное и конечное звенья этой цепочки, легко видеть, что значения чисел A и B (при которых данные равенства верны при всех x) являются решением системы линейных уравнений

$$\left\{\begin{array}{rcrcr} A & + & B & = & 1 \\ 3A & + & 2B & = & 4 \end{array}\right. \qquad \Rightarrow \qquad \left\{\begin{array}{lcr} A & = & 2 \\ B & = & -1 \end{array}\right. .$$

Теперь, используя таблицы 6.6.1 и 6.6.2, получаем

$$\int \frac{(x-4)\,dx}{x^2-5x+6} = \int \left(\frac{2}{x-2} - \frac{1}{x-3} \right) dx =$$

$$= 2\int \frac{dx}{x-2} - \int \frac{dx}{x-3} = 2\ln|x-2| - \ln|x-3| + C = \ln \frac{(x-2)^2}{|x-3|} + C\ .$$

Заметим, что этот алгоритм иногда называют «методом разложения на простейшие множители».

Следующие два примера показывают как можно действовать в случае, когда знаменатель дробно-рациональной функции не удается разложить на линейные множители.

Пример 6.2.2.3. Найти определенный интеграл

$$\int\limits_2^{2\sqrt{3}} \frac{dx}{x^2+4}\,.$$

Решение. По формуле 2° (из таблицы 6.2.2) выполним следующие преобразования соответствующего неопределенного интеграла

$$\int \frac{dx}{x^2+4} = \frac{1}{4}\int \frac{dx}{1+\left(\dfrac{x}{2}\right)^2} = \frac{1}{2}\int \frac{\dfrac{dx}{2}}{1+\left(\dfrac{x}{2}\right)^2} =$$

Теперь используем формулу 2° таблицы 5.1.2, определение дифференциала и правило 4° из таблицы 6.2.2, считая $u = \dfrac{x}{2}$.

$$= \frac{1}{2}\int \frac{d\left(\dfrac{x}{2}\right)}{1+\left(\dfrac{x}{2}\right)^2} = \frac{1}{2}\operatorname{arctg}\frac{x}{2} + C\ .$$

Заключительное равенство следует из формулы 8° таблицы 6.2.1.

Наконец, по формуле Ньютона-Лейбница находим определенный интеграл.

$$\int_2^{2\sqrt{3}} \frac{dx}{x^2+4} = \frac{1}{2}\operatorname{arctg}\frac{x}{2}\bigg|_2^{2\sqrt{3}} = \frac{1}{2}\operatorname{arctg}(\sqrt{3}) - \frac{1}{2}\operatorname{arctg}(1) = \frac{1}{2}\cdot\frac{\pi}{3} - \frac{1}{2}\cdot\frac{\pi}{4} = \frac{\pi}{24}\,.$$

Пример 6.2.2.4. Найти

$$\int_2^3 \frac{x^2+x+1}{x^2+1}\,dx\,.$$

Решение. Поскольку

$$\frac{x^2+x+1}{x^2+1} = 1 + \frac{x}{x^2+1}\,,$$

то применяя последовательно формулы 1° и 4° таблицы 6.2.2, получим

$$\int \frac{x^2+x+1}{x^2+1}\,dx = \int\left(1+\frac{x}{x^2+1}\right)dx = x + \int\frac{x}{x^2+1}\,dx =$$

$$= x + \frac{1}{2}\int\frac{d(x^2)}{x^2+1} = x + \frac{1}{2}\int\frac{d(x^2+1)}{x^2+1} = x + \frac{1}{2}\ln\left|x^2+1\right| + C\,.$$

Здесь, при использовании 4° мы полагали $u = x^2+1$, откуда следуют равенства для дифференциалов: $d(x^2+1) = d(x^2) = 2x\,dx$.

Теперь найдем значение определенного интеграла. По формуле Ньютона-Лейбница

$$\int_2^3 \frac{x^2+x+1}{x^2+1}\,dx = \left(x + \frac{1}{2}\ln\left(x^2+1\right)\right)\bigg|_2^3 =$$

$$= \left(3+\frac{1}{2}\ln 10\right) - \left(2+\frac{1}{2}\ln 5\right) = 1 + \frac{1}{2}\ln\frac{10}{5} = 1 + \ln\sqrt{2}\,.$$

Иногда в процессе интегрирования оказывается эффективным одновременное использование различных свойств неопределенного интеграла, приводимых в таблице 6.2.2.

Пример 6.2.2.5. Найти

$$\int_1^4 (\sqrt{x}+1)^3\,dx\ .$$

Решение. По формулам 1° и 2° в таблице 6.2.2, выражающим неопределенный интеграл от суммы функций через интегралы от слагаемых, и используя формулу «куб суммы двух чисел» (см. §1.1, п.2°), получаем

$$\int (\sqrt{x}+1)^3\,dx\ = \int \Big((\sqrt{x})^3 + 3(\sqrt{x})^2 + 3\sqrt{x} + 1\Big)\,dx\ =$$

$$= \int (\sqrt{x})^3\,dx\ +\ 3\int (\sqrt{x})^2\,dx\ +\ 3\int \sqrt{x}\,dx\ +\ \int 1\cdot\,dx\ =$$

Теперь используем первую формулу таблицы 6.2.1 и правила действий со степенями с дробным показателем.

$$= \int x^{3/2}\ dx + 3\int x\ dx + 3\int x^{1/2}\ dx + \int 1\cdot\,dx\ = \frac{2}{5}x^{5/2} + \frac{3}{2}x^2 + 2x^{3/2} + x + C\ .$$

Наконец, по формуле Ньютона-Лейбница найдем требуемое значение определенного интеграла.

$$\int_1^4 (\sqrt{x}+1)^3\,dx\ = \left(\frac{2}{5}x^{5/2} + \frac{3}{2}x^2 + 2x^{3/2} + x\right)\Bigg|_1^4 =$$

$$= \left(\frac{2}{5}\cdot 4^{5/2} + \frac{3}{2}\cdot 4^2 + 2\cdot 4^{3/2} + 4\right) - \left(\frac{2}{5}\cdot 1^{5/2} + \frac{3}{2}\cdot 1^2 + 2\cdot 1^{3/2} + 1\right) = \frac{519}{10}\,.$$

Пример 6.2.2.6. Найти

$$\int_0^\pi x\cdot\cos x\,dx\ .$$

Решение. Для нахождения неопределенного интеграла применим правило интегрирования по частям, то есть формулу 3° из таблицы 6.2.2. Будем полагать, что $f(x) = \sin x$, а $g(x) = x$. Тогда соответствующий неопределенный интеграл можно преобразовать следующим образом.

$$\int x\cdot\cos x\,dx\ = \int x\cdot(\sin x)'\,dx\ = x\cdot\sin x - \int (x)'\cdot\sin x\,dx\ =$$

Поскольку $(x)' = 1$, по шестой формуле таблицы 6.2.1 получаем

$$= x \cdot \sin x - \int \sin x\, dx \ = x \cdot \sin x - (-\cos x) = x \cdot \sin x + \cos x + C \ .$$

По формуле Ньютона-Лейбница имеем

$$\int\limits_0^\pi x \cdot \cos x\, dx \ = (x \cdot \sin x + \cos x)\Big|_0^\pi =$$

$$= (\pi \cdot \sin \pi + \cos \pi) - (0 \cdot \sin 0 + \cos 0) = (-1) - (1) = -2 \ .$$

Пример 6.2.2.7. Найти

$$\int\limits_0^1 \operatorname{arctg} x\, dx \ .$$

Решение. Применим формулу 3° *интегрирования по частям,* используя равенство $(x)' = 1$,

$$\int \operatorname{arctg} x\, dx \ = \int 1 \cdot \operatorname{arctg} x\, dx \ = x \operatorname{arctg} x - \int x \cdot \frac{1}{1+x^2} dx \ =$$

затем – формулу 4° таблицы 6.2.2, что с учетом $u = x^2$ дает

$$= x \operatorname{arctg} x - \frac{1}{2} \int \frac{d(x^2)}{1+x^2} dx \ = x \operatorname{arctg} x - \frac{1}{2} \ln(1+x^2) + C \ .$$

Наконец, по формуле Ньютона-Лейбница,

$$\int\limits_0^1 \operatorname{arctg} x\, dx \ = \left(x \operatorname{arctg} x - \frac{1}{2} \ln(1+x^2) \right)\Bigg|_0^1 =$$

$$= \left(1 \operatorname{arctg} 1 - \frac{1}{2} \ln(1+1^2) \right) - \left(0 \operatorname{arctg} 0 - \frac{1}{2} \ln(1+0^2) \right) = \frac{\pi}{4} - \ln \sqrt{2} \ .$$

В завершение обзора приведем пример, в котором оказывается целесообразным двукратное использование правила интегрирования по частям.

Пример 6.2.2.8. Найти

$$\int\limits_0^\pi e^x \sin x\, dx \ .$$

Решение. Применив *интегрирования по частям,* получим

$$\int e^x \sin x\, dx \ = e^x(-\cos x) - \int e^x(-\cos x)\, dx \ ,$$

то есть выражение содержащее интеграл, который ничуть не проще исходного. Однако повторное интегрирование по частям дает соотношение

$$\int e^x \sin x\, dx \ = -e^x \cos x + e^x \sin x - \int e^x \sin x\, dx \ ,$$

из которого следует

$$2\cdot \int e^x \sin x\, dx \ = -e^x \cos x + e^x \sin x + C \quad \Rightarrow \quad \int e^x \sin x\, dx \ = \frac{1}{2} e^x (\sin x - \cos x) + C$$

По формуле Ньютона-Лейбница

$$\int\limits_0^{\pi} e^x \sin x\, dx \ = \left(\frac{1}{2} e^x (\sin x - \cos x) \right) \Big|_0^{\pi} =$$

$$= \left(\frac{1}{2} e^{\pi} (\sin \pi - \cos \pi) \right) - \left(\frac{1}{2} e^{0} (\sin 0 - \cos 0) \right) = \frac{e^{\pi} + 1}{2} \, .$$

В заключение обзора методов интегрирования следует отметить, во-первых, что достаточно часто пункты A), B) и C) объединяют, чтобы получить более компактную форму записи. Это можно делать, однако, соблюдая определенные правила. Например, полагая $u = \sin x$, будем иметь

$$\int\limits_0^{\frac{\pi}{2}} e^{\sin x} \cos x\, dx \ = \int\limits_0^{1} e^u\, du \ = e^u \Big|_0^1 = e - 1 \ .$$

Обратите внимание, что при замене переменной (x на u) *пределы интегрирования также меняются*: вместо $[0, \frac{\pi}{2}]$ нужно брать $[0, 1]$, поскольку при замене $u = \sin x$ имеет место $\sin 0 = 0$ и $\sin \frac{\pi}{2} = 1$.

Во-вторых, подсчет значений определенных интегралов иногда можно упростить, использовав равенства

$$\int_{-a}^{a} f(x)\,dx \;=\; 2\cdot\int_{0}^{a} f(x)\,dx\;, \tag{6.2.3.1}$$

если функция $f(x)$ – четная, и

$$\int_{-a}^{a} f(x)\,dx \;=\; 0,$$

в случае нечетной $f(x)$. Проверьте справедливость этих равенств самостоятельно.

Рассмотрим теперь примеры прикладных задач, решение которых сводится к нахождению значения некоторого определенного интеграла.

Пример 6.2.2.9. Какую работу надо совершить, чтобы вытащить цилиндрическую пробку из горлышка бутылки (рис. 6.3)? Длина пробки равна L, сила трения пробки о поверхность горлышка, когда пробка целиком находится внутри горлышка, равна F и пропорциональна площади соприкосновения поверхностей пробки и горлышка, когда пробка вытащена лишь частично.

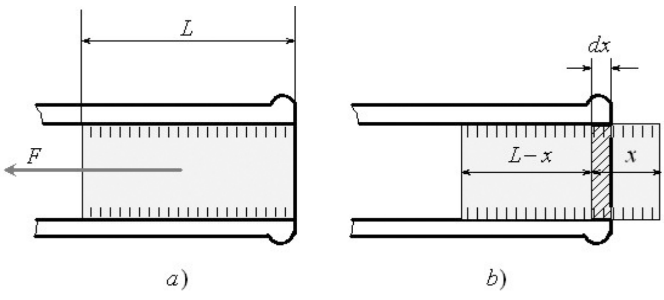


Рис. 6.3. Какую работу надо совершить, чтобы вытащить пробку?

Решение. Сила сопротивления движению (сила трения) максимальна в начальном положении пробки (см. рис. 6.3*a*.) По мере извлечения пробки из горлышка величина этой силы уменьшается, так как уменьшается площадь соприкосновения поверхностей пробки и горлышка.

Пусть пробка вытащена на длину x, тогда (см. рис. 6.3b) согласно условию задачи величина силы трения будет $F(x) = F \cdot \dfrac{L-x}{L}$, а работа по вытаскиванию пробки на малом перемещении dx равна $dA = F(x)\,dx$. Пробка будет вытащена, когда x станет равным L, поэтому полная работа равняется определенному интегралу $A = \int\limits_0^L F(x)\,dx$. Подставляя, получаем

$$A = \int\limits_0^L \left(F \cdot \frac{L-x}{L}\right) dx \;=\; F\int\limits_0^L 1 \cdot dx \;-\; \frac{F}{L}\int\limits_0^L x\,dx \;=$$

$$= F\left(x\Big|_0^L\right) - \frac{F}{L}\left(\frac{x^2}{2}\Big|_0^L\right) = FL - \frac{F}{L}\cdot\frac{L^2}{2} = \frac{1}{2}FL\,.$$

Пример 6.2.2.10. За какое время опорожнится заполненная водой цилиндрическая бочка высоты $H = 1.5$м и радиуса $R = 0.5$м через отверстие в ее основании радиуса $r = 1$см?

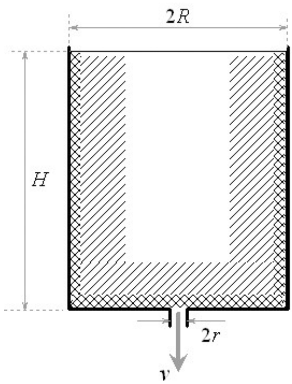


Рис. 6.4. За какое время вытечет вся вода из бочки?

Решение. Площадь горизонтального сечения воды в бочке $S = \pi R^2$, а площадь отверстия слива $s = \pi r^2$. Пусть за малое время dt понижение уровня воды в бочке составит dh, при этом $S\,dh = sv\,dt$, поскольку количество вытекшей жидкости очевидно равно уменьшению объема воды в бочке. С другой стороны, по закону Торричелли, скорость истечения жидкости из сосуда, когда ее уровень в нем равен h, определяется формулой $v = \sqrt{2gh}$. Поэтому

$$dt = \frac{S}{sv} dh = \frac{1}{\sqrt{2g}} \left(\frac{R}{r} \right)^2 \cdot \frac{dh}{\sqrt{h}} .$$

Бочка опустеет, когда понижение уровня станет равным H, то есть за время T, равное определенному интегралу,

$$T = \int\limits_0^H \frac{1}{\sqrt{2g}} \left(\frac{R}{r} \right)^2 \cdot \frac{dh}{\sqrt{h}} = \frac{1}{\sqrt{2g}} \left(\frac{R}{r} \right)^2 \cdot \int\limits_0^H \frac{dh}{\sqrt{h}} =$$

значение которого найдем по формуле 1° таблицы 6.2.1 при $a = -\dfrac{1}{2}$,

$$= \frac{1}{\sqrt{2g}} \left(\frac{R}{r} \right)^2 \cdot \left(2\sqrt{h} \Big|_0^H \right) = \left(\frac{R}{r} \right)^2 \sqrt{\frac{2H}{g}} .$$

При $H = 1.5$м, $R = 0.5$м, $r = 1$см и, приняв $g \approx 10 \dfrac{\text{м}}{c^2}$, получим $T \approx 1370c$ или, примерно, 23мин.

Пример 6.2.2.11. («Задача Коши») Найти функцию $F(x)$ такую, что

$$F'(x) = e^{2x} + \cos 3x \quad \text{ и } \quad F(0) = \frac{3}{2} .$$

Решение. Множество всех функций $F(x)$, для которых $F'(x) = e^{2x} + \cos 3x$, является неопределенным интегралом

$$\int \left(e^{2x} + \cos 3x \right) \, dx = \frac{1}{2} e^{2x} + \frac{1}{3} \sin 3x + C .$$

Если в эту формулу подставить значение $x = 0$ и приравнять полученную величину $\dfrac{3}{2}$, то получим равенство

$$\frac{1}{2} e^{2 \cdot 0} + \frac{1}{3} \sin(3 \cdot 0) + C = \frac{3}{2} ,$$

из которого находим, что $C = 1$. Значит искомая функция имеет вид

$$F(x) = \frac{1}{2}e^{2x} + \frac{1}{3}\sin 3x + 1 \ .$$

Завершая рассмотрение практического использования определенных интегралов, приведем пример геометрической задачи.

Пример 6.2.2.12. Пусть требуется найти площадь криволинейной фигуры, координаты точек которой в прямоугольной декартовой системе координат удовлетворяют условиям

$$\begin{cases} 0 \le y \le x(x-3)^2, \\ 0 \le x \le 3 \ . \end{cases}$$

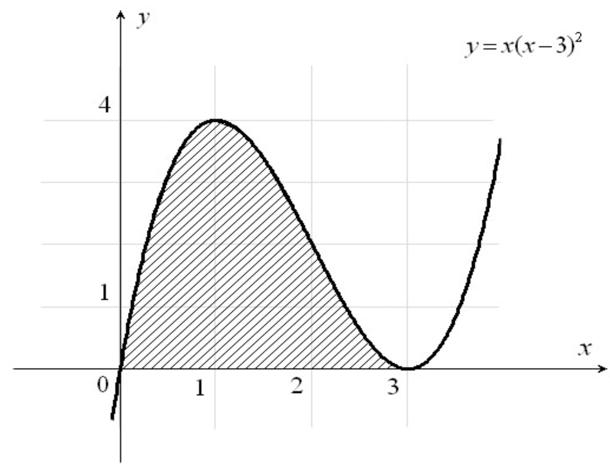


Рис. 6.5. Чему равна площадь криволинейной фигуры?

Решение. Глядя на рисунок 6.5, нетрудно заметить, что точки, координаты которых удовлетворяют данным ограничениям, принадлежат заштрихованной криволинейной фигуре, ограниченной сверху графиком функции $y = f(x)$, а снизу – графиком функции $y = 0$, то есть просто, осью Ox.

Найдем величину площади этой фигуры. Согласно формуле (6.1.2.1) имеем

$$S = \int_0^3 x(x-3)^2 d\,x \ .$$

Приведем вначале подинтегральную функцию к виду, удобному для использования таблиц 6.2.1 и 6.2.2.

$$\int x(x-3)^2 d\,x = \int \left(x^3 - 6x^2 + 9x\right) d\,x.$$

Затем, по таблицам 6.2.1 и 6.2.2 находим неопределенный интеграл

$$\int \left(x^3 - 6x^2 + 9x\right) d\,x = \int x^3 d\,x - 6\int x^2 d\,x + 9\int x d\,x = \frac{x^4}{4} - 2x^3 + \frac{9x^2}{2} + C\ .$$

Наконец, используем формулу Ньютона-Лейбница

$$S = \left(\frac{x^4}{4} - 2x^3 + \frac{9x^2}{2}\right)\Bigg|_0^3 = \left(\frac{3^4}{4} - 2\cdot 3^3 + \frac{9\cdot 3^2}{2}\right) - \left(\frac{0^4}{4} - 2\cdot 0^3 + \frac{9\cdot 0^2}{2}\right)\ .$$

И окончательно

$$S = \frac{81}{4} - 2\cdot 27 + \frac{81}{2} = \frac{27}{4}\,.$$

6.3. Несобственные интегралы

В предыдущих параграфах мы имели дело с двумя основными типами интегралов: определенным и неопределенным. Рассмотрим теперь еще один класс интегралов, называемых *несобственными.*

В приложениях достаточно часто возникает ситуация, когда на промежутке интегрирования (a, B) либо не удается построить интегральную сумму (6.1.1.1), либо ее предел при $N \to \infty$ не существует, однако существуют определенные интегралы $\int\limits_a^b f(x)\,dx$ при *любых фиксированных* a и b таких, что $a < b < B$. Тогда можно дать

Определение 6.3.1. *Несобственным интегралом* от функции $f(x)$ на промежутке от a до B называется предел вида

$$\lim_{b\to B}\int\limits_a^b f(x)\,dx\ ,$$

которого используется (по исторически сложившимся причинам, к сожалению внешне совпадающий с обозначением для определенного интеграла) символ

$$\int\limits_a^B f(x)\,dx\ .$$

Точка B при этом может быть как конечной, так и бесконечной.

Аналогично может быть определен несобственный интеграл для нижнего предела интегрирования:

$$\int\limits_A^b f(x)\,dx \;=\; \lim_{a\to A}\int\limits_a^b f(x)\,dx\ ,$$

Если предел, указанный в определении 6.3.1 существует, то этот несобственный интеграл принято называть *сходящимся,* иначе говорят о *не сходящимся* (или *расходящимся*) несобственном интеграле.

Пример 6.3.1. Исследовать на сходимость несобственный интеграл

$$\int\limits_0^{+\infty}\frac{dx}{x^2+1}\ .$$

Решение. Имеем, согласно формуле 8° из таблицы 6.2.1

$$\int\limits_0^{+\infty}\frac{dx}{x^2+1} = \lim_{b\to+\infty}\int\limits_0^{b}\frac{dx}{x^2+1} = \lim_{b\to+\infty}\left(\operatorname{arctg}x\Big|_0^b\right) = \lim_{b\to+\infty}\operatorname{arctg}b = \frac{\pi}{2}\ .$$

Значит, данный несобственный интеграл сходится.

Пример 6.3.2. Исследовать на сходимость несобственный интеграл

$$\int\limits_0^{1}\frac{dx}{\sqrt[3]{x^2}}\ .$$

Решение. Используя формулу 1° из таблицы 6.2.1, получим

$$\int\limits_0^{1}\frac{dx}{\sqrt[3]{x^2}} = \lim_{a\to0}\int\limits_a^{1}\frac{dt}{\sqrt[3]{x^2}} = \lim_{a\to0}\left(3\sqrt[3]{x}\Big|_a^1\right) = \lim_{a\to0}\left(3-3\sqrt[3]{a}\right) = 3\ .$$

И этот несобственный интеграл также сходится.

Приведенные примеры демонстрируют, что несобственные интегралы могут быть двух видов: «несобственные интегралы на неограниченном промежутке интегрирования» (как в примере 6.3.1) и «несобственные интегралы от неограниченных функций<<, подобные случаю рассмотренному в примере 6.3.2.

Пример 6.3.3. Выяснить, при каких значениях параметра p несобственный интеграл

$$\int\limits_1^{+\infty} \frac{dx}{x^p}$$

сходится.

Решение. Снова используем определение 6.3.1 и формулы таблицы 6.2.1. При $p \neq 1$

$$\lim_{b\to+\infty} \int\limits_1^{b} \frac{dx}{x^p} = \lim_{b\to+\infty} \left(\frac{x^{1-p}}{1-p} \Bigg|_1^b \right) =$$

$$= \lim_{b\to+\infty} \left(\frac{b^{1-p}}{1-p} - \frac{1}{1-p} \right) = \left[\begin{array}{ll} \dfrac{1}{1-p}, & \text{если } p > 1 \ , \\ \infty \ , & \text{если } p < 1 \ , \end{array} \right.$$

Если же $p = 1$, то интеграл

$$\lim_{b\to+\infty} \int\limits_1^{b} \frac{dx}{x} = \lim_{b\to+\infty} \ln|x| \Bigg|_1^b = \lim_{b\to+\infty} (\ln b - \ln 1) = \lim_{b\to+\infty} \ln b = +\infty.$$

то есть он расходится. Следовательно, исходный несобственный интеграл сходится при $p > 1$ и расходится при $p \leq 1$.

Пример 6.3.4. Выяснить, при каких значениях параметра p несобственный интеграл

$$\int\limits_0^{1} \frac{dx}{x^p}$$

сходится.

Решение. По определению 6.3.1 и формуле 1° таблицы 6.2.1, при $p \neq 1$

$$\lim_{a\to 0} \int\limits_a^{1} \frac{dx}{x^p} = \lim_{a\to 0} \left(\frac{x^{1-p}}{1-p} \Bigg|_a^1 \right) =$$

$$= \lim_{a\to 0}\left(\frac{1}{1-p} - \frac{a^{1-p}}{1-p}\right) = \left[\begin{array}{ll} \dfrac{1}{1-p}, & \text{если } p < 1 , \\ \infty , & \text{если } p > 1 , \end{array}\right.$$

Если же $p = 1$, то интеграл

$$\lim_{a\to 0}\int\limits_a^1 \frac{dx}{x} = \lim_{a\to 0} \ln|x| \Big|_a^1 = \lim_{a\to 0}(\ln 1 - \ln a) = \lim_{a\to 0}(-\ln a) = +\infty.$$

и он также расходится. В итоге исходный интеграл сходится при $p < 1$ и расходится при $p \geq 1$.

В значительном числе практически важных задач оказывается необходимым лишь установление факта сходимости или расходимости несобственного интеграла, *без вычисления его значения.* В этих случаях можно попытаться применить *признаки сходимости для несобственных интегралов.*

Рассмотрим непрерывные функции $f(x)$ и $g(x)$, определенные для любого x принадлежащего промежутку от a до B и такие, что справедливо неравенство $0 \leq f(x) \leq g(x)$. Тогда

$$\int\limits_a^b f(x)\,dx \ \leq \ \int\limits_a^b g(x)\,dx$$

и в силу сохранения нестрогих неравенств при предельном переходе (см. теорему «о двух милиционерах»), можно утверждать, что из сходимости несобственного интеграла $\int\limits_a^B g(x)\,dx$ следует сходимость и несобственного интеграла $\int\limits_a^B f(x)\,dx$, а из расходимости интеграла $\int\limits_a^B f(x)\,dx$ вытекает расходимость интеграла $\int\limits_a^B g(x)\,dx$.

Следует также отметить, что в случае, например, расходимости интеграла от функции $f(x)$ заключение о сходимости или расходимости интеграла от $g(x)$ сделать невозможно.

Проиллюстрируем использование этих правил следующими примерами.

Пример 6.3.5. Не находя значения несобственного интеграла

$$\int\limits_1^{+\infty} \frac{dx}{x\sqrt{x} + \dfrac{1}{x^4}} \,,$$

выяснить, сходится он или не сходится.

Решение. На промежутке $[1, +\infty)$ справедливо неравенство

$$0 \le \frac{1}{x\sqrt{x} + \dfrac{1}{x^4}} \le \frac{1}{x\sqrt{x}} \,,$$

поэтому из сходимости интеграла

$$\int\limits_1^{+\infty} \frac{dx}{x\sqrt{x}}$$

(см. пример 6.3.3, при $p = \dfrac{3}{2} > 1$) будет следовать сходимость интеграла

$$\int\limits_1^{+\infty} \frac{dx}{x\sqrt{x} + \dfrac{1}{x^4}} \,.$$

Пример 6.3.6. Не находя значения несобственного интеграла

$$\int\limits_0^{\frac{1}{2}} \frac{dx}{x\sqrt{x} - x^4} \,,$$

выяснить, сходится он или не сходится.

Решение. На промежутке $(0, \dfrac{1}{2}]$ справедливо неравенство

$$\frac{1}{x\sqrt{x} - x^4} \ge \frac{1}{x\sqrt{x}} \ge 0 \,,$$

поэтому из расходимости интеграла

$$\int\limits_0^{\frac{1}{2}} \frac{dx}{x\sqrt{x}}$$

(см. пример 6.3.4, при $p = \dfrac{3}{2} > 1$) будет следовать расходимость интеграла

$$\int\limits_0^{\frac{1}{2}} \frac{dx}{x\sqrt{x} - x^4}\,.$$

В заключение продемонстрируем различие между определенным и несобственным интегралом на примере решения следующих задач:

Пример 6.3.7. Найти площадь фигуры, ограниченной графиком функции $y = xe^{-x}$:

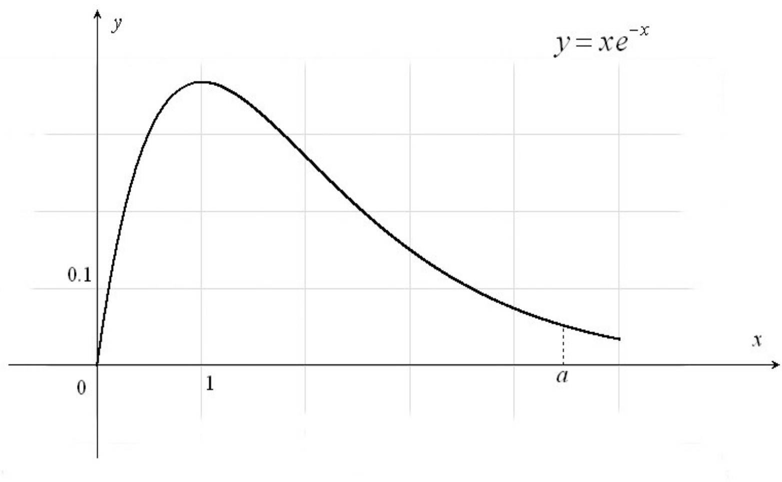


Рис. 6.6. Площадь фигуры, ограниченной графиком $y = xe^{-x}$.

а) при $0 \le x \le 3$,
б) при $0 \le x < +\infty$.

Решение. а) Функция $y = xe^{-x}$ имеет неотрицательные значения на рассматриваемых промежутках (см. рис. 6.6,) поэтому (согласно §6.1.2) площадь под графиком для $0 \le x \le a$ равна *определенному интегралу*

$$\int\limits_0^a xe^{-x}\,dx\,.$$

Заметим, что $e^{-x} = (-e^{-x})'$. Тогда, воспользовавшись формулой 3° таблицы 6.2.2 – интегрирования по частям, получим

$$\int_0^a xe^{-x}\,dx = x\left(-e^{-x}\right)\Big|_0^a - \int_0^a \left(-e^{-x}\right)\,dx =$$

$$= -xe^{-x}\Big|_0^a + \int_0^a e^{-x}\,dx = \left(-xe^{-x} - e^{-x}\right)\Big|_0^a = 1-(a+1)e^{-a}\,.$$

Значит, при $a = 3$ искомая площадь равна

$$S = 1 - \frac{4}{e^3}\,.$$

б) при неограниченных значениях x искомая площадь выражается *несобственным* интегралом

$$\int_0^{+\infty} xe^{-x}\,dx = \lim_{a\to+\infty}\int_0^a xe^{-x}\,dx = \lim_{a\to+\infty}\left(1-\frac{a+1}{e^a}\right) = 1\,.$$

Пример 6.3.8. Найти с какой силой притягивается точечная масса M

а) отрезком тонкой нити длины L, находящейся на расстоянии H от M. Масса единицы длины нити равна μ,

б) той же нитью, но неограниченной длины.

Решение. а) Выберем прямоугольную систему координат, такую, что нить расположена на оси Ox, а массивная точка M – на оси Oy в точке B (см. рис.6.7.) Рассмотрим отрезок нити малой длины dx, середина которого совпадает с точкой A, с координатами $\{x, 0\}$. Поскольку масса этого отрезка $\mu\ dx$, то по закону всемирного тяготения величина силы притяжения между отрезком и точкой M равна $\left|\overrightarrow{dF}\right| = \gamma\dfrac{M\mu\,dx}{(AB)^2}$, где γ – гравитационная постоянная и (по теореме Пифагора) $(AB)^2 = x^2 + H^2$. Сила $\overrightarrow{dF}$ является векторной величиной и ее следует суммировать по «правилу параллелограмма».

В рассматриваемом случае взаимодействие между материальной точкой и нитью симметрично относительно оси Oy – силы действующие вдоль оси Ox взаимно уничтожаются, поэтому следует суммировать лишь силы dF_y, параллельные оси Oy. С учетом этой симметрии

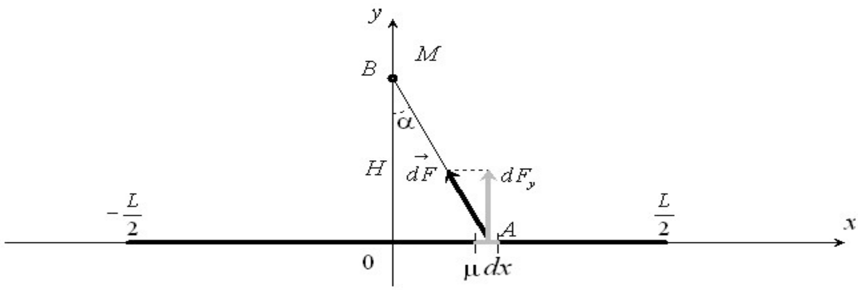


Рис. 6.7. Какова сила притяжения между точечной массой M и тонкой нитью?

сила притяжения между M и нитью длины L будет равна *определенному* интегралу

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} \gamma \frac{M\mu \cos\alpha \, dx}{x^2+H^2}, \quad \text{равного в силу (6.2.3.1)} \quad 2\gamma M\mu H \int_0^{\frac{L}{2}} \frac{dx}{(x^2+H^2)\sqrt{x^2+H^2}},$$

поскольку $dF_y = \left|\overrightarrow{dF}\right| \cos\alpha$, а $\cos\alpha = \dfrac{H}{\sqrt{x^2+H^2}}$.

Используя таблицы интегралов, получаем

$$F = 2\gamma M\mu H \int_0^{\frac{L}{2}} \frac{dx}{(x^2+H^2)\sqrt{x^2+H^2}} =$$

$$= 2\gamma M\mu \left(\frac{x}{H\sqrt{x^2+H^2}} \right) \Bigg|_0^{\frac{L}{2}} = \frac{2\gamma M\mu}{H} \cdot \frac{L}{\sqrt{L^2+4H^2}} .$$

б) Для случая бесконечной нити, полная сила притяжения выражается *несобственным* интегралом

$$F = 2\gamma M\mu H \int_0^{+\infty} \frac{dx}{(x^2+H^2)\sqrt{x^2+H^2}} =$$

$$= \frac{2\gamma M\mu}{H}\cdot \lim_{L\to+\infty} \frac{L}{\sqrt{L^2+4H^2}} = \frac{2\gamma M\mu}{H}\cdot \lim_{L\to+\infty} \frac{1}{\sqrt{1+\left(\frac{2H}{L}\right)^2}} = \frac{2\gamma M\mu}{H}.$$

6.4. Ряды

6.4.1. Числовые ряды

Понятие *числового ряда,* основанное на приводимых ниже определениях, расширяет возможность применения операции сложения на случай неограниченного числа слагаемых.

Определение 6.4.1.1. Пусть дана некоторая числовая последовательность $\{a_n\}$. Тогда выражение вида

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots = \sum_{k=1}^{\infty} a_k$$

называется *числовым рядом,* а числа $a_1, a_2, a_3, \ldots a_n, \ldots$ – *членами числового ряда.*

Определение 6.4.1.2. Суммы первых n членов числового ряда $\sum\limits_{k=1}^{\infty} a_k$

$$S_n = a_1 + a_2 + a_3 + \ldots + a_n = \sum_{k=1}^{n} a_k$$

называются *частичными суммами* этого ряда.

Определение 6.4.1.1 вызывает два вполне естественных вопроса:

- что такое "сумма"бесконечного числа чисел ?
- если такая "сумма"существует, то каковы ее свойства?

Попробуем оценить возможные ответы на эти вопросы.

Пусть отрезок оси вещественных чисел $[0, 1]$ последовательно делится на две части по следующему правилу:

на первом шаге он делится пополам, а на каждом последующем шаге пополам делится правый из получившихся отрезков.

Требуется найти длину левого отрезка при неограниченном увеличении числа шагов этой процедуры.

Значения длин левых отрезков будут образовывать числовую последовательность вида:

$$L_1 = \frac{1}{2}\,;$$

$$L_2 = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}\,;$$

$$L_3 = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}\,;$$

$$L_4 = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}\,;$$

$$\ldots$$

Таким образом, мы приходим к равенству, которое позволяет формально записать, что

$$\lim_{N\to+\infty} L(N) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = \sum_{k=1}^{+\infty} \frac{1}{2^k}\,.$$

Интуитивно ясно, что длина левого отрезка будет приближаться к числу единица. С этим был не согласен древнегреческий философ Зенон, справедливо полагавший, что человеческой жизни не хватит для выполнения бесконечно большого числа сложений. Впрочем, строго судить его не стоит, ибо он не знал (в отличие от нас) что такое геометрическая прогрессия и не был знаком с теорией пределов числовых последовательностей.

В рассматриваемом случае просто

$$L(N) = \frac{\dfrac{1}{2} - \dfrac{1}{2^{N+1}}}{1 - \dfrac{1}{2}} \qquad \text{и} \qquad \lim_{N\to+\infty} L(N) = 1\,.$$

Определение 6.4.1.3. Если существует $\lim\limits_{n\to+\infty} S_n = A$, то числовой ряд $\sum\limits_{k=1}^{+\infty} a_k$ называется *сходящимся,* а число A – *суммой ряда.* В противном случае числовой ряд называется *расходящимся.*

Приведем примеры числовых рядов. В курсе элементарной математики рассматривается геометрическая прогрессия (см. п.9° §1.1) – числовая последовательность, $n-$й член которой равен $b_n = b_1 q^{n-1}$. В случае $|q| < 1$ (то есть когда эта геометрическая прогрессия бесконечно убывающая,) она порождает сходящийся числовой ряд

$$\sum_{k=1}^{+\infty} b_1 q^{k-1} = b_1 + b_1 q + b_1 q^2 + b_1 q^3 + b_1 q^4 + \ldots \ ,$$

сумма которого равна $A = \dfrac{b_1}{1-q}$.

Иногда, зная сумму одного ряда, можно найти сумму другого. Убедитесь самостоятельно, что из формулы

$$\sum_{k=1}^{+\infty} q^{k-1} = 1 + q + q^2 + q^3 + q^4 + \ldots = \frac{1}{1-q}$$

следует, что

$$\sum_{k=1}^{+\infty} k \cdot q^k = q + 2q^2 + 3q^3 + 4q^4 + \ldots = \frac{q}{(1-q)^2} \, . \qquad (6.4.1.1)$$

Другими примерами числовых рядов могут служить:

1°. $\sum\limits_{k=1}^{+\infty} \dfrac{1}{k^2} = 1 + \dfrac{1}{2^2} + \dfrac{1}{3^2} + \dfrac{1}{4^2} + \ldots = 1 + \dfrac{1}{4} + \dfrac{1}{9} + \dfrac{1}{16} + \ldots \ ,$

2°. $\sum\limits_{k=1}^{+\infty} \dfrac{1}{k} = 1 + \dfrac{1}{2} + \dfrac{1}{3} + \dfrac{1}{4} + \ldots \ ,$

3°. $\sum\limits_{k=1}^{+\infty} k e^{-k} = \dfrac{1}{e} + \dfrac{2}{e^2} + \dfrac{3}{e^3} + \dfrac{4}{e^4} + \ldots \ ,$

4°. $\sum\limits_{k=0}^{+\infty} \cos k = 1 + \cos 1 + \cos 2 + \cos 3 + \ldots \ ,$

5°. $\sum\limits_{k=1}^{+\infty} \dfrac{\sin k}{k} = \sin 1 + \dfrac{\sin 2}{2} + \dfrac{\sin 3}{3} + \dfrac{\sin 4}{4} + \ldots \ .$

В курсе высшей математики показывается, что ряды 1°, 3° и 5° сходящиеся, а ряды 2° и 4° расходящиеся.

Приведем без доказательства и проиллюстрируем примерами некоторые признаки (критерии), позволяющие делать заключение о сходимости или расходимости числовых рядов.

1°. Если числовой ряд $\sum\limits_{k=1}^{+\infty} a_k$ сходится, то $\lim\limits_{k\to+\infty} a_k = 0$.

2°. Если $0 \le a_n \le b_n \ \forall n$, то из сходимости ряда $\sum\limits_{k=1}^{+\infty} b_k$ следует сходимость ряда $\sum\limits_{k=1}^{+\infty} a_k$, а из расходимости ряда $\sum\limits_{k=1}^{+\infty} a_k$ вытекает расходимость ряда $\sum\limits_{k=1}^{+\infty} b_k$.

3°. Пусть функция $f(x) \ge 0$ при всех $x \in [1, +\infty)$. Тогда несобственный интеграл $\int\limits_1^{+\infty} f(x)\,dx$ и ряд $\sum\limits_{k=1}^{\infty} f(k)$ сходятся или расходятся одновременно.

4°. (Признак *Лейбница.*) Пусть $\forall k \ : \ a_k \ge 0$, тогда для сходимости ряда $\sum\limits_{k=0}^{+\infty} (-1)^k a_k$ *достаточно,* чтобы $\forall k \ : \ a_k \ge a_{k+1}$ и $\lim\limits_{k\to+\infty} a_k = 0$.

Тот факт, что свойства выражений с бесконечным числом слагаемых могут отличаться от обычных сумм, иллюстрирует следующий пример. Рассмотрим числовой ряд (сходящийся, как следует из п.4°, по признаку Лейбница к некоторому числу A).

$$\sum_{k=0}^{+\infty} \frac{(-1)^k}{k+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots = A\ .$$

Изменим в записи ряда порядок суммирования. Объединим слагаемые в группы по три следующим образом

$$\left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \cdots + \left(\frac{1}{2n-1} - \frac{1}{4n-2} - \frac{1}{4n}\right) + \cdots\ .$$

Если в каждой паре скобок выполнить первое вычитание, то мы получим

$$\left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{6} - \frac{1}{8}\right) + \cdots + \left(\frac{1}{4n-2} - \frac{1}{4n}\right) + \cdots =$$

$$= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots\right) = \frac{1}{2}\sum_{k=0}^{+\infty} \frac{(-1)^k}{k+1} = \frac{A}{2}\,.$$

То есть перегруппировка слагаемых может привести к изменению суммы ряда.

Рассмотрим еще несколько примеров, иллюстрирующих свойства рядов.

Пример 6.4.1.1. Числовой ряд

$$\sum_{k=0}^{\infty} \sin k = 0 + \sin 1 + \sin 2 + \sin 3 + \ldots\ ,$$

в силу критерия 1° расходится, поскольку $\lim\limits_{k\to+\infty} \sin k \neq 0$.

Пример 6.4.1.2. Числовой ряд

$$\sum_{k=1}^{\infty} \frac{1}{2^k + k} = \frac{1}{2^1+1} + \frac{1}{2^2+2} + \frac{1}{2^3+3} + \frac{1}{2^4+4} + \ldots\ ,$$

является сходящимся, так как для любого номера k справедливо неравенство

$$\frac{1}{2^k + k} \leq \frac{1}{2^k}\,.$$

При этом ряд

$$\sum_{k=1}^{\infty} \frac{1}{2^k} = \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \ldots$$

сходится как сумма членов бесконечно убывающей геометрической прогрессии с $b = q = \dfrac{1}{2}$, но тогда по критерию 2° будет сходиться и исходный ряд.

Пример 6.4.1.3. Выяснить при каких значениях параметра a будет сходиться ряд

$$\sum_{k=1}^{\infty} \frac{1}{k^a\sqrt{k}} = 1 + \frac{1}{2^a\sqrt{2}} + \frac{1}{3^a\sqrt{3}} + \frac{1}{4^a\sqrt{4}} + \ldots\ ,$$

Согласно критерию 3° ряд

$$\sum_{k=1}^{\infty} \frac{1}{k^a \sqrt{k}}$$

будет сходиться или расходиться одновременно с несобственным интегралом

$$\int_{1}^{+\infty} \frac{1}{x^a \sqrt{x}}\, dx = \int_{1}^{+\infty} \frac{dx}{x^{a+\frac{1}{2}}}\,.$$

Используя решение примера 6.3.3, можно утверждать, что последний интеграл сходится при $a + \dfrac{1}{2} > 1$, то есть при $a > \dfrac{1}{2}$, и соответственно расходится при $a \le \dfrac{1}{2}$. Значит данный числовой ряд также сходится при $a > \dfrac{1}{2}$ и расходится при $a \le \dfrac{1}{2}$.

Пример 6.4.1.4. Числовой ряд

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{\sqrt{k}} = 1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} + \ldots$$

сходится в силу признака Лейбница (4°), поскольку последовательность $\left\{ a_k = \dfrac{1}{\sqrt{k}} \right\}$ неотрицательная и монотонно убывающая при $k \to +\infty$.

6.4.2. Функциональные ряды

В §3.2.5 были рассмотрены *функциональные последовательности,* то есть числовые последовательности вида $\{f_n(x)\}$ или $f_1(x), f_2(x), \ldots f_n(x), \ldots$, задаваемые при помощи функций $f_n(x)$, $\forall x \in$

$[a, b]$. Будем говорить, что функция $F(x)$ является предельной для сходящейся функциональной последовательности $\{f_n(x)\}$, если

$$\lim_{n\to+\infty} f_n(x) = F(x),\ \forall x \in [a, b]\ .$$

Аналогично случаю числовой последовательности, используя понятие функциональной последовательности, можно ввести понятие функционального ряда.

Определение 6.4.2.1. Пусть дана некоторая функциональная последовательность $\{f_n(x)\}$. Тогда выражение вида

$$f_1(x) + f_2(x) + f_3(x) + \ldots + f_n(x) + \ldots = \sum_{k=1}^{\infty} f_k(x)$$

называется *функциональным рядом,* а функции

$$f_1(x), f_2(x), f_3(x), \ldots f_n(x), \ldots$$

– *членами функционального ряда.*

Определение 6.4.2.2. Суммы первых n членов функционального ряда $\sum\limits_{k=1}^{\infty} f_k(x)$

$$S_n(x) = f_1(x) + f_2(x) + f_3(x) + \ldots + f_n(x) = \sum_{k=1}^{n} f_k(x)$$

называются *частичными суммами* этого ряда.

Определение 6.4.3.3. Если существуют $\lim\limits_{n\to+\infty} S_n(x) = F(x),\ \forall x \in [a, b]$, то функциональный ряд $\sum\limits_{k=1}^{\infty} f_k(x)$ называется *сходящимся,* а функция $F(x)$ – *суммой ряда* или же его *предельной функцией.*

Очевидно, что все сформулированные ранее свойства числовых рядов будут справедливы и для функциональных рядов, ибо для каждого фиксированного $x_0 \in [a, b]$ рассматриваемый функциональный ряд превращается в числовой. Поэтому далее ограничимся лишь примерами использования функциональных рядов.

Основная область применения функциональных рядов – аппроксимация (то есть запись при помощи *приближенной,* но *простой* формулы) функций, которые либо принципиально не представимы при помощи основных элементарных функций, или же когда такое представление существует, но его практическое использование оказывается черезмерно сложным.

Существуют два основных типа задач представления функций при помощи функциональных рядов:

а) аппроксимация степенными функциями в некоторой (вообще говоря, не малой) окрестности фиксированной точки;

б) аппроксимация при помощи полной (чаще всего, тригонометрической) системы функций на некотором промежутке.

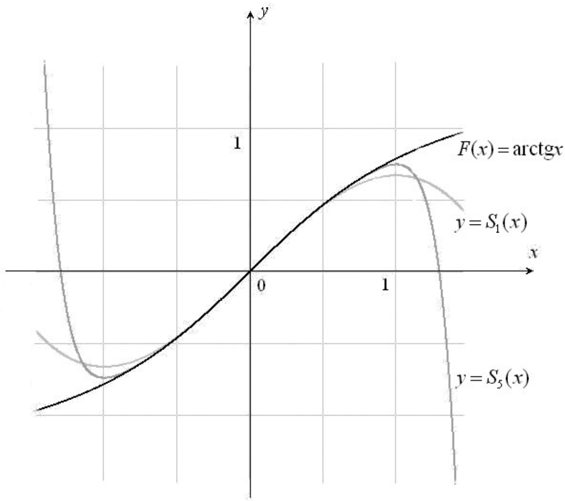


Рис. 6.8. Аппроксимация функции $\operatorname{arctg} x$ частичными суммами ряда.

Проиллюстрируем оба этих подхода следующими примерами.

Пример 6.4.2.1. Для функции $F(x) = \operatorname{arctg} x$ в любой точке x интервала $(-1, 1)$ будет справедливо

$$\operatorname{arctg} x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1} \;=\; x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots .$$

Заметим, что правую часть равенства

$$\Big(\operatorname{arctg} x\Big)' = \frac{1}{1+x^2}$$

при $x \in (-1, 1)$ можно рассматривать как сумму всех членов бесконечно убывающей геометрической прогрессии с первым членом равным 1 и знаменателем $q = -x^2$. То есть будет верно соотношение

$$\Big(\operatorname{arctg} x\Big)' = 1 - x^2 + x^4 - x^6 + x^8 - \ldots = \sum_{k=0}^{\infty} (-1)^k x^{2k} \ .$$

Взяв неопределенный интеграл от его обеих частей, получим

$$\operatorname{arctg} x = \sum_{k=0}^{\infty} (-1)^k \int x^{2k}\, dx = C + \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1} \ .$$

Поскольку $\operatorname{arctg} 0 = 0$, то и $C = 0$.

Таким образом мы приходим к равенству

$$\operatorname{arctg} x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1} \ ,$$

которое иллюстрируется рис.6.8, позволяющим сравнить график функции $F(x) = \operatorname{arctg} x$ с графиками зависимости от x *частичных сумм* (см. определение 6.4.1.2) полученного аппроксимирующего степенного ряда для двух случаев, соответственно с $n = 1$

$$S_1(x) = x - \frac{x^3}{3}$$

и при $n = 5$

$$S_5(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \frac{x^{11}}{11} \ .$$

Для справки заметим, что существуют обширные таблицы, содержащие аппроксимации и других элементарных функций в виде степенных рядов. Например, показательная функция для любого x может быть представлена в виде

$$F(x) = e^x = 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \ , \qquad (6.4.2.1)$$

а логарифмическая функция – рядом

$$F(x) = \ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1} x^k}{k} = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots \ , \qquad (6.4.2.2)$$

который сходится, но только при $|x| < 1$.

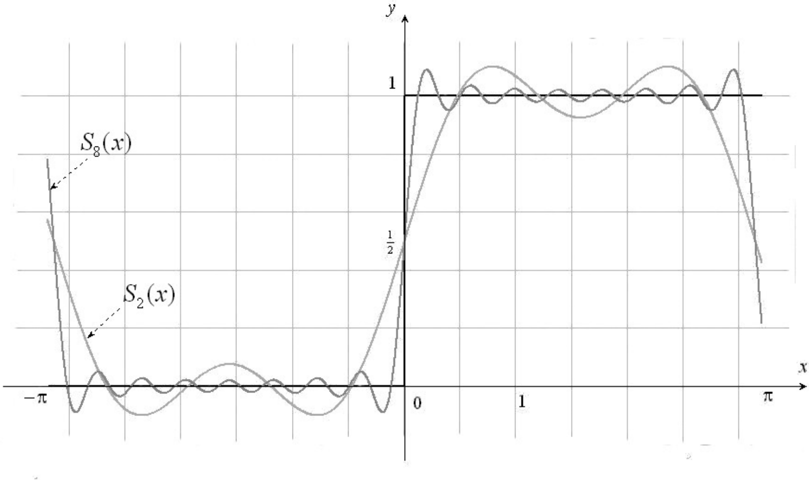


Рис. 6.9. Аппроксимация функции–«ступеньки» рядом Фурье.

Пример 6.4.2.2. Функция–«ступенька»: $F(x) = \left[\begin{array}{l} 0\ ,\ \text{если}\ -\pi < x < 0\ , \\ 1\ ,\ \text{если}\quad 0 < x < \pi\ , \end{array}\right.$ в любой точке $x \in (-\pi, 0) \bigcup (0, \pi)$ аппроксимируется тригонометрическим рядом (называемым часто *рядом Фурье*) по формуле

$$F(x) = \frac{1}{2} + \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{\sin(2k-1)x}{2k-1} =$$

$$= \frac{1}{2} + \frac{2}{\pi} \sin x + \frac{2}{3\pi} \sin 3x + \frac{2}{5\pi} \sin 5x + \ldots\ .$$

Эту формулу иллюстрирует рис.6.9, позволяющий сравнить график функции $F(x)$ с графиками зависимости от x частичных сумм ряда Фурье для $n = 2$, когда

$$S_2(x) = \frac{1}{2} + \frac{2}{\pi} \sin x + \frac{2}{3\pi} \sin 3x$$

и при $n = 8$, когда

$$S_8(x) = \frac{1}{2} + \frac{2}{\pi} \sin x + \frac{2}{3\pi} \sin 3x + \frac{2}{5\pi} \sin 5x + \frac{2}{7\pi} \sin 7x +$$

$$+\frac{2}{9\pi}\sin 9x+\frac{2}{11\pi}\sin 11x+\frac{2}{13\pi}\sin 13x+\frac{2}{15\pi}\sin 15x\ .$$

Обратите внимание, что, во–первых, хотя функция–«ступенька» и разрывна при $x=0$, аппроксимирующий ее ряд в этой точке сходится к значению $\frac{1}{2}$. Во–вторых, аппроксимация функции в примере 6.4.2.2 выполняется не в некоторой окрестности точки области определения (как в примере 6.4.2.1), а на более сложном множестве – объединении двух интервалов: $(-\pi,0)\bigcup(0,\pi)$.

В заключение укажем еще на одну, практически важную область применения рядов: представления функций, которые невозможно задать в виде формулы, записываемой лишь при помощи основных элементарных функций.

Пример 6.4.2.3. Рассмотрим так называемую *функцию Лапласа*

$$\Phi(x)=\sqrt{\frac{2}{\pi}}\int\limits_0^x e^{-\dfrac{t^2}{2}}\,dt\ , \tag{6.4.2.3}$$

часто используемую в теории вероятностей и математической статистике, значения которой обычно приводятся в справочных таблицах.

Хотя эта функция задается при помощи «неберущегося» интеграла (то есть первообразная подынтегральной функции не выражается через основные элементарные функции), можно попытаться представить $\Phi(x)$ в виде некоторого степенного ряда, выполнив, например, следующие преобразования, основанные на использовании формулы (6.4.2.1).

$$\Phi(x)=\sqrt{\frac{2}{\pi}}\int\limits_0^x e^{-\dfrac{t^2}{2}}\,dt\ =\sqrt{\frac{2}{\pi}}\int\limits_0^x\left(1+\sum_{k=1}^{\infty}\frac{\left(-\dfrac{t^2}{2}\right)^k}{k!}\right)dt\ =$$

$$=\sqrt{\frac{2}{\pi}}\left(\int\limits_0^x dt+\sum_{k=1}^{\infty}\frac{(-1)^k}{k!\,2^k}\int\limits_0^x t^{2k}\,dt\right)=C+\sqrt{\frac{2}{\pi}}\left(x+\sum_{k=1}^{\infty}\frac{(-1)^k\,2^{-k}}{(2k+1)\,k!}\cdot x^{2k+1}\right).$$

Поскольку $\Phi(0)=0$, то константу интегрирования следует положить равной нулю. Поэтому, окончательно

$$\Phi(x) = \sqrt{\frac{2}{\pi}} \left(x + \sum_{k=1}^{\infty} \frac{(-1)^k \, 2^{-k}}{(2k+1) \; k!} \cdot x^{2k+1} \right) \; = \; \sqrt{\frac{2}{\pi}} \left(x - \frac{1}{6} x^3 + \frac{1}{40} x^5 - \ldots \right). \tag{6.4.2.4}$$

Глава 7.

Введение в теорию вероятностей

7.1. Случайные события и вероятности

7.1.1. Случайные события и операции с ними

Теория вероятностей, возникшая как раздел математики в XVII веке, допускает (как и другие разделы) аксиоматическое построение. Однако, принимая во внимание специфику и задачи данного курса, представляется возможным ограничиться менее формальным методом изучения понятий, выводов и количественных соотношений в теории вероятностей.

Вначале уточним базовые понятия *опыта* и *события.* Опыт (испытание, эксперимент) – это возникновение или преднамеренное создание определенного набора условий (быть может умозрительных), приводящих к некоторому исходу (результату). Исход опыта называется событием.

Иногда приходится сталкиваться с ситуацией, когда исход того или иного опыта (испытания) не может быть предсказан с гарантируемой достоверностью. Причиной этого может являться недостаточная информация как об условиях постановки данного опыта, так и о характеристиках его осуществления. Причем эта недостаточность нередко имеет принципиально неустранимый характер, не позволяющая в процессе исследования корректно применять детерминированные схемы причинно-следственного анализа. События такого вида принято называть *случайными.*

Вместе с тем, практика показывает, что достаточно часто исход опыта может быть точно предсказан вне зависимости от наличия или отсутствия информации о нем в полном объеме. Например, событие, заключающееся в реализации в случае выстрела по мишени какой-либо одной из следующих двух ситуаций:

- попадание в мишень;
- промах,

гарантированно будет иметь место, независимо от степени нашей информированности о меткости стрелка, свойствах оружия и т.п.

Также (независимо от уровня нашей информированности) очевидно: событие заключающееся в том, что при выстреле имели место одновременно

- попадание в мишень;
- промах,

никогда не произойдет.

События первого из упомянутых типов принято называть *достоверными,* и обозначать Ω, а второго – *невозможными,* обозначаемыми как $\emptyset$.

Если два события взаимно исключают друг друга, то они называются *несовместными* (например, попадание в цель при выстреле и промах.)

Заметим, что комбинация событий также является событием, а обратное, вообще говоря, неверно. То есть возможна альтернатива – два взаимоисключающих друг друга случая: либо событие представимо как некоторая комбинация других событий, либо такое представление невозможно. События второго вида будем называть *элементарными.*

При исследовании случайных событий оказывается удобным использование следующих операций с ними.

Определение 7.1.1.1. Будем говорить, что наступление события A *благоприятствует* (влечет за собой) B, если из наступления события A следует наступление события B. Данный факт обозначается как $A \to B$

или $A \subseteq B$. В этом случае, каждое элементарное событие, входящее в A будет входить и в B. Наконец, естественно принять, что события A и B *равны* ($A = B$), если одновременно $A \to B$ и $B \to A$, то есть если они состоят из одних и тех же элементарных событий.

Определение 7.1.1.2. *Суммой* (или *объединением*) двух событий A и B называется событие $A+B$ (или $A \bigcup B$), заключающееся в осуществлении *хотя бы одного* из событий A или B. Это определение естественным образом обобщается на случай произвольного числа событий. *Суммой* (или *объединением*) событий $A_1, A_2, \dots A_n$ называется событие $\sum\limits_{k=1}^{n} A_k$ (или $\bigcup\limits_{k=1}^{n} A_k$), заключающееся в осуществлении хотя бы одного из событий $A_1, A_2, \dots A_n$.

Определение 7.1.1.3. *Произведением* (или *пересечением*) двух событий A и B называется событие $A \cdot B$ (или $A \bigcap B$,) заключающееся в осуществлении как события A, так и события B. На случай произвольного числа событий это определение обобщается так *произведением* (или *пересечением*) событий $A_1, A_2, \dots A_n$ называется событие $\prod\limits_{k=1}^{n} A_k$ (или $\bigcap\limits_{k=1}^{n} A_k$), заключающееся в осуществлении *всех* событий $A_1, A_2, \dots A_n$.

Определение 7.1.1.4. Пусть событие $\overline{A}$ состоит в том, что событие A не происходит, тогда эти события называются *противоположными,* а событие $\overline{A}$ – *дополнительным* к событию A.

Определение 7.1.1.5. *Разностью* двух событий A и B называется событие $A \backslash B$ (или $A - B$), состоящее в наступлении события A и ненаступлении события B. Иначе говоря, справедливо равенство $A \backslash B = A \cdot \overline{B}$.

Отметим, что два события A и B являются несовместными тогда и только тогда, когда их произведение является невозможным событием. То есть $A \cdot B = \emptyset$. Например, несовместными являются события A и $\overline{A}$.

Приведем теперь графическую интерпретацию операций с событиями. Пусть событие A заключается в принадлежности некоторой точки плоскости меньшему кругу, а событие B – большему. На рис. 7.1 заштрихованные области графически представляют результат выполнения введенных операций с этими событиями.

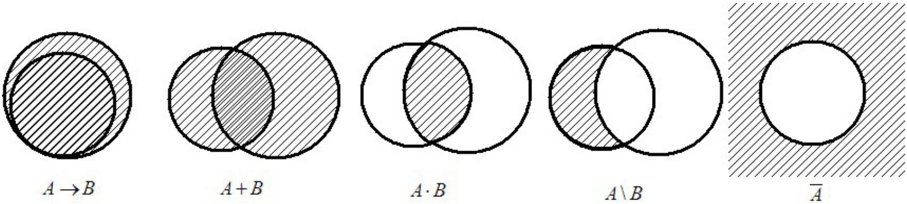


Рис. 7.1. Геометрический пример операций с событиями.

Отметим полезные равенства:

- события A и B несовместны тогда и только тогда, когда $A \cdot B = \emptyset$ (или $A \bigcap B = \emptyset$).
- пусть события $\{A_k \ k = 1, 2, \ldots n\}$ *всевозможные* исходы некоторого испытания, тогда $\sum\limits_{k=1}^{n} A_k = \Omega$.

Для противоположных событий будут справедливы соотношения $A \cdot \overline{A} = \emptyset$, $A + \overline{A} = \Omega$ и $\overline{(\overline{A})} = A$, $\forall A$.

В заключение дадим определение полной группы событий.

Определение 7.1.1.7. Попарно несовместные события образуют *полную группу,* если в результате опыта одно из них обязательно произойдет. Или, в терминах операций с событиями, $\{A_1, A_2, \ldots A_n\}$ – полная группа событий, если

$$\begin{cases} A_i \cdot A_j = \emptyset, \text{ при } i \neq j \ , \\ \sum\limits_{k=1}^{n} A_k = \Omega \ . \end{cases}$$

Пример 7.1.1.1. Проиллюстрируем использование операций с событиями полезным соотношением, называемым *формулой де Моргана.* Для любого набора событий A_1, A_2, $\ldots$ A_N имеет место равенство $\overline{A_1 + A_2 + \ldots + A_N} = \overline{A_1} \cdot \overline{A_2} \cdot \ldots \cdot \overline{A_N}$.

Действительно, пусть элементарное событие $B \subseteq \overline{(A_1 + A_2 + \ldots + A_N)}$, тогда $B \not\subseteq (A_1 + A_2 + \ldots + A_N)$ и

$$\forall k = [1, N] : \ B \subseteq \overline{A_k} \qquad \Rightarrow \qquad B \subseteq \overline{A_1} \cdot \overline{A_2} \cdot \ldots \cdot \overline{A_N} \ .$$

Обратно, пусть $B \subseteq \overline{A_1} \cdot \overline{A_2} \cdot \ldots \cdot \overline{A_N}$, тогда $\forall k = [1, N]:\ B \subseteq \overline{A_k}$ и

$$B \not\subseteq (A_1 + A_2 + \ldots + A_N) \qquad \Rightarrow \qquad B \subseteq \overline{(A_1 + A_2 + \ldots + A_N)}\ .$$

Значит события $\overline{A_1 + A_2 + \ldots + A_N}$ и $\overline{A_1} \cdot \overline{A_2} \cdot \ldots \cdot \overline{A_N}$ равны.

7.1.2. Вероятность случайного события

Помимо «крайних» случаев: достоверных и невозможных событий, существуют, порождаемые в ходе неоднократно проводимого некоторого опыта, случайные события, для которых хотя и нет возможности достоверно предсказать исход каждого конкретного испытания, но наблюдается *количественная устойчивость отношения* числа наступлений определенного исхода к полному числу опытов при проведении достаточно большой серии испытаний.

Средняя величина этого отношения может использоваться как количественная оценка возможности наступления этого исхода при проведении единственного опыта. Данную характеристику принято называть *статистической вероятностью* случайного события.

Поскольку проведение бесконечно большого числа испытаний принципиально невозможно, то статистическая интерпретация понятия вероятности оказывается полезной для оценки ее величины в тех случаях, когда существование вероятности каким-то образом обосновано или принимается априорно.

На практике используются различные определения вероятности случайного события. В нашем курсе мы ограничимся, так называемым, классическим определением. *Классической схемой* (или схемой случаев) называется опыт (испытание), в котором полное число элементарных исходов конечно и все они равновозможны. Элементарный исход называется благоприятствующим событию A, если он влечет за собой наступление этого исхода.

Определение 7.1.2.1. Классической *вероятностью случайного события* A называется число

$$P(A) = \frac{m}{n}$$

являющееся отношением m – числа элементарных исходов, благоприятствующих этому событию A, к n – полному числу элементарных исходов классической схемы.

Пример 7.1.2.1. Рассмотрим в качестве опыта (испытания) бросание игральной кости – симметричного кубика, на каждой из граней которого указано число очков: 1, 2, 3, 4, 5 и 6. После бросания одна из граней кости оказывается верхней и считается выпавшим то число очков, которое указано на этой верхней грани. Найдем вероятность события A, заключающегося в выпадении четного числа очков при однократном бросании кости.

В рассматриваемом опыте шесть равновозможных элементарных исходов, из которых три благоприятствуют событию A. Значит,

$$P(A) = \frac{3}{6} = \frac{1}{2}\,.$$

Иногда классическое определение вероятности удается применить в случаях, когда число элементарных, равновероятных исходов не является конечным.

Пусть каждый *элементарный исход* можно изобразить *точкой* некоторого отрезка X длины $L(X)$ на прямой (или фигуры площади $S(X)$ на координатной плоскости, или же тела объема $V(X)$ в пространстве) , а элементарный *исход, благоприятствующий* событию A – *точкой* другого отрезка (фигуры или тела) Y, содержащегося в первом (то есть $Y \subseteq X$), имеющего длину $L(Y)$ (соответственно, площадь $S(Y)$ или объем $V(Y)$.) Тогда для подсчета вероятности события A можно использовать формулы

$$P(A) = \frac{\text{Длина отрезка } (Y)}{\text{Длина отрезка } (X)} = \frac{L(Y)}{L(X)}\,,$$

$$P(A) = \frac{\text{Площадь фигуры } (Y)}{\text{Площадь фигуры } (X)} = \frac{S(Y)}{S(X)}\,, \tag{7.1.2.1}$$

$$P(A) = \frac{\text{Объем тела } (Y)}{\text{Объем тела } (X)} = \frac{V(Y)}{V(X)}\,.$$

В этом случае принято говорить о *геометрической вероятности* события A. Поясним это понятие при помощи следующих задач.

Пример 7.1.2.2. На отрезке $X:\ [-2,5]$ числовой прямой случайным (равновероятным) образом выбирают точку. Какова вероятность события, что выбранная точка изображает число x, не меньшее, чем 3?

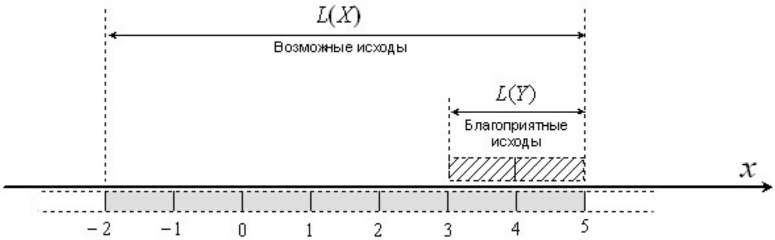


Рис. 7.2. Оценка геометрической вероятности условия $x \geq 3$.

Решение. Благоприятный исход – это событие, состоящее в том, что выбранная точка будет изображать число $x \geq 3$, имеет место, если эта точка (см. рис.7.2) попадает на отрезок $Y: \ [3, 5]$, длина которого равна $L(Y) = 2$. Длина же отрезка, содержащего точки, соответствующие всевозможным исходам, по условию задачи равна $L(X) = 7$. Значит, по первой из формул (7.1.2.1),

$$P(A) = \frac{L(Y)}{L(X)} = \frac{2}{7}\,.$$

Пример 7.1.2.3. Два человека договорились встретиться в определенном месте от 17 до 18 часов. При этом каждый обязался после прихода на место встречи ожидать другого 20 минут. Какова вероятность встречи этих людей, если каждый из них придет на место встречи равновозможно в любой момент указанного интервала времени?

Решение. Введем следующие обозначения: t_1 – время (в часах), через которое, считая от 17 часов, первый человек прибыл к месту встречи; t_2 – время, через которое, второй человек прибыл к месту встречи. На рис. 7.3 событие, состоящее в том, что первый пришел в момент t_1, а второй в момент t_2, изобразится точкой с координатами (t_1, t_2). Условие осуществления события A – того, что один из них дождется другого, имеет вид

$$|t_2 - t_1| \leq \frac{1}{3}\,.$$

Фигура X в данном случае есть квадрат со стороной 1, а фигура Y – затененный вытянутый шестиугольник. Площадь фигуры X очевид-

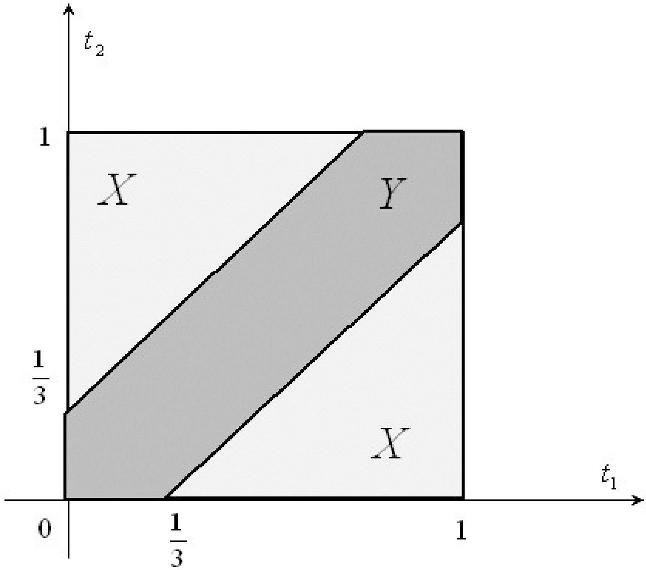


Рис. 7.3. Оценка геометрической вероятности встречи.

но равна 1. Площадь шестиугольника равна площади квадрата минус площади двух равных светло-серых треугольников, каждую из которых обозначим $S_\triangle$. Поэтому, согласно второй из формул (7.1.2.1),

$$P(A) = \frac{S(Y)}{S(X)} = \frac{1 - 2 \cdot S_\triangle}{1} = 1 - 2 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{5}{9}\,.$$

7.1.3. Элементы комбинаторики

Классическое определение вероятности случайного события во многих случаях приводит как к задаче подсчета числа возможных элементарных исходов, так и количества элементарных исходов благоприятствующих некоторому событию. В математике задачи такого класса принято называть *комбинаторными.* Формулировка этих задач, как правило, содержит вопрос: «Сколькими способами...?»

Рассмотрим вначале задачи определения числа подмножеств, образуемых для множеств *не содержащих одинаковых элементов.*

Определение 7.1.3.1. Пусть имеется множество, состоящее из n элементов. Каждую *упорядоченную* выборку из этого множества, содержащую k элементов, называют *размещением* из n элементов по k элементов (иногда говорят: «размещение из n по k»).

В рассматриваемом случае очевидно, что $0 \leq k \leq n$. Число всех размещений из n по k обозначается A_n^k. Если учесть, что при $k = 0$ существует только одно размещение – пустое множество $\emptyset$, то справедливо равенство

$$A_n^k = \left[\begin{array}{ll} 1, & \text{если } k = 0, \\ n(n-1)(n-2)\ldots(n-k+1), & \text{если } k > 0. \end{array} \right.$$

Как известно, произведение всех натуральных чисел от 1 до n принято обозначать

$$1 \cdot 2 \cdot 3 \cdot \ldots \cdot n = n!$$

Тогда формулу полного числа размещений можно записать в виде

$$A_n^k = \frac{n!}{(n-k)!}.$$

Заметим, что $A_0^0 = 1$, поэтому по определению полагают $0! = 1$ и формула для числа размещений будет верной для любых $0 \leq k \leq n$.

Пример 7.1.3.1. Сколькими способами можно трех человек рассадить в двадцать кресел?

Решение. Очевидно, что люди суть объекты разные, поэтому искомое число вариантов равно числу *размещений* из 20 по 3, то есть

$$A_{20}^3 = \frac{20!}{17!} = \frac{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16 \ \cdots \ 3 \cdot 2 \cdot 1}{17 \cdot 16 \ \cdots \ 3 \cdot 2 \cdot 1} = 20 \cdot 19 \cdot 18 = 6840.$$

Определение 7.1.3.2. Размещение из n элементов по n называется *перестановкой* из n элементов.

Число всех перестановок из n элементов $P_n = n!$.

Пример 7.1.3.2. Сколькими способами можно 10 человек рассадить в 10 кресел?

Решение. Искомое число вариантов есть число всевозможных *перестановок* из 10 элементов, то есть

$$P_{10} = 10! = 10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 3\,628\,800 \ .$$

Согласно определению 7.1.3.1, размещения могут отличаться друг от друга как составом своих элементов, так и порядком их следования.

Определение 7.1.3.3. Если порядок следования элементов в подмножестве не существенен, то такая выборка k элементов из n называется *сочетанием* из n элементов по k элементов (иногда говорят: «сочетание из n по k«).

Формула для C_n^k – числа всех сочетаний из n по k имеет вид

$$C_n^k = \frac{n!}{(n-k)! \cdot k!} \,.$$

Заметим, что при использовании этой формулы часто оказывается полезным очевидное равенство $C_n^k = C_n^{n-k}$.

Пример 7.1.3.3. Сколькими способами можно выбрать три кресла из имеющихся двадцати для последующей рассадки в них трех человек?

Решение. В отличие от людей (сравните с примером 7.1.3.1), все кресла одинаковые, поэтому искомое число способов формирования тройки кресел есть число *сочетаний* из 20 по 3 , которое равно

$$C_{20}^3 = \frac{20!}{17! \cdot 3!} = \frac{20 \cdot 19 \cdot 18}{1 \cdot 2 \cdot 3} = 1\,140 \,.$$

Теперь приведем формулы числа размещений, перестановок и сочетаний в случае, когда в состав подмножества-выборки можно включать одинаковые элементы. Чтобы отличать данные два случая, во втором будем использовать обозначения:

- $\hat{A}_n^k$ для числа размещений из n элементов по k ;
- $\hat{P}_n$ для числа перестановок из n элементов ;
- $\hat{C}_n^k$ для числа сочетаний из n элементов по k ;

Число *размещений* из n элементов по k с повторениями будет определяться равенством $\hat{A}_n^k = n^k$, поскольку для подмножества, содержащего k элементов, имеется по n возможностей выбора каждого из этих k элементов.

Если в наборе из n элементов содержится k различных, причем $i-$й из них встречается n_i раз, то общее число *перестановок* следует уменьшить в $n_i!$ раз для каждого $i = 1, 2, \ldots k$, поскольку при перестановке местами одинаковых элементов подмножество не изменится. Значит

$$\hat{P}_n = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!} \,.$$

Наконец, для числа *сочетаний* из n элементов по k будет справедлива формула $\hat{C}_n^k = C_{n+k-1}^k$. Действительно, если предположить, что в выборке из n элементов по k могут быть одинаковые, то выбор выполняется уже не по одному разу для каждого из n элементов, а еще до $k-1$ раз, и искомое число вариантов равно числу сочетаний (без повторений) из $n+k-1$ элементов по k.

Пример 7.1.3.4. Номер банкноты в 1000 рублей определенной серии состоит из семи десятичных цифр. Какое максимальное число банкнот данной серии может быть выпущено в обращение?

Решение. Поскольку каждая цифра номера может иметь одно из десяти значений, то полное число вариантов таких номеров равно $\hat{A}_{10}^7 = 10^7$.

Пример 7.1.3.5. Сколько существует различных шестизначных кодовых комбинаций, содержащих по три цифры 0 и 1 ?

Решение. Цифры 0 и 1 встречаются в комбинации по три раза, поэтому полное число этих комбинаций

$$\hat{P}_6 = \frac{6!}{3! \cdot 3!} = \frac{720}{6 \cdot 6} = 20 .$$

Пример 7.1.3.6. В подарочный набор вкладывают по три шоколадки из имеющихся десяти сортов. Сколько существует вариантов таких подарочных наборов, если набор можно вкладывать шоколадки как одинакового, так и различного сортов?

Решение. Максимальное число вариантов подарочного набора будет

$$\hat{C}_{10}^3 = C_{10+3-1}^3 = C_{12}^3 = \frac{12!}{9! \cdot 3!} = \frac{12 \cdot 11 \cdot 10}{6} = 220 .$$

7.1.4. Вероятности суммы и произведения событий

Из определения 7.1.2.1 очевидно вытекают следующие соотношения.

1°. Для любого случайного события A $\quad 0 \leq P(A) \leq 1$.
2°. $P(\emptyset) = 0$.
3°. $P(\Omega) = 1$.

Основываясь на определении классической вероятности, покажем также, что имеют место следующие соотношения.

Теорема 7.1.4.1. **Вероятность суммы несовместных событий.**

> Вероятность суммы двух *несовместных* событий равна сумме вероятностей этих событий: $P(A+B) = P(A) + P(B)$.

Действительно, пусть события A и B таковы, что из общего числа исходов n событию A благоприятствуют m_A исходов, а событию B – m_B исходов. Поскольку эти события несовместные, то событию $A+B$ будут благоприятствовать $m_A + m_B$ исходов из общего числа исходов n. Тогда, согласно определению 7.1.2.1, будут выполнены равенства

$$P(A+B) = \frac{m_A + m_B}{n} = \frac{m_A}{n} + \frac{m_B}{n} = P(A) + P(B) .$$

Во многих случаях вероятность некоторого события может зависеть от того, произошло другое событие или нет. Иначе говоря, наступление второго события меняет вероятность наступления первого.

Определение 7.1.4.1. Вероятность события A, вычисленная при условии, что произошло событие B, называется *условной вероятностью события* A и обозначается $P_B(A)$.

Убедимся, что формула для подсчета условной вероятности имеет вид

$$P_B(A) = \frac{P(A \cdot B)}{P(B)} .$$

Пусть из n –общего числа исходов некоторого испытания, m благоприятствуют событию B, причем k из числа последних благоприятствуют также и событию A, то есть эти k исходов благоприятствуют событию $A \cdot B$. В силу определения 7.1.2.1 это означает что

$$P(B) = \frac{m}{n}, \quad P_B(A) = \frac{k}{m} \quad \text{и} \quad P(A \cdot B) = \frac{k}{n} . \tag{7.1.4.1}$$

Перемножив почленно первое и второе равенства в (7.1.4.1), получаем

$$P(B) \cdot P_B(A) = \frac{k}{n} = P(A \cdot B)$$

и, следовательно, можно утверждать, что справедлива

Теорема 7.1.4.2. **Вероятность произведения событий.**

Вероятность произведения двух событий равна произведению вероятности одного события на условную вероятность другого, вычисленную при условии, что первое произошло: $P(A \cdot B) = P_B(A) \cdot P(B), \quad P(B) > 0$ или $P(A \cdot B) = P_A(B) \cdot P(A), \quad P(A) > 0$.

Заметим, что возможен случай, когда вероятность наступления события A *не меняется* при наступлении события B. При этом $P(A{\cdot}B) = P(A) \cdot P(B)$ и естественно дать

Определение 7.1.4.2. События A и B называются *независимыми,* если

$$P(A \cdot B) = P(A) \cdot P(B) .$$

Получим теперь формулу, обобщающую утверждение теоремы 7.1.4.1, для вероятности суммы двух событий, не являющихся обязательно несовместными. Напомним, что события называются *совместными,* если появление одного из них в некотором испытании не исключает и появление второго.

Имеет место

Теорема 7.1.4.3. **Вероятность суммы событий.**

Вероятность суммы двух событий определяется по формуле: $P(A + B) = P(A) + P(B) - P(A \cdot B)$.

Доказательство.

В силу определений 7.1.1.2 и 7.1.1.3 событие A произойдет тогда и только тогда, когда произойдет одно из двух *несовместных* событий: $A \cdot \overline{B}$ и $A \cdot B$. В силу теоремы 7.1.4.1 в этом случае будет справедливо равенство $P(A) = P(A \cdot \overline{B}) + P(A \cdot B)$. Из которого следует, что $P(A \cdot \overline{B}) = P(A) - P(A \cdot B)$.

Аналогично, событие B произойдет тогда и только тогда, когда произойдет одно из двух *несовместных* событий: $\overline{A} \cdot B$ и $A \cdot B$. И по теореме 7.1.4.1, будет верно равенство $P(B) = P(\overline{A} \cdot B) + P(A \cdot B)$, в силу которого $P(\overline{A} \cdot B) = P(B) - P(A \cdot B)$.

Наконец, событие $A + B$ произойдет тогда и только тогда, когда произойдет одно из трех попарно *несовместных* событий: $A \cdot \overline{B}$, $\overline{A} \cdot B$ или $A \cdot B$. Поэтому $P(A + B) = P(A \cdot \overline{B}) + P(\overline{A} \cdot B) + P(A \cdot B)$.

Подставляя в последнее равенство выражения $P(A \cdot \overline{B}) = P(A) - P(A \cdot B)$ и $P(\overline{A} \cdot B) = P(B) - P(A \cdot B)$, получаем доказываемое соотношение

$$P(A + B) = P(A) + P(B) - P(A \cdot B) .$$

В случае, когда число событий больше двух (например, N) вероятность их суммы удобно находить по формуле

$$P(A_1 + A_2 + A_3 + \ldots + A_N) = 1 - P\left(\overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3} \cdot \ldots \cdot \overline{A_N}\right)$$

или, что то же самое,

$$P\left(\sum_{k=1}^{N} A_k\right) = 1 - P\left(\prod_{k=1}^{N} \overline{A_k}\right) ,$$

получаемой из формулы де Моргана (см. пример 7.1.1.1).

Продемонстрируем использование теорем 7.1.4.1 – 7.1.4.3 при решении задач.

Пример 7.1.4.1. При каком N – числе последовательных бросаний игральной кости, вероятность события A, заключающегося в том, что хотя бы один раз выпадет шесть очков, будет больше, чем $\dfrac{2}{3}$?

Решение. Пусть событие A_k состоит в том, что шесть очков выпало при $k-$м бросании. Тогда $A = A_1 + A_2 + A_3 + \ldots + A_N$. Найдем вероятность события $B = \overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3} \cdot \ldots \cdot \overline{A_N}$, заключающегося в том, что за N бросаний шесть очков *не выпадут ни разу.* Поскольку бросания являются независимыми событиями, а условия бросания одинаковы для всех из них, то

$$\begin{aligned} P(B) &= P\left(\overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3} \cdot \ldots \cdot \overline{A_N}\right) = \\ &= P\left(\overline{A_1}\right) \cdot P\left(\overline{A_2}\right) \cdot P\left(\overline{A_3}\right) \cdot \ldots \cdot P\left(\overline{A_N}\right) = (1-p)^N , \end{aligned}$$

где $p = \dfrac{1}{6}$ – вероятность выпадения шести очков при одном бросании. Поскольку $P(A) + P(B) = 1$, нам необходимо найти значения N, для которых $P(B) < \dfrac{1}{3}$. Иначе говоря, требуется решить неравенство

$$(1-p)^N < \frac{1}{3} \qquad \text{или} \qquad \left(\frac{5}{6}\right)^N < \frac{1}{3} .$$

Решение (см. п.6° §1.1) имеет вид

$$N > \log_{\frac{5}{6}} \frac{1}{3} = \frac{\ln \frac{1}{3}}{\ln \frac{5}{6}} = \frac{\ln 1 - \ln 3}{\ln 5 - \ln 6} \approx \frac{(-1.097)}{(-0.182)} \approx 6.036 .$$

Наконец, учитывая, что число N натуральное, приходим к заключению, что вероятность, хотя бы однократного выпадения шести очков, оказывается больше $\frac{2}{3}$ при $N \geq 7$.

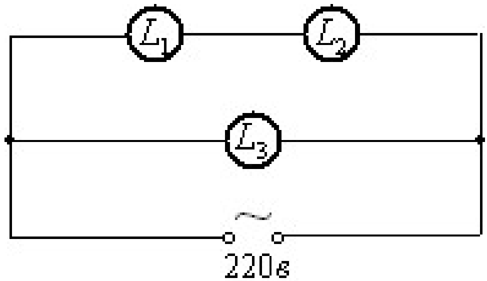


Рис. 7.4. Схема включения электроламп в задаче 7.1.4.2.

Пример 7.1.4.2. Для освещения некоторого помещения используются три электролампы L_1, L_2 и L_3, включенные по схеме показанной на рис. 7.4. Вероятности выхода из строя («перегорания») этих ламп в течение года эксплуатации не зависят друг от друга и равны соответственно $\frac{1}{3}$, $\frac{1}{4}$ и $\frac{2}{5}$. Какова вероятность, что после года эксплуатации это помещение окажется без освещения?

Решение. Пусть событие A_k состоит в том, что лампа с номером k исправно проработает в течение года. Тогда событие $\overline{A_k}$ заключается в том, что лампа с номером k перегорит в течение этого периода. События A_k и $\overline{A_k}$ очевидно образуют полную группу. Заметим также, что лампа не будет гореть либо в случае, когда она вышла из строя, либо, когда ветвь цепи, в которую она включена, разорвана. Последнее означает, например, что в случае перегорания лампы L_1, лампа L_2 гореть не будет, даже, если она исправна.
К концу года для каждой из ламп с номером $k = 1, 2, 3$ произойдет либо событие A_k, либо событие $\overline{A_k}$. Составим таблицу 7.1.4.1 *всевозможных* исходов, записав их формулы при помощи операций с событиями и подсчитав вероятности каждого из исходов.

Для того, чтобы ответить на вопрос задачи, необходимо найти вероятность события, заключающегося в появлении одного из исходов с номерами 4, 6 и 8. Для этого следует применить теорему 7.1.4.1, которая дает

$$P(\text{ Освещения нет }) = \frac{1}{15} + \frac{1}{10} + \frac{1}{30} = \frac{1}{5}\,.$$

Нетрудно заметить, что все восемь исходов попарно несовместны и образуют полную группу, поскольку осуществление *хотя бы одного из них* – достоверное событие. Это позволяет также сделать заключение, что

$$P(\text{ Освещение есть }) = 1 - \frac{1}{5} = \frac{4}{5}\,.$$

(Таблица 7.1.4.1)

Номер исхода	Событие	Его вероятность	Есть ли освещение?
1	$A_1 \cdot A_2 \cdot A_3$	$\frac{2}{3} \cdot \frac{3}{4} \cdot \frac{3}{5} = \frac{3}{10}$	Есть
2	$A_1 \cdot A_2 \cdot \overline{A_3}$	$\frac{2}{3} \cdot \frac{3}{4} \cdot \frac{2}{5} = \frac{2}{10}$	Есть
3	$A_1 \cdot \overline{A_2} \cdot A_3$	$\frac{2}{3} \cdot \frac{1}{4} \cdot \frac{3}{5} = \frac{1}{10}$	Есть
4	$A_1 \cdot \overline{A_2} \cdot \overline{A_3}$	$\frac{2}{3} \cdot \frac{1}{4} \cdot \frac{2}{5} = \frac{1}{15}$	Нет
5	$\overline{A_1} \cdot A_2 \cdot A_3$	$\frac{1}{3} \cdot \frac{3}{4} \cdot \frac{3}{5} = \frac{3}{20}$	Есть
6	$\overline{A_1} \cdot A_2 \cdot \overline{A_3}$	$\frac{1}{3} \cdot \frac{3}{4} \cdot \frac{2}{5} = \frac{1}{10}$	Нет
7	$\overline{A_1} \cdot \overline{A_2} \cdot A_3$	$\frac{1}{3} \cdot \frac{1}{4} \cdot \frac{3}{5} = \frac{1}{20}$	Есть
8	$\overline{A_1} \cdot \overline{A_2} \cdot \overline{A_3}$	$\frac{1}{3} \cdot \frac{1}{4} \cdot \frac{2}{5} = \frac{1}{30}$	Нет

7.1.5. Условные вероятности для полных групп событий

Пусть набор событий $B_1,\ B_2,\ \ldots\ B_N$ образует полную группу, то есть одно из этих событий обязательно происходит, но все они попарно несовместны. И пусть событие A может произойти с любым (но только с одним!) из событий этой полной группы. Тогда, используя операции с событиями, можно написать, что

$$A = A \cdot B_1 + A \cdot B_2 + \ldots + A \cdot B_N = \sum_{k=1}^{N} A \cdot B_k \ .$$

Поскольку слагаемые в этой сумме несовместные события, то будет справедливо равенство (см. теорему 7.1.4.1)

$$P(A) = P(A \cdot B_1) + P(A \cdot B_2) + \ldots + P(A \cdot B_N) = \sum_{k=1}^{N} P(A \cdot B_k) \ .$$

Теперь в каждом из слагаемых применим теорему 7.1.4.2, предполагая, что известны условные вероятности $P_{B_k}(A)$ – вероятности осуществления события A, при условии, что произошло событие B_k. Получим соотношение

$$\begin{aligned} P(A) &= P_{B_1}(A) \cdot P(B_1) + P_{B_2}(A) \cdot P(B_2) + \ldots + P_{B_N}(A) \cdot P(B_N) = \\ &= \sum_{k=1}^{N} P_{B_k}(A) \cdot P(B_k) \ , \end{aligned}$$

называемую *формулой полной вероятности* и обобщающую теорему 7.1.4.2 на случай полной группы событий.

Рассмотрим теперь задачу определения вероятности осуществления события B_k при условии, что событие A произошло. Иными словами, требуется найти условную вероятность вида $P_A(B_k)$.

Из теоремы 7.1.4.2 следует, что

$$P(A \cdot B_k) = P_{B_k}(A) \cdot P(B_k) = P_A(B_k) \cdot P(A) \ .$$

Поэтому при $P(A) > 0$ и всех $k = [1, N]$ будет верно равенство

$$P_A(B_k) = \frac{P_{B_k}(A) \cdot P(B_k)}{P(A)} \ ,$$

которое с учетом формулы полной вероятности можно записать в виде

$$P_A(B_k) = \frac{P_{B_k}(A) \cdot P(B_k)}{P_{B_1}(A) \cdot P(B_1) + P_{B_2}(A) \cdot P(B_2) + \ldots + P_{B_N}(A) \cdot P(B_N)}$$

или

$$P_A(B_k) = \frac{P_{B_k}(A) \cdot P(B_k)}{\sum\limits_{i=1}^{N} P_{B_i}(A) \cdot P(B_i)} .$$

Полученная формула носит название *формулы Байеса.*

При использовании формул полной вероятности и Байеса следует понимать, что условные вероятности $P_{B_k}(A)$ и $P_A(B_k)$ суть *принципиально различные* количественные характеристики случайных событий. Чтобы подчеркнуть это различие, первую из них принято называть *априорной* (доопытной) условной вероятностью, а вторую – *апостериорной* (послеопытной).

Поясним это различие следующим примером.

Пример 7.1.5.1. 1) Найти вероятность попадания в мишень при выстреле из случайно выбранного одного из трех пистолетов. Первый из пистолетов выбирается с вероятностью $\frac{1}{2}$, второй – с вероятностью $\frac{1}{3}$ и третий – с вероятностью $\frac{1}{6}$. Вероятности попадания для каждого пистолета равны $\frac{1}{4}$, $\frac{1}{4}$ и $\frac{1}{2}$ соответственно. 2)С какой вероятностью при выстреле был выбран третий пистолет, если имело место попадание в мишень?

Решение. Пусть события $B_1,\ B_2,\ B_3$ состоят в выборе соответственно первого, второго или третьего пистолета. Событие A – попадание в мишень при выстреле. Для ответа на первый вопрос применим формулу полной вероятности с параметрами:

$$P(B_1) = \frac{1}{2},\ P(B_2) = \frac{1}{3},\ P(B_3) = \frac{1}{6},\ P_{B_1}(A) = \frac{1}{4},\ P_{B_2}(A) = \frac{1}{4},\ P_{B_3}(A) = \frac{1}{2}.$$

Получаем

$$P(A) = \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{6} = \frac{7}{24} .$$

Ответ на второй вопрос можно получить при помощи формулы Байеса.

Поскольку $P(A) = \frac{7}{24}$, а для третьего пистолета

$$P_{B_3}(A) \cdot P(B_3) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12},$$

то третий пистолет (при условии, что попадание в цель произошло!) был выбран с вероятностью

$$P_A(B_3) = \frac{\frac{1}{12}}{\frac{7}{24}} = \frac{2}{7}\,.$$

7.2. Последовательности случайных событий

7.2.1. Схема испытаний Бернулли

На практике достаточно часто возникает необходимость оценить вероятность какого-либо исхода случайного события не для единичного испытания, а для некоторой серии опытов. Рассмотрим основные типы таких серий.

Путь проводятся n *одинаковых* и *независимых друг от друга* опытов, в результате каждого из которых может произойти событие A с постоянной вероятностью p, или же не произойти с вероятностью $q = 1 - p$. Такая последовательность испытаний называется *схемой Бернулли*. Ее примером может служить серия последовательного неоднократного бросания игральной кости (см. примеры 7.1.2.1 и 7.1.4.1), когда событие A заключается, скажем, в выпадении шести очков.

Найдем $P_n(m)$ – вероятность того, что событие A произойдет в серии Бернулли ровно m раз. Предположим вначале, что событие A произошло в первых m испытаниях и не произошло в $n - m$ последних. Поскольку все эти испытания независимые, то вероятность такого события (согласно формуле, содержащейся в определении 7.1.4.2) будет равна $p^m q^{n-m}$. Такая же вероятность будет, например, и для комбинации исходов, когда в первых $n-m$ испытаниях событие A не происходило, а произошло в последних m. Понятно, что это же значение будет иметь данная вероятность и *для любой комбинации исходов,* в которой событие A наступит m раз и не наступит в $n-m$ опытах.

Полное число комбинаций исходов, удовлетворяющих условию задачи, равно (см. §7.1.3) $C_n^m = \dfrac{n!}{(n-m)!\,m!}$, а поскольку все эти исходы *попарно несовместны,* то, в силу теоремы 7.1.4.1, искомая вероятность будет

$$P_n(m) = C_n^m p^m q^{n-m} = \frac{n!}{(n-m)!\,m!}\,p^m q^{n-m}\ . \qquad (7.2.1.1)$$

Эта формула носит название *формулы Бернулли.* Она справедлива для любого значения m такого, что $0 \le m \le n$.

Отметим наиболее важные свойства формулы (7.2.1.1) и следствия из нее.

1°. Вероятность того, что событие A в серии Бернулли произошло *не менее* m раз, дается формулой

$$R_n(m) = \sum_{k=0}^{m} P_n(k) = \sum_{k=0}^{m} \frac{n!}{(n-k)!\,k!}\,p^k q^{n-k}\ . \qquad (7.2.1.2)$$

поскольку для разных k исходы серий Бернулли являются попарно несовместными событиями. Причем в случае, когда $m = n$, последняя формула дает (см. бином Ньютона, п.8° в §1.1)

$$R_n(n) = \sum_{k=0}^{n} P_n(k) = \sum_{k=0}^{n} \frac{n!}{(n-k)!\,k!}\,p^k q^{n-k} = (p+q)^n = 1\ ,$$

поскольку это событие – достоверное.

2°. Можно показать, что m^* – наиболее вероятное число появлений события A в серии Бернулли, состоящей из n испытаний, удовлетворяет системе неравенств $(n+1)p-1 \le\ m^* < (n+1)p$.

3°. Число N испытаний, которое нужно провести в серии Бернулли, для того, чтобы событие A произошло с вероятностью не меньшей чем r, хотя бы один раз, находится по формуле

$$N \ge \frac{\ln(1-r)}{\ln(1-p)}\,.$$

Проверьте эту формулу для примера 7.1.4.1.

Использование формулы Бернулли проиллюстрируем следующими примерами.

Пример 7.2.1.1. При играх с равносильным противником и невозможности ничейного исхода, что вероятнее: 1) выиграть три игры в серии из четырех игр или пять игр в серии из восьми; 2) выиграть не менее, чем три игры в серии из четырех игр, или не менее, чем пять игр в серии из восьми?

Решение. 1) Поскольку противник равносильный, то $p = q = \dfrac{1}{2}$. Тогда вероятность выиграть три игры из четырех, согласно формуле (7.2.1.1) будет

$$P_4(3) = C_4^3 \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^1 = \frac{4!}{1!\,3!} \cdot \frac{1}{2^4} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{1\ \cdot 3 \cdot 2 \cdot 1} \cdot \frac{1}{16} = \frac{1}{4}\,,$$

а вероятность выиграть пять игр из восьми

$$P_8(5) = C_8^5 \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^3 = \frac{7}{32}\,.$$

Значит, более вероятно выиграть три игры из четырех, чем пять игр из восьми.

2) Вероятность выиграть не менее трех игр из четырех, согласно формуле (7.2.1.2) составит

$$R_4(3) = P_4(3) + P_4(4) = \frac{1}{4} + \frac{1}{16} = \frac{5}{16}\,,$$

в то время как вероятность выигрыша не менее пяти игр из восьми будет

$$R_8(5) = P_8(5) + P_8(6) + P_8(7) + P_8(8) = \frac{7}{32} + \frac{7}{64} + \frac{1}{32} + \frac{1}{256} = \frac{93}{256}\,.$$

То есть вероятность выиграть не менее пяти игр из восьми оказывается больше вероятности выиграть не менее трех игр из четырех.

Пример 7.2.1.2. При покупке билетов некоторой лотереи в среднем один из 15 билетов оказывается выигрышным. При покупке какого количества билетов вероятность того, что не менее чем два из них оказались выигрышными будет больше $\dfrac{1}{2}$?

Решение. Пусть было куплено N билетов, каждый из которых с вероятностью $\dfrac{1}{15}$ оказывается выигрышным, а с вероятностью $\dfrac{14}{15}$ – нет.

Рассмотрим процедуру покупки билетов как схему испытаний Бернулли, "успехом"в которых является покупка невыигрышного билета, и найдем вероятность того, что среди купленных билетов либо вообще не окажется выигрышных, либо выигрышным будет только один. Заметим, что это событие является противоположным к событию A, заключающемуся в том, что не менее чем два билета оказались выигрышными.

Вероятность того, что, среди купленных билетов не окажется выигрышных согласно формуле (7.2.1.1) равна

$$P_N(N) = C_N^N \cdot \left(\frac{14}{15}\right)^N \cdot \left(\frac{1}{15}\right)^0 = \frac{N!}{0!\,N!} \cdot \left(\frac{14}{15}\right)^N \cdot \left(\frac{1}{15}\right)^0 = \left(\frac{14}{15}\right)^N ,$$

а вероятность того, что, среди купленных билетов окажется лишь один выигрышный будет

$$P_N(N-1) = C_N^{N-1} \cdot \left(\frac{14}{15}\right)^{N-1} \cdot \left(\frac{1}{15}\right)^1 = \frac{N!}{(N-1)!\,1!} \cdot \left(\frac{14}{15}\right)^{N-1} \cdot \left(\frac{1}{15}\right)^1 = \frac{N}{14} \cdot \left(\frac{14}{15}\right)$$

Тогда

$$P(\overline{A}) = P_N(N) + P_N(N-1) = \left(\frac{14}{15}\right)^N + \frac{N}{14} \cdot \left(\frac{14}{15}\right)^N = \left(\frac{14}{15}\right)^N \cdot \left(1 + \frac{N}{14}\right) .$$

Остается выяснить, при каких значениях N будет справедливо неравенство

$$P(\overline{A}) = \left(\frac{14}{15}\right)^N \cdot \left(1 + \frac{N}{14}\right) < \frac{1}{2} .$$

Поскольку N натуральное число, то проще всего составить таблицу значений $P(\overline{A})$ для различных N

N	5	10	20	24	25	30
$P(\overline{A})$	0.9612	0.8599	0.6111	0.5182	0.4964	0.3065

из которой следует, что вероятность того, что не менее чем два из купленных билетов окажутся выигрышными будет больше $\frac{1}{2}$ при $N \geq 25$.

7.2.2. Формула Пуассона

Ответ на второй вопрос в примере 7.2.1.1 наглядно демонстрирует рост затрат вычислительных усилий при использовании формулы Бернулли, когда значение n увеличивается. При больши́х n использование этой формулы практически невозможно и оказывается необходимым применение альтернативных методов оценки величины $P_n(m)$.

Один из таких методов формулируется как

Теорема 7.2.2.1. **Теорема Пуассона.**

> Если вероятность p – наступления события A, постоянна и достаточно мала, а полное число испытаний n достаточно велико, то имеет место приближенное равенство
>
> $$P_n(m) \approx \frac{\lambda^m \cdot e^{-\lambda}}{m!} \, ,$$
>
> где $\lambda = np$.

Доказательство. По условию теоремы $p = \dfrac{\lambda}{n}$. Подставив это выражение для p в формулу Бернулли, получим

$$P_n(m) = C_n^m \cdot \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} =$$
$$= \frac{\lambda^m}{n^m} \cdot \frac{n(n-1)(n-2)\ldots(n-m+1)}{m!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} =$$
$$= \frac{\lambda^m}{m!} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \ldots \cdot \frac{n-m+1}{n} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} =$$
$$= \frac{\lambda^m}{m!} \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \ldots \cdot \left(1 - \frac{m-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} .$$

Заметим, что при достаточно малом m

$$\lim_{n\to\infty} \left(1 - \frac{k}{n}\right) = 1 \,, \quad \forall k : \ 0 \le k \le m-1$$

и

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-m} = 1 \,,$$

а потому *все скобки, кроме предпоследней,* при достаточно большом n можно считать равными единице. С другой стороны, (см. §3.2.4, так называемый, *второй замечательный предел*)

$$\lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

и оценка вероятности $P_n(m)$ в формуле Бернулли будет иметь вид

$$P_n(m) \approx \frac{\lambda^m \cdot e^{-\lambda}}{m!} \,,$$

что и доказывает теорему Пуассона.

Пример 7.2.2.1. Пивоваренное предприятие произвело и отгрузило заказчику партию в 100 000 бутылок пива. Вероятность того, что одна случайно выбранная из этой партии бутылка окажется разбитой, равна 0,0001. Какова вероятность, что в отгруженной партии число разбитых бутылок окажется равным 1) трем, 2) семи и 3) шестнадцати?

Решение. Поскольку процедуру последовательной проверки всех бутылок (на предмет выявления разбитых) можно рассматривать как схему испытаний Бернулли, в которой событие A заключается в том, что проверяемая бутылка оказывается разбитой, а число таких испытаний (проверок) достаточно велико, то имеет смысл применить формулу Пуассона.

В данном случае значение параметра $\lambda = np = 100000 \cdot 0,0001 = 10$. Найдем величину множителя $e^{-\lambda}$, которую можно подсчитать при помощи калькулятора, взять из справочных таблиц или же оценить ее при помощи основного логарифмического тождества (см. п.6° §1.1). Действительно, зная, что $\lg e \approx 0,434$, легко найти

$$e^{-\lambda} = \left(10^{\lg e}\right)^{-\lambda} = 10^{-\lambda \cdot \lg e} \quad , \text{поскольку } e = 10^{\lg e} \ .$$

В нашем случае

$$e^{-10} = 10^{-10 \cdot \lg e} = 10^{-4,34} \approx 4,5 \cdot 10^{-5} = 0,000045 \ .$$

и, следовательно,

$$P_{100000}^{3} = \frac{10^3 \cdot e^{-10}}{3!} = \frac{1000 \cdot 0,000045}{6} = 0,0076 \ .$$

Аналогично получаем, что $P_{100000}^{7} = 0,0900$ и $P_{100000}^{16} = 0,0220$.

Анализируя решение задачи 7.2.2.1, можно предположить, что существует *наиболее вероятное* число разбитых бутылок. Иначе говоря, для некоторого m значение P_n^m будет иметь максимальное значение. Попробуем оценить это значение m следующим образом: выясним, при каком условии функция $F(\lambda) = \dfrac{\lambda^m \cdot e^{-\lambda}}{m!}$ будет иметь максимум. Для этого найдем значение λ , при котором производная от функции $F(\lambda)$ обращается в ноль (см. §5.4 – *необходимое условие экстремума функции.*)

$$F'(\lambda) = \left(\frac{\lambda^m \cdot e^{-\lambda}}{m!}\right)'_{\lambda} = \frac{1}{m!} \cdot m\lambda^{m-1}e^{-\lambda} + \frac{1}{m!} \cdot \lambda^m \left(-e^{-\lambda}\right) =$$
$$= \frac{\lambda^{m-1}e^{-\lambda}}{m!}(m - \lambda) = 0 \ .$$

Значит, максимум достигается при $\lambda = m$ и наиболее вероятное число разбитых бутылок в партии, состоящей из 100000 (как в примере

7.2.2.1), равно десяти. Эта оценка согласуется с условием, что в среднем из десяти тысяч бутылок одна оказывается разбитой – в партии состоящей из ста тысяч бутылок в среднем десять будут разбитыми.

7.2.3. Цепи Маркова

В схеме Бернулли результат каждого испытания не зависел от результатов других испытаний, однако вполне может оказаться, что такое предположение не соответствует реальности. Поэтому естественно возникает идея обобщения схемы Бернулли, простейший вариант которого заключается в следующем.

Пусть проводится серия испытаний, в каждом из которых обязательно происходит одно из двух несовместных событий A или B. Напомним, что такой набор исходов называется *полной группой событий,* и для них справедливы равенства

$$\begin{cases} A + B = \Omega, \\ A \cdot B = \emptyset. \end{cases}$$

И пусть, кроме того, исход $k+1$-го испытания случайным образом зависит *только от исхода* k-го. Используя введенную ранее терминологию, это предположение можно сформулировать так: существуют условные вероятности (см. §7.1.5)

$$P_{A_k}(A_{k+1}) = p_1^1 \,, \quad P_{B_k}(A_{k+1}) = p_2^1 \,,$$

$$P_{A_k}(B_{k+1}) = p_1^2 \,, \quad P_{B_k}(B_{k+1}) = p_2^2 \,,$$

где события A_k и B_k заключаются в том, что в результате k-го испытания произошло соответственно событие A или событие B.

При этом предполагается, что вероятности p_1^1, p_1^2, p_2^1, p_2^2 *одинаковы для всех* k. Наконец, поскольку события A_k и B_k образуют полную группу, то справедливы равенства

$$\begin{cases} p_1^1 + p_1^2 = 1, \\ p_2^1 + p_2^2 = 1. \end{cases}$$

Серия испытаний описанного вида называется *цепью Маркова.* В случае двух исходов, каждый шаг цепи Маркова описывается квадратной матрицей второго порядка

$$\|S\| = \left\| \begin{array}{cc} p_1^1 & p_1^2 \\ p_2^1 & p_2^2 \end{array} \right\|,$$

называемой *марковской матрицей перехода,* в то время как условные вероятности p_1^1, p_1^2, p_2^1, p_2^2 принято называть *марковскими переходными вероятностями.* [1]

Выясним теперь, как связаны вероятности событий A и B для двух последовательных испытаний в цепи Маркова. Иными словами, найдем выражения для вероятностей $P(A_{k+1})$ и $P(B_{k+1})$ через вероятности $P(A_k)$ и $P(B_k)$.

Согласно формуле полной вероятности (см. §7.1.5), для любого k имеют место соотношения

$$\begin{cases} P(A_{k+1}) = P_{A_k}(A_{k+1}) \cdot P(A_k) + P_{B_k}(A_{k+1}) \cdot P(B_k)\ , \\ \\ P(B_{k+1}) = P_{A_k}(B_{k+1}) \cdot P(A_k) + P_{B_k}(B_{k+1}) \cdot P(B_k) \end{cases}$$

или

$$\begin{cases} P(A_{k+1}) = p_1^1 \cdot P(A_k) + p_2^1 \cdot P(B_k)\ , \\ \\ P(B_{k+1}) = p_1^2 \cdot P(A_k) + p_2^2 \cdot P(B_k)\ . \end{cases}$$

Полученные равенства можно записать в более простой форме, используя операции с матрицами. Действительно

$$\left\| \begin{matrix} P(A_{k+1}) \\ \\ P(B_{k+1}) \end{matrix} \right\| = \left\| \begin{matrix} p_1^1 & p_2^1 \\ \\ p_2^1 & p_2^2 \end{matrix} \right\|^{\mathrm{T}} \cdot \left\| \begin{matrix} P(A_k) \\ \\ P(B_k) \end{matrix} \right\|$$

или же, совсем просто [2]

$$\left\| \begin{matrix} P(A_{k+1}) \\ \\ P(B_{k+1}) \end{matrix} \right\| = \|S\|^{\mathrm{T}} \cdot \left\| \begin{matrix} P(A_k) \\ \\ P(B_k) \end{matrix} \right\| .$$

Заметим, что, зная матрицу перехода $\|S\|$ для одного звена марковской цепи, можно, используя операцию произведения матриц, найти матрицу перехода для любого числа ее звеньев. Например,

$$\left\| \begin{matrix} P(A_4) \\ \\ P(B_4) \end{matrix} \right\| = \|S\|^{\mathrm{T}} \cdot \left\| \begin{matrix} P(A_3) \\ \\ P(B_3) \end{matrix} \right\| = \|S\|^{\mathrm{T}} \cdot \left(\|S\|^{\mathrm{T}} \cdot \left\| \begin{matrix} P(A_2) \\ \\ P(B_2) \end{matrix} \right\| \right) =$$

[1] Обратите внимание, что в записи, например, p_1^2 число 2 является не показателем степени, а индексом – номером одного из событий, образующих полную группу исходов для каждого испытания в марковской серии.

[2] Напомним, что при транспонировании матрицы ее столбцы превращаются в строки с сохранением порядка их следования.

$$= \|S\|^{\mathrm{T}} \cdot \left(\|S\|^{\mathrm{T}} \cdot \left(\|S\|^{\mathrm{T}} \cdot \left\| \begin{array}{c} P(A_1) \\ \\ P(B_1) \end{array} \right\| \right) \right) .$$

Откуда следует, что

$$\left\| \begin{array}{c} P(A_4) \\ \\ P(B_4) \end{array} \right\| = \left(\|S\|^{\mathrm{T}} \right)^3 \cdot \left\| \begin{array}{c} P(A_1) \\ \\ P(B_1) \end{array} \right\| .$$

И нетрудно придти к заключению, что $\|S_k\|$ – матрица перехода между первым и k-м испытанием в марковской серии, определяется равенством

$$\|S_k\|^{\mathrm{T}} = \underbrace{\|S\|^{\mathrm{T}} \cdot \|S\|^{\mathrm{T}} \cdot \ldots \cdot \|S\|^{\mathrm{T}}}_{k-1} = \left(\|S\|^{\mathrm{T}} \right)^{k-1} .$$

Примером цепи Маркова, в которой полная группа исходов каждого испытания может состоять более чем из двух событий, служит последовательность наблюдений за колонией жуков в «задаче о жуках» (см. §2.3). Действительно, если воспользоваться статистическим определением вероятности, то для каждого жука, условные вероятности быть обнаруженным в процессе наблюдения в одной из трех возможных сред обитания, зависят только от его «местонахождения» в момент предыдущего наблюдения и имеют значения:

$$\begin{array}{ll}
P(\textit{на берегу} \to \textit{на берегу}) & = p_1^1 = 0.3 \, , \\
P(\textit{на берегу} \to \textit{в воздухе}) & = p_1^2 = 0.5 \, , \\
P(\textit{на берегу} \to \textit{на воде}) & = p_1^3 = 0.4 \, , \\
P(\textit{в воздухе} \to \textit{на берегу}) & = p_2^1 = 0.4 \, , \\
P(\textit{в воздухе} \to \textit{в воздухе}) & = p_2^2 = 0.1 \, , \\
P(\textit{в воздухе} \to \textit{на воде}) & = p_2^3 = 0.2 \, , \\
P(\textit{на воде} \to \textit{на берегу}) & = p_3^1 = 0.3 \, , \\
P(\textit{на воде} \to \textit{в воздухе}) & = p_3^2 = 0.4 \, , \\
P(\textit{на воде} \to \textit{на воде}) & = p_3^3 = 0.4 \, ,
\end{array}$$

считая, что состояние «на берегу» имеет номер 1, состояние «в воздухе» – номер 2, а состояние «на воде» – номер 3. При этом квадратную матрицу *третьего порядка*

$$\|A\| = \left\| \begin{array}{ccc} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{array} \right\| = \left\| \begin{array}{ccc} p_1^1 & p_1^2 & p_1^3 \\ p_2^1 & p_2^2 & p_2^3 \\ p_3^1 & p_3^2 & p_3^3 \end{array} \right\|^{\mathrm{T}}$$

в уравнении (2.3.2) мы рассматриваем как транспонированную матрицу перехода $\|S\|$ от одного состояния колонии жуков к другому. Поскольку каждый жук *обязательно* обнаруживается при наблюдении *только в одной* из трех возможных сред обитания, то эти события (обнаружить жука «на берегу», обнаружить жука «в воздухе» и обнаружить жука «на воде<<) образуют полную группу. Связь между состояниями колонии для двух последовательных наблюдений с номерами k и $k+1$ при помощи матричных операций записывается в виде

$$\left\| \begin{array}{c} x_{1,k+1} \\ x_{2,k+1} \\ x_{3,k+1} \end{array} \right\| = \|S\|^{\mathrm{T}} \cdot \left\| \begin{array}{c} x_{1,k} \\ x_{2,k} \\ x_{3,k} \end{array} \right\| = \left\| \begin{array}{ccc} p_1^1 & p_1^2 & p_1^3 \\ p_2^1 & p_2^2 & p_2^3 \\ p_3^1 & p_3^2 & p_3^3 \end{array} \right\|^{\mathrm{T}} \cdot \left\| \begin{array}{c} x_{1,k} \\ x_{2,k} \\ x_{3,k} \end{array} \right\| =$$

$$= \left\| \begin{array}{ccc} 0.3 & 0.5 & 0.4 \\ 0.4 & 0.1 & 0.2 \\ 0.3 & 0.4 & 0.4 \end{array} \right\| \cdot \left\| \begin{array}{c} x_{1,k} \\ x_{2,k} \\ x_{3,k} \end{array} \right\|, \qquad (7.2.3.1)$$

где $x_{j,k}$ – число жуков, обнаруженных в среде номер j в процессе k-го наблюдения.

Наконец отметим следующее интересное свойство данной цепи Маркова. В примере 3.2.5.3 (см. §3.2.5) показано, что при достаточно большом числе наблюдений доли общего числа жуков, обнаруживаемых в разных средах обитания, перестают меняться. Это не означает, что жуки «угомонились», они по-прежнему продолжают менять среду обитания в соответствии с формулами (7.2.3.1), но эти изменения *взаимно компенсируются.* Иначе говоря, существует некоторое предельное распределение жуков по средам

$$\left\| \begin{array}{c} x_1^\star \\ x_2^\star \\ x_3^\star \end{array} \right\| = \lim_{k\to\infty} \left\| \begin{array}{c} x_{1,k} \\ x_{2,k} \\ x_{3,k} \end{array} \right\| .$$

В примере 3.2.5.3 также показано, что при использованных конкретных значениях вероятностей перехода

$$\left\| \begin{array}{c} x_1^\star \\ x_2^\star \\ x_3^\star \end{array} \right\| = \left\| \begin{array}{c} 92 \\ 60 \\ 86 \end{array} \right\|$$

и, кроме того, этот столбец является (см. §2.4) *собственным вектором* матрицы $\|S\|^{\mathrm{T}}$, соответствующим *собственному значению* $\lambda = 1$.

Анализ марковских цепей, так же как и использование схемы испытаний Бернулли, позволяет получать оценки вероятностей различных серий случайных событий. Рассмотрим следующую задачу.

Пример 7.2.3.1. Два стрелка A и B стреляют поочередно по мишени, следуя правилу: при попадании стрелок получает право на внеочередной выстрел; при промахе – право выстрела переходит к противнику. Вероятности попадания и промаха стрелка A равны соответственно $p_A^+ = \frac{3}{5}$ и $p_A^- = \frac{2}{5}$, а для стрелка B они равны $p_B^+ = \frac{4}{7}$ и $p_B^- = \frac{3}{7}$. Кто и с какой вероятностью будет делать второй и третий выстрелы, если первым стрелял A?

Решение. Данное соревнование представляет собой марковскую цепь с матрицей перехода для одного шага

$$\|S\| = \left\| \begin{matrix} \frac{3}{5} & \frac{2}{5} \\ \\ \frac{3}{7} & \frac{4}{7} \end{matrix} \right\| .$$

Пусть $P(A_k)$ – вероятность того, что k-й выстрел делает стрелок A, а $P(B_k)$ – вероятность того, что k-й выстрел делает стрелок B. Тогда по формуле полной вероятности (см. §7.1.5)

$$\begin{aligned} P(A_{k+1}) &= \frac{3}{5} \cdot P(A_k) + \frac{3}{7} \cdot P(B_k) \ , \\ P(B_{k+1}) &= \frac{2}{5} \cdot P(A_k) + \frac{4}{7} \cdot P(B_k) \ . \end{aligned} \tag{7.2.3.2}$$

В матричной форме соотношения (7.2.3.2) имеют вид

$$\left\| \begin{matrix} P(A_{k+1}) \\ \\ P(B_{k+1}) \end{matrix} \right\| = \left\| \begin{matrix} \frac{3}{5} & \frac{3}{7} \\ \\ \frac{2}{5} & \frac{4}{7} \end{matrix} \right\| \cdot \left\| \begin{matrix} P(A_k) \\ \\ P(B_k) \end{matrix} \right\| .$$

По условию $P(A_1) = 1$ и $P(B_1) = 0$, поэтому $P(A_2) = \frac{3}{5}$ и $P(B_2) = \frac{2}{5}$.

Для третьего выстрела по формулам (7.2.3.2) находим

$$P(A_3) = \frac{3}{5} \cdot P(A_2) + \frac{3}{7} \cdot P(B_2) = \frac{9}{25} + \frac{6}{35} = \frac{93}{175}\,,$$

$$P(B_3) = \frac{2}{5} \cdot P(A_2) + \frac{4}{7} \cdot P(B_2) = \frac{6}{25} + \frac{8}{35} = \frac{82}{175}\,.$$

Таким образом, второй выстрел стрелки A и B делают соответственно с вероятностями $\frac{3}{5}$ и $\frac{2}{5}$, а третий выстрел – с вероятностями $\frac{93}{175}$ и $\frac{82}{175}$.

Более детальный анализ задачи 7.2.3.1 позволяет заключить, что при неограниченном продолжении соревнования вероятности выстрелов для A и B становятся постоянными (не зависящими от номера выстрела). Действительно, пусть

$$\lim_{k\to\infty} \left\| \begin{array}{c} P(A_k) \\ P(B_k) \end{array} \right\| = \left\| \begin{array}{c} P(A) \\ P(B) \end{array} \right\| .$$

Тогда, переходя к пределу при $k \to \infty$ в равенствах (7.2.3.2), получим

$$\begin{cases} P(A) = \dfrac{3}{5} \cdot P(A) + \dfrac{3}{7} \cdot P(B)\,, \\ \\ P(B) = \dfrac{2}{5} \cdot P(A) + \dfrac{4}{7} \cdot P(B)\,. \end{cases} \qquad (7.2.3.3)$$

или в матричном виде

$$\left\| \begin{array}{cc} \dfrac{3}{5} & \dfrac{3}{7} \\ \\ \dfrac{2}{5} & \dfrac{4}{7} \end{array} \right\| \cdot \left\| \begin{array}{c} P(A) \\ P(B) \end{array} \right\| = \left\| \begin{array}{c} P(A) \\ P(B) \end{array} \right\| .$$

Если сравнить это матричное уравнение с формулой (2.4.1) §2.4, то можно заметить, что столбец $\left\| \begin{array}{c} P(A) \\ P(B) \end{array} \right\|$ является *собственным вектором* матрицы

$$\left\| \begin{array}{cc} \dfrac{3}{5} & \dfrac{3}{7} \\ \\ \dfrac{2}{5} & \dfrac{4}{7} \end{array} \right\| ,$$

который отвечает *собственному значению* $\lambda = 1$.

Нетрудно также видеть, что система уравнений (7.2.3.3) приводит к условию

$$\frac{2}{5}P(A) = \frac{3}{7}P(B),$$

а поскольку $P(A)$ и $P(B)$ суть вероятности событий, образующих *полную группу* (см. определение 7.1.1.6), то $P(A) + P(B) = 1$, что окончательно дает

$$P(A) = \frac{15}{29} \quad \text{и} \quad P(B) = \frac{14}{29}\,.$$

Интерпретируя статистически полученные предельные значения вероятностей, можно сказать, что A стреляет несколько чаще, чем B. Это согласуется с условием задачи, поскольку вероятность попадания при одном выстреле у A больше, чем у B. Действительно,

$$p_A^+ = \frac{3}{5} = \frac{21}{35} > \frac{20}{35} = \frac{4}{7} = p_B^+\,.$$

Глава 8.

Случайные величины

8.1. Дискретные случайные величины

8.1.1. Определение дискретной случайной величины

Если некоторая количественная характеристика ξ (читается – «кси») может принимать различные значения, причем событие, заключающееся в том, что эта характеристика оказалась равной некоторому конкретному числу, является *случайным*, то ξ принято называть *случайной величиной.*

Определение 8.1.1.1 Случайная величина называется *дискретной,* если все ее значения можно пронумеровать. Иначе она называется *непрерывной.*

Например, при пяти последовательных бросаниях симметричной монеты, число «выпадений герба вверх» является случайной величиной, могущей принять одно из шести значений: $0, 1, 2, 3, 4, 5$. Случайными величинами также могут быть: число бракованных изделий в составе одной партии продукции определенного вида, время ожидания пассажиром поезда метро, величина отклонения точки попадания пули в мишень от центра мишени и т.п. Заметим, что в приведенных примерах первые две случайные величины дискретные, а две последние – непрерывные.

Число значений, принимаемых дискретной случайной величиной при проведении некоторого испытания может оказаться хотя и счетным, но *неограниченным.* К примерам дискретной случайной величины с неограниченным множеством значений можно отнести: число последовательных выстрелов в мишень «до первого промаха» или же число последовательных бросаний симметричной монеты, при которых происходит лишь выпадение «герба вверх».

8.1.2. Закон распределения дискретной случайной величины

Пусть случайная величина ξ может принимать одно из N значений x_k, где $k = [1, N]$, то есть могут происходить случайные события вида $(\xi = x_k)$, для которых *существуют* вероятности $P(\xi = x_k) = p_k$. Тогда для описания случайной величины ξ можно, например, использовать таблицу

Значение ξ	x_1	x_2	$\ldots$	x_N
Вероятность события $(\xi = x_k)$	p_1	p_2	$\ldots$	p_N

содержащую перечисление (в порядке возрастания и без повторений) *всех* возможных значений ξ и соответствующих им вероятностей событий $(\xi = x_k)$. В этом случае будем говорить, что задан *закон распределения вероятности значений случайной величины* ξ (или просто, ее *распределение*) в табличной форме. В общем же случае можно дать

Определение 8.1.2.1. *Законом распределения вероятности значений случайной величины* называется соотношение (или правило), сопоставляющее каждому ее возможному значению, величину вероятности события, заключающемуся в том, что данная случайная величина принимает именно это значение.

Распределение случайной величины является наиболее полной ее характеристикой, поэтому на практике часто саму случайную величину задают ее распределением. На практике используются три основных способа задания (или описания) распределения случайной величины:

– таблица распределения;
– формула распределения;
– функция распределения.

Рассмотрим эти способы подробнее на следующих примерах.

Пример 8.1.2.1. Табличная форма записи закона распределения случайной величины ξ, равной числу выпавших очков при бросании игральной кости, имеет очевидную форму

Число выпавших очков	1	2	3	4	5	6
Вероятность этого события	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Пример 8.1.2.2. Найти закон распределения случайной величины ξ, равной числу проросших семян из числа трех посеянных, если вероятность прорасти для одного семени равна $\frac{2}{3}$.

Решение. Очевидно, что случайная величина ξ – число проросших семян, может принять одно из следующих четырех значений: 0, 1, 2 или 3. Поскольку все семена прорастают независимо друг от друга и притом с равными вероятностями, то можно использовать схему испытаний Бернулли. По формуле (7.2.1.1) имеем

$$P(\xi = 0 : \text{ни одно семя не проросло }) \;=\; C_3^0\left(\frac{2}{3}\right)^0\left(\frac{1}{3}\right)^3 \;=\; 1\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{3} \;=\; \frac{1}{27}.$$

Аналогично

$$P(\xi = 1 : \text{проросло только одно семя }) \;=\; C_3^1\left(\frac{2}{3}\right)^1\left(\frac{1}{3}\right)^2 \;=\; 3\cdot\frac{2}{3}\cdot\frac{1}{9} \;=\; \frac{2}{9}.$$

$$P(\xi = 2 : \text{проросли два из трех семян }) \;=\; C_3^2\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^1 \;=\; \frac{3\cdot 2}{2}\cdot\frac{4}{9}\cdot\frac{1}{3} \;=\; \frac{4}{9}.$$

$$P(\xi = 3 : \text{проросли все семена }) \;=\; C_3^3\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^0 \;=\; \frac{3\cdot 2\cdot 1}{1\cdot 2\cdot 3}\cdot\frac{8}{27} \;=\; \frac{8}{27}.$$

Поэтому таблица, задающая закон распределения, будет такой

Число проросших семян	0	1	2	3
Вероятность этого события	$\frac{1}{27}$	$\frac{2}{9}$	$\frac{4}{9}$	$\frac{8}{27}$

Запись закона распределения случайной величины в виде таблицы удобна лишь в случае небольших значений N. Если N велико или количество значений, которые может принимать дискретная случайная величина, не ограничено, то для описания или задания случайных величин можно использовать или *формулу,* связывающую вероятность некоторого значения случайной величины с его номером, или особую функцию, называемую *функцией распределения.*

Следующие примеры иллюстрируют формульный способ описания случайной величины.

Пример 8.1.2.3. Распределение числа очков ξ, выпавших при однократном бросании игральной кости, задается формулой

$$P(\xi = k) = \frac{1}{6}\,, \quad \forall\, k = 1, 2, 3, 4, 5, 6\ .$$

В этом примере ξ равновероятно принимает одно из шести возможных значений.

Пример 8.1.2.4. Распределение ξ – числа «успехов» в серии из n испытаний Бернулли (с вероятностью успеха при однократном испытании, равном p) имеет вид

$$P_n(\xi = k) = C_n^k\cdot p^k\cdot(1-p)^{n-k} = \frac{n!}{k!\cdot(n-k)!}\cdot p^k\cdot(1-p)^{n-k}$$

Такое распределение случайной величины, могущей принять одно из $n+1$ значения, называют *биномиальным распределением.*

Пример 8.1.2.5. Пусть случайная величина ξ есть число выстрелов по мишени «до первого промаха». Если вероятность попадания при одном выстреле равна p и выстрелы являются независимыми друг от друга испытаниями, то вероятность события, заключающегося в том, что при первых $k-1$ выстрелах имело место попадание, а на k-ом выстреле произошел промах, дается формулой

$$P(\xi = k) = p^{k-1}(1-p)\ , \quad \forall k = 0, 1, 2, 3, \ldots, k, \ldots \ .$$

Пример 8.1.2.6. Пусть случайная величина ξ может принимать неограниченное число значений, являющихся целыми неотрицательными числами $k = 0, 1, 2, 3, \ldots, k, \ldots$, с вероятностями, определяемыми по формуле

$$P(\xi = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}\ , \quad \lambda > 0\ ,$$

то говорят, что эта случайная величина *распределена по закону Пуассона* с параметром λ.

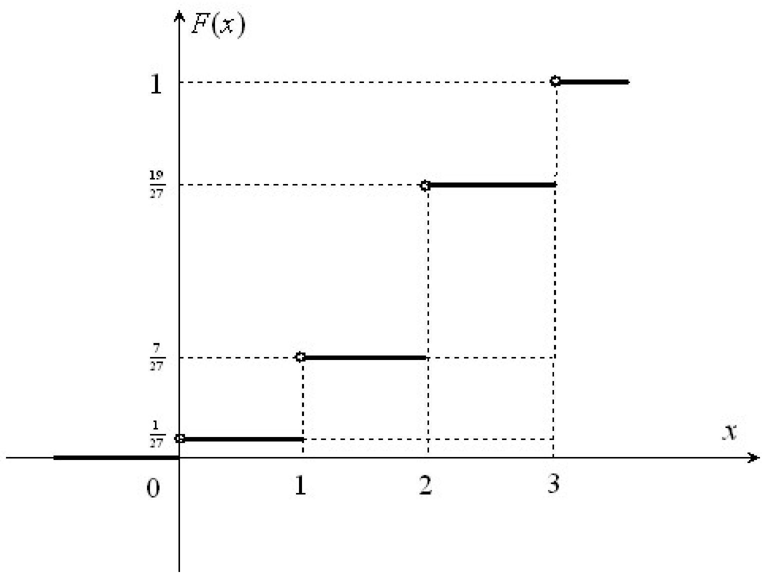


Рис. 8.1. Функция распределения числа проросших семян (пример 8.1.2.2.)

Рассмотрим теперь третий способ задания распределения случайной величины. Пусть ξ принимает одно из значений x_k, причем эти значения *упорядочены по возрастанию,* то есть, $x_k \leq x_{k+1}, \ \forall k$, и пусть вероятность события $(\xi = x_k)$ равна p_k. Тогда, в силу несовместности этих событий при разных k и согласно теореме 7.1.4.1 (см. §7.1.4), будет иметь место равенство

$$P(\xi < x_n) \ = \ p_1 + p_2 + \ldots + p_{n-2} + p_{n-1} \ = \ \sum_{k=1}^{n-1} p_k \ ,$$

ибо очевидно, что $P(\xi < x_1) = 0$.

Определение 8.1.2.2. *Функцией распределения вероятности значений случайной величины* ξ (или просто, *функцией распределения*) называется функция $F_\xi(x)$, значение которой для любого $x \in (-\infty, +\infty)$ равно $P(\xi < x)$ – вероятности того, что ξ окажется меньше, чем x.

Заметим, что на практике индекс ξ в записи функции распределения часто опускают и пишут просто $F(x) = P(\xi < x)$.

Из определения 8.1.2.2 непосредственно вытекают следующие *свойства функции распределения*:

1°. Область определения $F(x)$ – все действительные числа;

2°. Область значений – $[0, 1]$, поскольку $0 \leq F(x) \leq 1$.

3°. $F(x)$ – неубывающая функция своего аргумента x;

4°. $\lim\limits_{x\to-\infty} F(x) = 0$ и $\lim\limits_{x\to+\infty} F(x) = 1$;

5°. $F(x)$ – непрерывная слева функция в любой точке x;

6°. Имеет место равенство $P(\xi \leq x_n) = F(x_n) + p_n$. Или, в другой форме, $F(x_{n+1}) - F(x_n) = p_n$.

Эти свойства иллюстрирует (см. рис.8.1) график функции распределения случайной величины ξ рассмотренной в примере 8.1.2.2.

В заключение приведем таблицу, показывающую какие формы закона распределения могут быть использованы для описания случайных величин различных типов.

Случ. величина: / *Способ описания:*	Дискретная, с конечным числом значений	Дискретная, с бесконечн. числом значений	Непрерывн., с бесконечн. числом значений
Таблица распределения	+	–	–
Формула распределения	+	+	–
Функция распределения	+	+	+

8.1.3. Числовые характеристики дискретной случайной величины

Распределение вероятности значений случайной величины является наиболее полной формой ее описания и дает исчерпывающую информацию о ее свойствах. Однако в большом числе практически важных случаев оказывается достаточным знание лишь некоторых специальных характеристик случайной величины, позволяющих количественно оценивать основные ее свойства. К основным из этих характеристик относятся *математическое ожидание, дисперсия* и *среднее квадратичное отклонение.* Дадим последовательно их определения и опишем свойства.

Математическое ожидание дискретной случайной величины.

Рассмотрим следующую проблему. Пусть два равносильных противника A и B решили сыграть серию из четырех игр, не допускающих ничейного исхода. Каждый игрок перед началом серии внес в призовой фонд по 30 долларов, который должен полностью достаться победителю серии, или быть поделенным поровну между A и B при итоговом счете 2:2. Сыграв три игры, две из которых выиграл A и одну – B, противники решили не играть четвертую, а поделить 60 долларов «по справедливости». Сколько должен в этой ситуации получить каждый из игроков?

Покажем, что естественный (на первый взгляд) способ дележа *пропорционально достигнутым результатам,* когда 40 долларов получает A и 20 долларов – B, не является справедливым. Для этого найдем законы распределения двух случайных величин ξ – ожидаемой суммы выигрыша для A, и η (читается – «эта») – ожидаемого выигрыша для B, если бы четвертая игра все же состоялась.

Поскольку игроки имеют равные шансы на выигрыш в четвертой партии, то итоговый счет с вероятностями $\frac{1}{2}$ мог бы быть либо 2:2, при котором игроки получают по 30 долларов каждый, либо 3:1 в пользу A, с его выигрышем в 60 долларов. Иначе говоря, таблицы распределения случайных величин ξ и η имеют вид

ξ – величина выигрыша A	30	60
Вероятность этого события	$\frac{1}{2}$	$\frac{1}{2}$

η – величина выигрыша B	0	30
Вероятность этого события	$\frac{1}{2}$	$\frac{1}{2}$

Таким образом, в предположении о *многократном* доигрывании серии, величины $\text{Ср}\,\xi$ и $\text{Ср}\,\eta$ – *усредненные вероятностные оценки выигрыша* для A и B, соответственно составили бы

$$\mathrm{Cp}\,\xi = \frac{30+60}{2} = \frac{1}{2}\cdot 30 + \frac{1}{2}\cdot 60 = 45 \quad \text{и} \quad \mathrm{Cp}\,\eta = \frac{0+30}{2} = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 30 = 15\,,$$

которые и являются искомым справедливым распределением призового фонда.

Обобщим полученный результат. Пусть $x_k, k = [1, N]$ – все возможные значения случайной величины ξ. И пусть из m – полного числа элементарных событий, состоящих в том, что ξ принимает какое-нибудь значение, в n_k случаях она принимала именно значение x_k. Тогда *усредненное* значение случайной величины ξ, получаемое по результатам m испытаний будет близким к

$$M\xi = \frac{n_1x_1 + n_2x_2 + \ldots + n_Nx_N}{m} = \frac{1}{m}\sum_{k=1}^{N} n_kx_k\,.$$

С другой стороны, нетрудно видеть, что

$$M\xi = \frac{n_1x_1 + n_2x_2 + \ldots + n_Nx_N}{m} = \frac{n_1}{m}x_1 + \frac{n_2}{m}x_2 + \ldots + \frac{n_N}{m}x_N$$

и

$$M\xi = p_1x_1 + p_2x_2 + \ldots + p_Nx_N = \sum_{k=1}^{N} p_kx_k\,,$$

поскольку, $m = n_1 + n_2 + \ldots + n_N = \sum\limits_{k=1}^{N} n_k$, а в силу определения 7.1.2.1,

$$p_k = \frac{n_k}{m} = \frac{n_k}{\sum\limits_{k=1}^{N} n_k}\,.$$

Таким образом для *средневероятностной оценки* значений дискретной случайной величины ξ, можно использовать количественную характеристику $M\xi$, дав

Определение 8.1.3.1. Для случайной величины ξ с распределением

Значение ξ	x_1	x_2	$\ldots$	x_N
Вероятность события $(\xi = x_k)$	p_1	p_2	$\ldots$	p_N

математическим ожиданием называется число

$$M\xi = \sum_{k=1}^{N} x_k \cdot p_k = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots + x_N \cdot p_N \ .$$

Пример 8.1.3.1. Найти математическое ожидание случайной величины ξ, равной числу выпавших очков при бросании игральной кости.

Решение. Согласно определению 8.1.3.1 и распределению в примере 8.1.2.1, имеем

$$M\xi = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2} = 3\frac{1}{2} \ .$$

Пример 8.1.3.2. Найти математическое ожидание случайной величины ξ, равной числу проросших семян из числа трех посеянных, если вероятность прорасти для одного семени равна $\frac{2}{3}$.

Решение. Согласно определению 8.1.3.1 и распределению в примере 8.1.2.2, имеем

$$M\xi = 0 \cdot \frac{1}{27} + 1 \cdot \frac{2}{9} + 2 \cdot \frac{4}{9} + 3 \cdot \frac{8}{27} = \frac{54}{27} = 2 \ .$$

Стоит отметить, что математическое ожидание дискретной случайной величины, вообще говоря, может совпадать, а может и не совпадать с каким-либо ее значением.

Рассмотрим теперь задачу, в которой требуется найти математическое ожидание случайной величины, могущей принимать *неограниченное* число значений.

Пример 8.1.3.3. Найти математическое ожидание случайной величины ξ, равной числу последовательных выстрелов по мишени «до первого промаха», если вероятность попадания при одном выстреле равна $\frac{5}{8}$.

Решение. Поскольку число возможных значений ξ не ограничено, то в определении 8.1.3.1 сумму конечного числа слагаемых следует заменить *числовым рядом.* Понятно, что для описания распределения ξ табличный способ в этом случае не подходит, и нужно вначале получить формулу для p_k – вероятности события, заключающегося в том, что первые $k-1$ выстрелов попали в цель, а k-й выстрел оказался неудачным. Приняв во внимание, что результаты выстрелов события независимые, в силу определения 7.1.4.2 получим

$$p_k = \left(\frac{5}{8}\right)^{k-1} \frac{3}{8} = \frac{3}{5}\left(\frac{5}{8}\right)^{k} .$$

Следовательно, искомое математическое ожидание будет равно *сумме ряда*

$$M\xi = \sum_{k=1}^{\infty} k \cdot \frac{3}{5}\left(\frac{5}{8}\right)^{k} .$$

Этот ряд *сходящийся,* причем, в силу формулы (6.4.1.1), имеем

$$M\xi = \frac{3}{5}\sum_{k=1}^{\infty} k\left(\frac{5}{8}\right)^{k} = \frac{3}{5}\cdot\frac{\frac{5}{8}}{\left(1-\frac{5}{8}\right)^{2}} = \frac{8}{3} = 2\frac{2}{3} .$$

Перечислим основные *свойства математического ожидания дискретной случайной величины.*

1°. Математическое ожидание постоянной случайной величины равно этой постоянной. То есть, если все значения ξ равны c, то и $M\xi = c$.

2°. Постоянный множитель можно выносить из под знака математического ожидания. $M(c\cdot\xi) = c\cdot M\xi$.

3°. Математическое ожидание произведения конечного числа *независимых* случайных величин равно произведению их математических ожиданий.

4°. Математическое ожидание суммы конечного числа случайных величин равно сумме их математических ожиданий.

Рассмотрим последнее утверждение подробнее для случая суммы двух случайных величин. Пусть случайная величина ξ принимает значения $x_k, k=[1,s]$ с вероятностями p_k, а случайная величина η принимает значения $y_m, m=[1,t]$ с вероятностями p_m. Новая случайная величина $\xi+\eta$ в силу теоремы 7.1.4.2 будет принимать значение x_k+y_m с вероятностью $p_{km} = p_k\cdot p_{m/k}$, где $p_{m/k} = P_{(\xi=x_k)}(\eta=y_m)$. Тогда

$$M(\xi+\eta) = \sum_{k=1}^{s}\sum_{m=1}^{t}(x_k+y_m)\cdot p_{km} =$$

$$\sum_{k=1}^{s}\left(\sum_{m=1}^{t} x_k \cdot p_{km}\right) + \sum_{m=1}^{t}\left(\sum_{k=1}^{s} y_m \cdot p_{km}\right) =$$

$$= \sum_{k=1}^{s} x_k \left(\sum_{m=1}^{t} p_{km}\right) + \sum_{m=1}^{t} y_m \left(\sum_{k=1}^{s} p_{km}\right) =$$

$$= \sum_{k=1}^{s} x_k \left(\sum_{m=1}^{t} p_k \cdot p_{m/k}\right) + \sum_{m=1}^{t} y_m \left(\sum_{k=1}^{s} p_k \cdot p_{m/k}\right) =$$

$$= \sum_{k=1}^{s} x_k \cdot p_k \left(\sum_{m=1}^{t} p_{m/k}\right) + \sum_{m=1}^{t} y_m \cdot p_m \left(\sum_{k=1}^{s} p_{m/k}\right) =$$

$$= \sum_{k=1}^{s} x_k \cdot p_k + \sum_{m=1}^{t} y_m \cdot p_m = M\xi + M\eta ,$$

поскольку

$$\sum_{m=1}^{t} p_{m/k} = 1 \quad \text{и} \quad \sum_{k=1}^{s} p_{m/k} = 1$$

в силу полноты систем событий $(\xi = x_k), k = [1, s]$ и $(\eta = y_m), m = [1, t]$.

Дисперсия и среднее квадратичное отклонение дискретной случайной величины.

Итак, математическое ожидание можно назвать количественной оценкой положения на числовой оси центра вероятностной группировки возможных значений случайной величины. Однако, достаточно часто возникает потребность количественно оценить также и усредненную по вероятности *степень отклонения* этих значений от данного центра.

Рассмотрим эту проблему на примере следующей задачи.

Пример 8.1.3.3. Некий гусар, выйдя из трактира на улицу, делает последовательно шаги в 1 метр. Каждый шаг он делает равновероятно либо вправо, либо влево, и независимо от предыдущих шагов. На каком наиболее вероятном расстоянии от выхода из трактира гусар окажется, пройдя N шагов?

Решение. Введем систему координат, начало O которой находится у выхода из трактира, а ось Ox направлена вправо вдоль улицы. Координата положения гусара является некоторой случайной величиной ξ_k, значение которой, после k выполненных гусаром шагов, равно x_k. Математическое ожидание $M\xi_k$ очевидно нулевое, поскольку шаги в право и влево равновероятны и взаимно компенсируются. Поэтому рассмотрим другую случайную величину $\eta_k = (\xi_k - M\xi_k)^2$, значение d_k которой есть *квадрат координаты* положения гусара после k-го шага.

Очевидно, что $x_1 = \pm 1$, поэтому $d_1 = x_1^2 = 1$. Допустим, что гусар прошел k шагов и оказался на расстоянии $\sqrt{d_k} = |x_k|$ от входа. Случайная величина η_{k+1} (на шаге $k+1$) будет иметь два равновероятных значения

$$d_{k+1}^{(1)} = (x_k + 1)^2 = x_k^2 + 2 \cdot x_k + 1 = d_k + 2 \cdot x_k + 1$$

и

$$d_{k+1}^{(2)} = (x_k - 1)^2 = x_k^2 - 2 \cdot x_k + 1 = d_k - 2 \cdot x_k + 1\ .$$

Поэтому

$$M\eta_{k+1} = \frac{1}{2} \cdot d_{k+1}^{(1)} + \frac{1}{2} \cdot d_{k+1}^{(2)} = d_k + 1\ .$$

Последнее равенство верно для любого k, поэтому из $d_1 = 1$ следует, что $M\eta_k = k$, $\forall k$. Значит через N шагов гусар будет наиболее вероятно находится на расстоянии $\sqrt{M\eta_N} = \sqrt{N}$ метров от выхода.

Таким образом в качестве количественной характеристики степени рассеяния значений случайной величины вокруг своего центра группирования можно использовать математическое ожидание *квадрата отклонения* ξ от $M\xi$.

Определение 8.1.3.2. Величина $M(\xi - M\xi)^2$ называется *дисперсией* случайной величины ξ и обозначается $D\xi$. Величина $\sigma_\xi = \sqrt{D\xi}$ называется *средним квадратичным отклонением* случайной величины ξ.

Среднее квадратичное отклонение используют для характеристики степени разброса значений ξ в тех случаях, когда необходимо измерять разброс в тех же единицах, что и саму случайную величину.

Естественно возникает вопрос: как найти дисперсию дискретной случайной величины по ее распределению? Ответ на него дает

Теорема 8.1.3.1. **Формула для подсчета значения дисперсии.**

Дисперсия случайной величины ξ может быть найдена по формуле: $$D\xi = M(\xi^2) - (M\xi)^2\ .$$

Доказательство. Использовав определение 8.1.3.2 и свойства математического ожидания, получим

$$D\xi = M(\xi - M\xi)^2 = M(\xi^2 - 2 \cdot \xi \cdot (M\xi) + (M\xi)^2) = M(\xi^2) - 2 \cdot M\xi \cdot (M\xi) + (M\xi)^2 =$$

$= M(\xi^2) - (M\xi)^2$.

Таким образом для случайной величины ξ с распределением

Значение ξ	x_1	x_2	$\ldots$	x_N
Вероятность события $(\xi = x_k)$	p_1	p_2	$\ldots$	p_N

дисперсия может быть подсчитана по формуле

$$D\xi = \sum_{k=1}^{N} \left(x_k^2 \cdot p_k\right) - \left(\sum_{k=1}^{N} x_k \cdot p_k\right)^2 =$$

$$= x_1^2 \cdot p_1 + x_2^2 \cdot p_2 + \ldots + x_N^2 \cdot p_N - (M\xi)^2 \ .$$

Пример 8.1.3.4. Найти дисперсию и среднее квадратичное отклонение для случайной величины ξ, равной числу выпавших очков при бросании игральной кости.

Решение. Используя распределение, приведенное в примере 8.1.2.1 и $M\xi = \dfrac{7}{2}$ (см. пример 8.1.3.1), имеем для дисперсии ξ

$$D\xi = \left(1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6}\right) - \left(\frac{7}{2}\right)^2 = \frac{35}{12} = 2\frac{11}{12} \ .$$

Среднее квадратичное отклонение ξ соответственно будет равно

$$\sigma_\xi = \sqrt{D\xi} = \sqrt{\frac{35}{12}} \approx 1.71 \ .$$

В заключение отметим некоторые, полезные для практических расчетов, *свойства дисперсии.*

1°. Дисперсия любой случайной величины неотрицательна: $D\xi \geq 0$.

2°. Дисперсия постоянной случайной величины равна нулю. То есть, если все значения ξ равны c, то $D\xi = 0$.

3°. Постоянный множитель можно выносить из под знака дисперсии, предварительно возведя его в квадрат: $D(c \cdot \xi) = c^2 \cdot D\xi$.

4°. Дисперсия суммы конечного числа *независимых* случайных величин равна сумме их дисперсий.

Проверим справедливость свойства 4° для двух независимых случайных величин ξ и η. Действительно, по определению дисперсии и используя свойства математического ожидания, получаем

$$D(\xi+\eta) = M\left(\xi+\eta-M(\xi+\eta)\right)^2 =$$

$$= M\left(\ \xi+\eta-M\xi-M\eta\ \right)^2 = M\left(\ \xi-M\xi+\eta-M\eta\ \right)^2 =$$

$$= M\Big(\ (\xi-M\xi)^2+2\cdot(\xi-M\xi)\cdot(\eta-M\eta)+(\eta-M\eta)^2\ \Big) =$$

$$= M\left((\xi-M\xi)^2\right)+2\cdot M(\xi-M\xi)\cdot M(\eta-M\eta)+M\left((\eta-M\eta)^2\right) = D(\xi)+D(\eta)\ ,$$

поскольку $M(\xi-M\xi) = M\xi-M\xi = 0$ и $M(\eta-M\eta) = M\eta-M\eta = 0$.

Выясните самостоятельно, в каком месте этих выкладок использовано условие независимости ξ и η.

Пример 8.1.3.5. Найти дисперсию и среднее квадратичное отклонение для случайной величины κ, равной числу «успехов» в серии из n испытаний по схеме Бернулли.

Решение. Напомним, что в схеме Бернулли последовательно проводятся n независимых друг от друга испытаний, в каждом их которых событие A происходит (имеет место «успех») с вероятностью p и не происходит – с вероятностью $q = 1-p$. Найдем вначале математическое ожидание числа «успехов»в этой серии. Имеем

$$M\kappa = M\left(\kappa_1+\kappa_2+\kappa_3+\ldots+\kappa_n\right)\ ,$$

где κ_k – случайная величина, равная числу «успехов»при *одном* k-м испытании. Очевидно, что κ_k принимает значение 1 с вероятностью p и значение 0 с вероятностью q . Значит $M\kappa_k = p\cdot 1+q\cdot 0 = p$, $\forall k = [1,n]$. Тогда, в силу независимости испытаний, получаем $M\kappa = np$.

Для подсчета дисперсии воспользуемся формулой $D\kappa = M(\kappa^2) - (M\kappa)^2$. Поскольку квадрат значения случайной величины κ_k может равняться либо $1^2 = 1$ с вероятностью p, либо – $0^2 = 0$ с вероятностью q, то $M(\kappa_k^2) = p\cdot 1+q\cdot 0 = p$, $\forall k = [1,n]$. Тогда, в силу формулы $D\kappa_k = M(\kappa_k^2)-(M\kappa_k)^2$, получим $D\kappa_k = p-p^2 = pq$, $\forall k = [1,n]$.

Наконец, используя свойство 4° – формулу дисперсии суммы независимых случайных величин, находим, что $D\kappa = npq$. Соответственно среднее квадратичное отклонение составляет $\sigma_\kappa = \sqrt{npq}$.

Продемонстрируем эффективность применения этих формул при подсчете дисперсии и среднего квадратичного отклонения для случайной величины ξ, равной числу проросших семян из числа трех посеянных, если вероятность прорасти для одного семени равна $\frac{2}{3}$. Распределение этой величины приведено в примере 8.1.2.2, а математическое ожидание подсчитано в задаче 8.1.3.2. В данном примере $p = \frac{2}{3}$, $q = \frac{1}{3}$, а $n = 3$. Поэтому

$$M\xi = 3 \cdot \frac{2}{3} = 2 \quad ; \qquad D\xi = 3 \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{3} \quad \text{и} \quad \sigma_\xi = \sqrt{\frac{2}{3}} \approx 0.816 \; .$$

Пример 8.1.3.6. Найти математическое ожидание, дисперсию и среднее квадратичное отклонение для дискретной случайной величины ξ, вероятности значений которой определены формулой Пуассона:

$$P(\xi = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \; , \quad \lambda > 0 \; .$$

Решение. Найдем вначале $M\xi$. По определению математического ожидания, использовав формулу (6.4.2.1) с $x = \lambda$ и введя новый индекс $n = k - 1$, получим

$$M\xi = \sum_{k=0}^{+\infty} k \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} = e^{-\lambda} \cdot \lambda \cdot \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} = e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda \; .$$

Для вычисления дисперсии воспользуемся теоремой 8.1.3.1, которая утверждает, что $D\xi = M(\xi^2) - (M\xi)^2$. Снова используя формулу (6.4.2.1) и введя два новых индекса суммирования $n = k-1$ и $m = n-1$ для функционального ряда, сходящегося к показательной функции, получим

$$M(\xi^2) = \sum_{k=0}^{+\infty} k^2 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} = e^{-\lambda} \cdot \lambda \cdot \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} + e^{-\lambda} \cdot \lambda^2 \cdot \sum_{m=0}^{+\infty} \frac{\lambda^m}{m!} =$$

$$= e^{-\lambda} \cdot \lambda \cdot e^{\lambda} + e^{-\lambda} \cdot \lambda^2 \cdot e^{\lambda} = \lambda + \lambda^2 \; .$$

Таким образом, дисперсия оказывается равной

$$D\xi = M(\xi^2) - (M\xi)^2 = (\lambda + \lambda^2) - \lambda^2 = \lambda \; ,$$

а среднее квадратичное отклонение $\sigma_\xi = \sqrt{\lambda}$.

8.2. Непрерывные случайные величины

8.2.1. Функция распределения и плотность вероятности непрерывной случайной величины

Кроме дискретных на практике встречаются случайные величины, значения которых не только нельзя перечислить, но и даже не удается перенумеровать. К случайным величинам такого рода можно, например, отнести вес человека некоторого фиксированного роста, отклонение точки попадания от центра мишени, время ожидания ответа телефонного абонента и т.д. Существенной особенностью случайных величин этого сорта является то, что их значения могут оказываться сколь угодно близкими друг к другу и заполнять в совокупности некоторый промежуток. Случайные величины, обладающие подобными свойствами, принято называть *непрерывными.*

В силу этих особенностей ни табличный, ни формульный способы описания распределения непрерывных случайных величин оказываются непригодными. Действительно, с одной стороны, непрерывные случайные величины имеют бесконечно большое число значений, что обуславливает невозможность использования таблиц распределения. С другой стороны, вероятность события, заключающегося в том, что непрерывная случайная величина приняла некоторое конкретное значение, в силу определения 7.1.2.1 *равна нулю,* и, следовательно, формульный способ также бесполезен. А вот, третий способ описания распределения вероятностей – при помощи функции распределения, оказывается эффективным и удобным и для непрерывных случайных величин.

Напомним, что *функцией распределения* случайной величины ξ называется функция $F(x)$, значения которой равны вероятности события, заключающегося в том, что ξ приняла значение, удовлетворяющее неравенству $\xi < x$. Иначе говоря,

$$F(x) = P(\xi < x) \ , \tag{8.2.1.1}$$

поэтому в дополнение к определению 8.1.1.1 можно дать

Определение 8.2.1.1. Случайная величина ξ называется *непрерывной,* если ее функция распределения непрерывна для любых x, за исключением, быть может, конечного числа точек разрыва первого рода.

Из этого определения и формулы (8.2.1.1) вытекает, что *функция распределения непрерывной случайной величины обладает следующими свойствами.*

1°. $F(x)$ определена и $0 \le F(x) \le 1$ для всех $x \in (-\infty, +\infty)$.

2°. $F(x)$ монотонно возрастает для всех $x \in (-\infty, +\infty)$.

3°. $\lim\limits_{x\to-\infty} F(x) = 0$ и $\lim\limits_{x\to+\infty} F(x) = 1$.

4°. Вероятность попадания значения непрерывной случайной величины ξ в промежуток $[a, b)$ равна

$$P(a \le \xi < b) = F(b) - F(a) \ . \tag{8.2.1.2}$$

На рисунке 8.2а) приведены примеры графиков функций распределения непрерывных случайных величин.

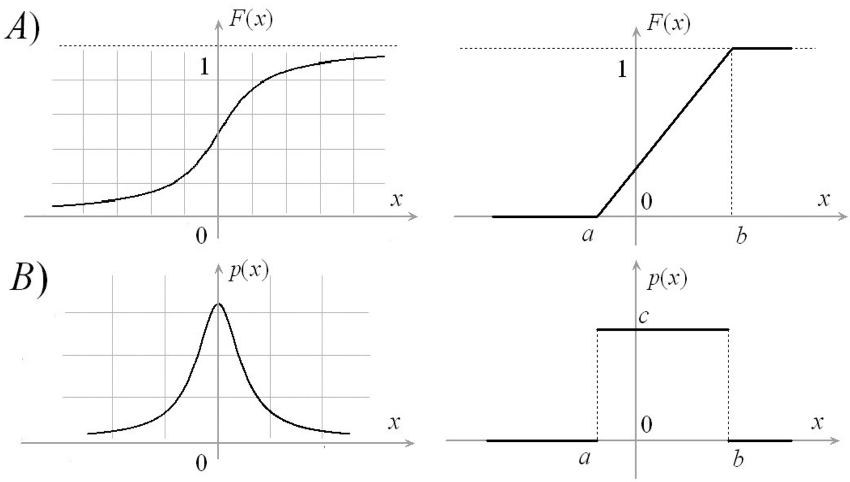


Рис. 8.2. Примеры графиков типичных функций распределения и плотности вероятности непрерывных случайных величин

Рассмотрим теперь случай, когда функция распределения непрерывной случайной величины не только непрерывна, но и дифференцируема, то есть имеет производную и дифференциал $\forall x \in (-\infty, +\infty)$, за исключением, быть может, конечного числа точек. Для такой случайной величины оказывается возможным и использование альтернативного способа описания функции распределения, не имеющего аналога для дискретного случая.

Оценим вероятность события, заключающегося в том, что случайная величина ξ примет значение из «малого» промежутка $[x, x+\Delta x)$. Для этого воспользуемся формулой (8.2.1.2), заменив приращение $F(x)$ ее дифференциалом,

$$P(x \leq \xi < x+\Delta x) = F(x+\Delta x) - F(x) \approx dF = F'(x) \cdot \Delta x \ .$$

Таким образом функция

$$p(x) = F'(x) \approx \frac{1}{\Delta x} \cdot P(x \leq \xi < x+\Delta x) \qquad (8.2.1.3.)$$

может использоваться для оценки «удельной» (то есть приходящейся на единицу длины промежутка значений) вероятности попадания значения ξ в малую окрестность значения x.

Эта оценка носит название *плотности вероятности* непрерывной случайной величины ξ и является, как и функция распределения, полным описанием ξ. Достаточно часто с практической точки зрения оказывается более удобным использовать плотность вероятности случайной величины вместо ее функции распределения.

Основные свойства плотность вероятности описываются следующим образом.

1°. $p(x) \geq 0$ для всех $x \in (-\infty, +\infty)$.

2°. Несобственный интеграл $\int\limits_{-\infty}^{+\infty} p(x)\ dx = 1$, то есть он сходится и его значение равно единице. Пользуясь геометрической интерпретацией смысла интеграла, можно утверждать, что *площадь фигуры,* ограниченная графиком функции $y = p(x)$ и осью Ox также равна единице.

3°. Кроме формулы (8.2.1.3) плотность вероятности и функция распределения связаны соотношением

$$F(x) = \int\limits_{-\infty}^{x} p(t)\ dt\ , \quad \forall\, x \in (-\infty, +\infty).$$

4°. Вероятность попадания значения непрерывной случайной величины ξ в промежуток (a, b) представляется определенным (или несобственным при бесконечных a или b) интегралом

$$P(a \leq \xi < b) = \int\limits_{a}^{b} p(x)\, dx\ . \qquad (8.2.1.4)$$

На рисунке 8.2в) приведены графики плотностей вероятности, соответствующие функциям распределения, показанным на рис.8.2а).

Для иллюстрации понятий функции распределения и плотности вероятности рассмотрим часто встречающиеся на практике случаи *равномерного* и *нормального* распределений непрерывной случайной величины.

Определение 8.2.1.2. Непрерывная случайная величина ξ называется *равномерно распределенной на отрезке* $[a, b]$, если ее плотность вероятности $p(x)$ постоянна на этом отрезке и равна нулю вне его, то есть

$$p(x) = \left[\begin{array}{ll} 0, & \text{если } x \leq a \ , \\ \dfrac{1}{b-a}, & \text{если } a < x \leq b \ , \\ 0, & \text{если } x > b \ . \end{array} \right.$$

Соответственно, функция распределения этой случайной величины ξ будет

$$F(x) = \left[\begin{array}{ll} 0, & \text{если } x \leq a \ , \\ \dfrac{x-a}{b-a}, & \text{если } a < x \leq b \ , \\ 1, & \text{если } x > b \ . \end{array} \right.$$

Графики функций $F(x)$ и $p(x)$ для равномерного распределения показаны справа на рисунке 8.2. Обратите внимание, что значение параметра $c = \dfrac{1}{b-a}$ выбирается (см. свойство 2° плотности вероятности) из условия *равенства единице* площади фигуры (в данном случае – прямоугольника), ограниченной графиком функции $y = p(x)$ и осью Ox.

Определение 8.2.1.3. Непрерывная случайная величина ξ называется распределенной *нормально* (или *по закону Гаусса*) с параметрами α и σ, если ее плотность вероятности $p(x)$ определена формулой

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\dfrac{(x-\alpha)^2}{2\sigma^2}} \ . \qquad (8.2.1.5)$$

Значение множителя перед экспонентой в (8.2.1.5) находится по свойству 2° плотности вероятности. Функция распределения этой случайной величины ξ представляется в виде «неберущегося» несобственного интеграла

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\dfrac{(t-\alpha)^2}{2\sigma^2}}\, dt \ .$$

Значения функции $F(x)$ находятся в этом случае либо по таблицам, либо при помощи функционального ряда (6.4.2.4). В случае, когда значения параметра $\alpha = 0$, а значение параметра $\sigma = 1$, принято говорить о *стандартном нормальном* законе распределения случайной величины ξ.

Графики функций $F(x)$ и $p(x)$ для нормального распределения показаны слева на рисунке 8.2. Геометрический смысл параметров α и σ иллюстрирует рис. 8.3. Заметим, что функцию распределения непрерывной случайной величины, распределенной по нормальному закону, часто записывают в виде

$$F(x) = \frac{1}{2} + \frac{1}{2}\,\Phi\left(\frac{x-\alpha}{\sigma}\right) \ ,$$

где

$$\Phi(x) = \sqrt{\frac{2}{\pi}} \int\limits_{0}^{x} e^{-\dfrac{t^2}{2}}\, dt$$

– так называемая *функция Лапласа*, значения которой имеются в справочниках или электронных таблицах.

Пример 8.2.1.1. Для непрерывной случайной величины ξ, нормально распределенной с параметрами $\alpha = 0$ и $\sigma = \dfrac{1}{\sqrt{2}}$, найти вероятность попадания ее значения на промежуток $(1, 3)$.

Решение. По формулам (8.2.1.4) и (8.2.1.5) находим (с использованием компьютера или калькулятора)

$$P(1 \le \xi < 3) = \int\limits_{1}^{3} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\dfrac{(x-a)^2}{2\sigma^2}}\, dx = \frac{1}{\sqrt{\pi}} \cdot \int\limits_{1}^{3} e^{-x^2}\, dx \approx 0.0786 \ .$$

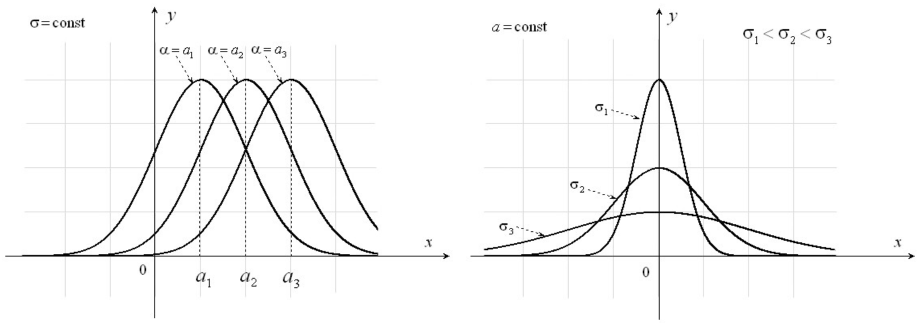


Рис. 8.3. Плотность вероятности при нормальном распределении с различными α и σ .

Пример 8.2.1.2. Плотность вероятности непрерывной случайной величины ξ задана равенством

$$p(x) = \frac{c}{4+x^2}\ .$$

Найти величину c и вероятность попадания значения случайной величины на интервал $(-\infty < \xi < 0)$.

Решение. Воспользуемся неопределенным интегралом, найденным при решении задачи 6.2.2.3.

$$\int \frac{dx}{x^2+4} = \frac{1}{2}\operatorname{arctg}\frac{x}{2} + C\ .$$

Найдем значение c по свойству 2° плотности вероятности непрерывной случайной величины (см. §8.2.1). Необходимо, чтобы

$$\int\limits_{-\infty}^{+\infty} \frac{c\,dx}{x^2+4} = 1\ .$$

Поскольку

$$\int\limits_{-\infty}^{+\infty} \frac{c\,dx}{x^2+4} = \frac{c}{2}\operatorname{arctg}\frac{x}{2}\Big|_{-\infty}^{+\infty} = \frac{c}{2}\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) = \frac{c\cdot\pi}{2} = 1\ , \quad \text{то} \quad c = \frac{2}{\pi}\ .$$

Наконец, по свойству 4° плотности вероятности непрерывной случайной величины (формула 8.2.1.4) искомая вероятность равна

$$P(-\infty < \xi < 0) = \int_{-\infty}^{0} p(x)\,dx = \frac{2}{\pi}\int_{-\infty}^{0} \frac{dx}{x^2+4} =$$

$$= \frac{1}{\pi}\operatorname{arctg}\frac{x}{2}\Big|_{-\infty}^{0} = \frac{1}{\pi}\left(0 - \left(-\frac{\pi}{2}\right)\right) = \frac{1}{2}\,.$$

8.2.2. Математическое ожидание и дисперсия непрерывной случайной величины

Пусть непрерывная случайная величина ξ имеет плотность вероятности $p(x)$. Тогда формула для математического ожидания дискретной случайной величины ξ^*, значения которой x_k мало отличаются от значений ξ в небольших окрестностях Δx, с учетом соотношения (8.2.1.2), в силу которого $p_k \approx p(x_k)\Delta x$, может быть преобразована к виду

$$M\xi^* = \sum_{k=1}^{n} p_k x_k \approx \sum_{k=1}^{n} (p(x_k)\Delta x)\, x_k = \sum_{k=1}^{n} x_k p(x_k)\Delta x\ .$$

Что, в свою очередь, позволяет использовать понятия интегральной суммы (6.1.1.1) и несобственного интеграла и дать

Определение 8.2.2.1. *Математическим ожиданием* непрерывной случайной величины ξ с плотностью вероятности $p(x)$ называется число (несобственный интеграл)

$$M\xi = \int_{-\infty}^{+\infty} xp(x)\ dx$$

Аналогично для дисперсии этой случайной величины естественно принять

Определение 8.2.2.2. *Дисперсией* непрерывной случайной величины ξ с плотностью вероятности $p(x)$ называется число (несобственный интеграл)

$$D\xi = \int_{-\infty}^{+\infty} (x - M\xi)^2\, p(x)\ dx$$

Предполагается, что несобственные интегралы в определениях 8.2.2.1 и 8.2.2.2 *сходятся абсолютно,* то есть сходятся также и интегралы от модуля подынтегральной функции. Однако для некоторых распределений данное условие не выполняется и математического ожидания или дисперсии не существует. Например, для случайной величины ξ, распределенной по так называемому *закону Коши,* с плотностью вероятности

$$p(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2} \,,$$

интеграл в формуле для $M\xi$ «берущийся»:

$$\int xp(x)\,dx = \frac{1}{\pi} \cdot \int \frac{x}{1+x^2}\,dx = \frac{1}{2\pi} \cdot \ln(1+x^2) + C \,.$$

Однако предел $\lim\limits_{x\to\infty} \ln(1+x^2)$ не существует, а, значит, не существует и $M\xi$.

Проиллюстрируем теперь использование определений 8.2.2.1 и 8.2.2.2 для конкретных случайных величин.

Пример 8.2.2.1. Найти математическое ожидание, дисперсию и среднее квадратичное отклонение для случайной величины τ, равной времени ожидания поезда метро пассажиром, приходящим на станцию в случайный момент времени, если поезда следуют с равными интервалами в 5 минут.

Решение. Пусть два последовательно идущих поезда прибывают на станцию в моменты времени $t_0 = 0$ и $t_1 = 5$, а пассажир появляется на станции равновероятно в любой из моментов $t \in [0, 5]$. Это означает, что τ – равномерно распределенная на $[0, 5]$ случайная величина и ее функция распределения и плотность вероятности τ в соответствии с определением 8.2.1.2 будут равны

$$F(t) = \left[\begin{array}{ll} 0, & \text{если } t \le 0 \,, \\ \dfrac{t}{5}, & \text{если } 0 < t \le 5 \,, \\ 1, & \text{если } t > 5 \,, \end{array}\right. \quad \text{и} \quad p(t) = \left[\begin{array}{ll} 0, & \text{если } t \le 0 \,, \\ \dfrac{1}{5}, & \text{если } 0 < t \le 5 \,, \\ 0, & \text{если } t > 5 \,. \end{array}\right.$$

Тогда математическое ожидание составит

$$M\tau = \int\limits_{-\infty}^{+\infty} tp(t)\,dt = \int\limits_0^5 t \cdot \left(\frac{1}{5}\right)\,dt = \frac{1}{5}\int\limits_0^5 t\,dt = \frac{1}{5} \cdot \left(\frac{t^2}{2}\Bigg|_0^5\right) = \frac{5}{2} \text{ мин},$$

а дисперсия

$$D\tau = \int\limits_{-\infty}^{+\infty} (t - M\tau)^2\, p(t)\, dt = \int\limits_{0}^{5} \left(t - \frac{5}{2}\right)^2 \cdot \left(\frac{1}{5}\right)\, dt = \frac{1}{5}\int\limits_{0}^{5} \left(t - \frac{5}{2}\right)^2\, dt =$$

$$= \frac{1}{5}\cdot \left(\frac{1}{3}\left(t - \frac{5}{2}\right)^3 \Bigg|_0^5\right) = \frac{1}{15}\cdot \left(\left(\frac{5}{2}\right)^3 - \left(-\frac{5}{2}\right)^3\right) = \frac{1}{15}\cdot \frac{125}{4} = \frac{25}{12}\ \text{мин}^2\ .$$

Наконец, среднее квадратичное отклонение будет равно

$$\sigma_\tau = \sqrt{D\tau} = \sqrt{\frac{25}{12}} = \frac{5}{2\sqrt{3}} \approx 1.44\ \text{мин}.$$

Итак, математическое ожидание времени ожидания поезда составляет 2.5 минуты при среднем квадратичном отклонении в 1.44 минуты.

Пример 8.2.2.2. Найти математическое ожидание для случайной величины ξ, распределенной по нормальному закону с параметрами $a = 0$ и $\sigma = \dfrac{1}{\sqrt{2}}$.

Решение. В данном случае $p(x) = \dfrac{1}{\sqrt{\pi}}\, e^{-x^2}$. Поэтому

$$M\xi = \int\limits_{-\infty}^{+\infty} x p(x)\, dx = \frac{1}{\sqrt{\pi}} \cdot \int\limits_{-\infty}^{+\infty} x\, e^{-x^2}\, dx =$$

$$= \frac{1}{2\sqrt{\pi}} \cdot \int\limits_{-\infty}^{+\infty} e^{-x^2}\, d\left(x^2\right) = \left(-\frac{1}{2\sqrt{\pi}}\right) e^{-x^2}\ \Bigg|_{-\infty}^{+\infty} = 0\ .$$

Стоит отметить, что в общем случае непрерывной случайной величины ξ, распределенной по *нормальному закону* (8.2.1.5) с параметрами a и σ (см. рис. 8.3), справедливы равенства

$$M\xi = a\ , \qquad D\xi = \sigma^2\ , \qquad \sigma_\xi = \sigma\ ,$$

то есть мы имеем пример того, что само распределение (функция распределения или плотность вероятности), если известен его тип, может быть описано лишь при помощи количественных характеристик случайной величины.

8.2.3. Вспомогательные количественные характеристики случайных величин

В теории вероятностей и математической статистике помимо основных количественных характеристик случайных величин – математического ожидания и дисперсии, используются и другие. Различие значений этих численных характеристик отражает отличия в свойствах распределений случайных величин.

Рассмотрим некоторые из этих характеристик наиболее часто используемые в теории вероятностей и математической статистике.

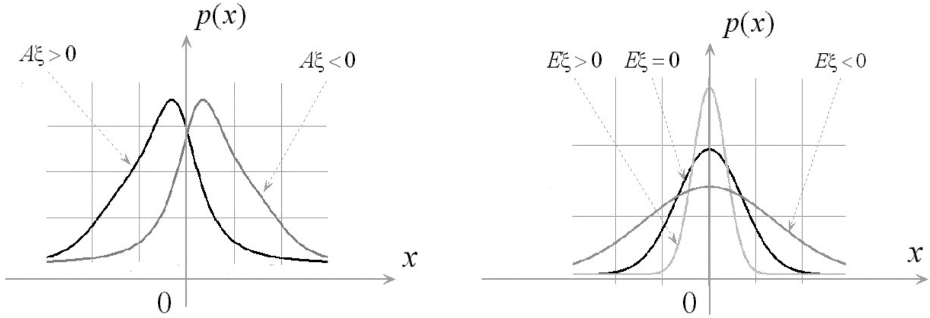


Рис. 8.4. Геометрическая интерпретация асимметрии и эксцесса.

Количественная характеристика ξ		
Название	Симв.	Описание
Начальный момент порядка k	$M\xi^k$	*Начальным моментом порядка k* непрерывной случайной величины называется число равное несобственному интегралу $\int\limits_{-\infty}^{+\infty} x^k p(x)\ dx$, или, для дискретной случайной величины, ряду (или конечной сумме) $\sum\limits_{k=1}^{+\infty} x_k^k p_k$.
Центральный момент порядка k	$\mu\xi^k$	*Центральным моментом порядка k* непрерывной случайной величины порядка k называется число равное несобственному интегралу $$\int\limits_{-\infty}^{+\infty} (x - M\xi)^k\, p(x)\ dx\ ,$$ или, для дискретной случайной величины, ряду (или конечной сумме) $$\sum_{k=1}^{+\infty} (x_k - M\xi)^k\, p_k\ .$$
Коэффициент асимметрии	$A\xi^k$	*Коэффициентом асимметрии* случайной величины ξ, оценивающим величину нарушения симметрии («скошенности») ее распределения относительно математического ожидания $M\xi$, называется число $$A\xi^k = \frac{\mu\xi^3}{\sqrt{(D\xi)^3}} = \frac{M(\xi - M\xi)^3}{\sqrt{(D\xi)^3}}\ .$$
Коэффициент эксцесса	$E\xi^k$	*Коэффициентом эксцесса* случайной величины ξ, оценивающим «островершинность» ее распределения в сравнении с нормальным, называется число $$E\xi^k = \frac{\mu\xi^4}{(D\xi)^2} - 3 = \frac{M(\xi - M\xi)^4}{(D\xi)^2} - 3\ .$$

Количественная характеристика ξ		
Название	Симв.	Описание
Коэффициент вариации	$V\xi$	*Коэффициентом вариации* случайной величины называется число $V\xi = \dfrac{\sqrt{D\xi}}{M\xi} \cdot 100\%$, позволяющее сравнивать степени рассеяния случайных величин, измеряемых в разных единицах.
Мода	$Mo\xi$	*Модой непрерывной* случайной величины называется такое ее значение, при котором плотность $p(x)$ имеет максимум. *Модой дискретной* случайной величины называется ее значение, имеющее наибольшую вероятность.
Медиана	$Me\xi$	*Медианой* непрерывной случайной величины ξ называется ее значение $Me\xi$ для которого $$P(\xi < Me\xi) = \frac{1}{2} = P(\xi > Me\xi)\ ,$$ то есть значения ξ большие и меньшие, чем $Me\xi$ равновероятны.
Квантиль	$Q\xi$	*$q-$квантилем* непрерывной случайной величины ξ называется ее значение $Q\xi$ для которого $$F(Q\xi) = q = P(\xi < Q\xi)\ .$$ При $q = \dfrac{1}{2}$ очевидно, что 0.5-квантиль является медианой для ξ.

На рисунке 8.4 приведены графики плотностей вероятности различных распределений. Например, по левому рисунку 8.4 очевидно, что «скошенности» графика вправо соответствуют отрицательные значения параметра асимметрии, а «скошенности» влево – положительные. Аналогично, правый график иллюстрирует тот факт, что «островершинность» графика плотности вероятности больше, чем у стандартной гауссовой кривой при положительных значениях эксцесса, и меньше –

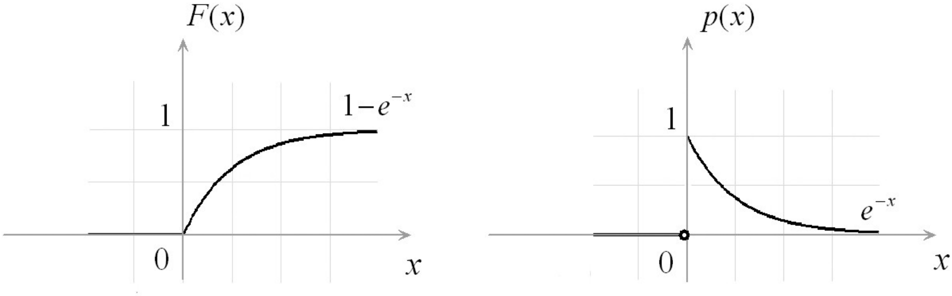


Рис. 8.5. Функция распределения и плотность вероятности экспоненциально распределенной случайной величины

при отрицательной. Заметим, что для *нормально распределенной* (то есть с плотностью вероятности (8.2.1.5)) случайной величины асимметрия и эксцесс имеют нулевые значения.

Пример 8.2.3.1. Найти все возможные численные характеристики случайной величины ξ, распределенной по *экспоненциальному* закону

$$F(x) = \left[\begin{array}{ll} 0, & \text{если } x < 0\ , \\ 1 - e^{-x}, & \text{если } x \geq 0\ . \end{array} \right.$$

Решение. Плотность вероятности для данного распределения будет иметь вид (графики $F(x)$ и $p(x)$ приведены на рис. 8.5)

$$p(x) = \left[\begin{array}{ll} 0, & \text{если } x < 0\ , \\ e^{-x}, & \text{если } x \geq 0\ . \end{array} \right.$$

Поэтому математическое ожидание определяется формулой

$$M\xi = \int\limits_{-\infty}^{+\infty} xp(x)\,dx = \int\limits_{0}^{+\infty} xe^{-x}\,dx = -xe^{-x}\Big|_0^{+\infty} + \int\limits_{0}^{+\infty} e^{-x}\,dx =$$

$$= -xe^{-x}\Big|_0^{+\infty} - e^{-x}\Big|_0^{+\infty} = 1\ .$$

Аналогично, используя последовательно интегрирование по частям (или же таблицу неопределенных интегралов), находим дисперсию

$$D\xi = M(\xi^2) - (M\xi)^2 = \left(\int\limits_{0}^{+\infty} x^2 e^{-x}\,dx \right) - 1 = 2 - 1 = 1\ .$$

Значит, для рассматриваемой случайной величины $D\xi = 1$ и $\sigma_\xi = \sqrt{D\xi} = 1$.
Наконец, в силу $D\xi = 1$, а также равенств

$$\int\limits_0^{+\infty} (x-1)^3 e^{-x}\,dx = 2 \quad \text{и} \quad \int\limits_0^{+\infty} (x-1)^4 e^{-x}\,dx = 9\,,$$

получаем значения коэффициентов асимметрии $A\xi = 2$ и эксцесса $E\xi = 6$.

Проверьте самостоятельно, что величины коэффициента вариации, моды, медианы и $q-$квантиля соответственно равны

$$V\xi = 1\,,\;\; Mo\xi = 0\,,\;\; Me\xi = \ln 2 \quad \text{и} \quad Q\xi = \ln\frac{1}{1-q}\,.$$

8.3. Функции и системы случайных величины

8.3.1. Функции случайных величин

Рассмотрим некоторую случайную величину ξ, распределение которой известно, и непрерывную функцию $y = f(x)$. Тогда $\eta = f(\xi)$ будет также являться случайной величиной, распределение которой полностью и однозначно может быть найдено по распределению случайной величины ξ. Поясним этот факт следующим примером.

Пример 8.3.1.1. Пусть дискретная случайная величина ξ, принимающая значения $x_1 = -1,\ x_2 = 0,\ x_3 = 1,\ x_4 = 2$, имеет распределение вида

Значение ξ	-1	0	1	2
Вероятность события $(\xi = x_k)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$

,

а функция $y = x^2 + 3$. Тогда $\eta = \xi^2 + 3$ может принимать значения $y_1 = 3,\ y_2 = 4,\ y_3 = 7$ с вероятностями

$$\begin{aligned} P(\eta = 3) &= \frac{1}{6}, \\ P(\eta = 4) &= \frac{1}{3} + \frac{1}{4} = \frac{7}{12}, \\ P(\eta = 7) &= \frac{1}{4}. \end{aligned}$$

Следовательно распределение случайной величины η будет иметь вид

Значение η	3	4	7
Вероятность события $(\eta = y_k)$	$\frac{1}{6}$	$\frac{7}{12}$	$\frac{1}{4}$

,

В заключение рассмотрим вопрос о функции, зависящей от *непрерывной* случайной величины. Пусть случайная величина η связана с непрерывной случайной величиной ξ, имеющей функцию распределения $F(x)$, соотношением $\eta = f(\xi)$, где $y = f(x)$ – некоторая непрерывная монотонно возрастающая функция. Тогда функция распределения случайной величины η будет иметь вид

$$G(y) = P(\ \eta < y\ ) = P(\ f(\xi) < y\ ) = P(\ \xi < g(y)\ ) = F(g(y))\ ,$$

где $x = g(y)$ – функция, *обратная*[1] к $y = f(x)$.

Формула для математического ожидания для случайной величины η записывается еще проще, если известна $p(x)$ – плотность вероятности непрерывной случайной величины ξ

$$M\eta = \int\limits_{-\infty}^{+\infty} f(x)p(x)\,dx\ .$$

[1]Напомним, что функция $g(x)$ является обратной к монотонной функции $f(x)$, если $g(f(x)) = f(g(x)) = x$, а не $\quad g(x) = \dfrac{1}{f(x)} \quad$!

8.3.2. Многомерные случайные величины и распределения вероятностей их значений

До сих пор мы рассматривали случайные события, заключающиеся в том, что некоторая величина принимала определенное значение. Однако, на практике достаточно часто встречаются ситуации, когда случайным образом принимают определенные значения несколько (и, быть может, неограниченное число) случайных величин. В этом случае принято говорить о *многомерной случайной величине* (или *случайном процессе*). В дальнейшем ради большей простоты и ясности изложения, но без существенной потери общности, ограничимся рассмотрением системы, состоящей лишь из двух случайных величин ξ и η.

Как и для одномерной случайной величины, основным и наиболее полным способом описания системы случайных величин является *распределение вероятностей* возможных значений ξ и η, представляемое двумерной таблицей, содержащей перечни как возможных значений, так и вероятностей событий, состоящих в том, что именно эти значения примут случайные величины, образующие рассматриваемую систему.

Предположим, что случайная величина ξ принимает одно из значений $x_1, x_2, \ldots, x_n$, а случайная величина η – одно из значений $y_1, y_2, \ldots, y_m$, и пусть $p_{ji} = P\left((\xi = x_i)\cdot(\eta = y_j)\right)$, то есть вероятность события, заключающегося в том, что ξ приняла значение x_i, а η – значение y_j. В этом случае говорят о *системе дискретных* случайных величин $\{\xi, \eta\}$ с распределением, задаваемом таблицей 8.3.2.1. [2]

Эта же таблица позволяет построить распределения каждой из случайных величин ξ и η по отдельности. Действительно, событие $(\xi = x_i)$, заключающегося в том,что ξ приняла значение x_i, представимо как сумма несовместных событий вида

$$(\xi = x_i)\cdot(\eta = y_1) + (\xi = x_i)\cdot(\eta = y_2) + \ldots + (\xi = x_i)\cdot(\eta = y_m)\ .$$

Поэтому, согласно теореме 7.1.4.1,

$$P(\xi = x_i) = \sum_{j=1}^{m} P\left((\xi = x_i)\cdot(\eta = y_j)\right) = \sum_{j=1}^{m} p_{ji}\ ,$$

[2] Как и в одномерном случае, значения случайных величин записываются в порядке их возрастания и без повторений.

и распределение одномерной случайной величины ξ может быть представлено в виде таблицы 8.3.2.2. Аналогичное одномерное распределение случайной величины η можно записать в виде таблицы 8.3.2.3.

Заметим, что поскольку события

$$(\xi = x_i) \cdot (\eta = y_j),\ \forall i = [1, n],\ \forall j = [1, m]$$

образуют полную группу, то справедливы равенства

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ji} = 1\ , \quad \sum_{i=1}^{n} p_i = 1\ , \quad \sum_{j=1}^{m} q_j = 1\ ,$$

где использованы следующие обозначения для одномерных случайных величин, входящих в систему

$$p_i = \sum_{j=1}^{m} p_{ji}\ , \quad q_j = \sum_{i=1}^{n} p_{ji}\ .$$

Значения: ξ η	x_1	x_2	$\dots$	x_i	$\dots$	x_n
y_1	p_{11}	p_{12}	$\dots$	p_{1i}	$\dots$	p_{1n}
y_2	p_{21}	p_{22}	$\dots$	p_{2i}	$\dots$	p_{2n}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_j	p_{j1}	p_{j2}	$\dots$	p_{ji}	$\dots$	p_{jn}
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
y_m	p_{m1}	p_{m2}	$\dots$	p_{mi}	$\dots$	p_{mn}

Таблица 8.3.2.1. Совместное (двумерное) распределение ξ и η.

Значение ξ	x_1	x_2	$\dots$	x_n
Его вероятность	$p_1 = \sum\limits_{j=1}^{m} p_{j1}$	$p_2 = \sum\limits_{j=1}^{m} p_{j2}$	$\dots$	$p_n = \sum\limits_{j=1}^{m} p_{jn}$

Таблица 8.3.2.2. Одномерное распределение ξ.

Значение η	y_1	y_2	$\dots$	y_m
Его вероятность	$q_1 = \sum\limits_{i=1}^{n} p_{1i}$	$q_2 = \sum\limits_{i=1}^{n} p_{2i}$	$\dots$	$q_m = \sum\limits_{i=1}^{n} p_{mi}$

Таблица 8.3.2.3. Одномерное распределение η.

Допустим теперь, что событие $\eta = y_j$ произошло. Выясним, какое распределение в этом случае будет иметь случайная величина ξ. Согласно формуле для вероятности произведения событий (см. теорему 7.1.4.2), используя введенные выше обозначения, получаем

$$P_{\eta=y_j}(\ \xi = x_i\ ) = \frac{P(\ (\xi = x_i)\cdot(\eta = y_j)\ )}{P(\eta = y_j)} = \frac{p_{ji}}{q_j}\,, \quad \forall i = [1,n]\ . \quad (8.3.2.1)$$

Это сопоставление называется *условным распределением случайной величины* ξ. Аналогичным способом определяется и условное распределение для случайной величины η.

Альтернативным способом описания системы случайных величин (не только дискретных, но и непрерывных) служит *функция распределения,* которая вводится аналогично одномерному случаю следующим определением.

Определение 8.3.2.1. *Функцией распределения системы двух случайных величин* ξ *и* η называется функция двух переменных

$$F(x,y) = P\,(\ (\xi < x)\cdot(\eta < y)\ )\ .$$

Перечислим основные *свойства функции распределения системы случайных величин* $F(x, y)$, вытекающие из ее определения.

1°. Область определения: $F(x, y)$ существует для всех $\{x, y\}$.

2°. Область значений: $0 \leq F(x, y) \leq 1$.

3°. $F(x, y)$ монотонно возрастает по каждому из своих аргументов.

4°. $F(-\infty, -\infty) = F(-\infty, +\infty) = F(+\infty, -\infty) = 0,\ F(+\infty, +\infty) = 1$.

5°. Распределения случайных величин ξ и η по отдельности равны соответственно $F(x, +\infty)$ и $F(+\infty, y)$.

В большом числе практически важных случаев для системы непрерывных случайных величин функцию распределения удается представить в виде

$$F(x, y) = \int\limits_{-\infty}^{x} \left(\int\limits_{-\infty}^{y} p(t_1, t_2)\, dt_1 \right) dt_2 \ .$$

Тогда функцию $p(x, y)$ принято называть *плотностью вероятности* системы непрерывных случайных величин. Ее можно выразить через функцию распределения следующим образом

$$p(x, y) = \frac{d}{dx} \left(\frac{dF(x, y)}{dy} \right) \ .$$

Напомним, что для функций зависящих от двух переменных вначале внутренняя производная по y вычисляется в в предположении, что $x = \mathrm{const}$, а затем, полученный результат дифференцируется по x в предположении $y = \mathrm{const}$.

Отметим основные свойства плотности вероятности системы двух непрерывных случайных величин, вытекающие из ее определения.

1°. $p(x, y) \geq 0$ для всех $\{x, y\}$.

2°. Имеет место равенство:

$$\int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{+\infty} p(x, y)\, dy \right) dx = 1 \ . \qquad (8.3.2.2)$$

3°. Одномерные плотности вероятности случайных величин ξ и η находятся по формулам

$$p(x) = \int\limits_{-\infty}^{+\infty} p(x,y)\,dy \qquad q(y) = \int\limits_{-\infty}^{+\infty} p(x,y)\,dx \ .$$

Из интуитивно очевидных соображений следует, что значения случайных величин образующих систему, могут в определенной степени зависеть друг от друга. Причем степень этой зависимости может изменяться от жесткой однозначной функциональной зависимости (см. §3.1) до полной независимости значений друг от друга. В том случае, когда при конкретном значении одной из случайных величин существует функция распределения (или плотность) вероятности другой, принято говорить о *вероятностной или стохастической форме* зависимости этих величин.

Определение 8.3.2.2. *Условной функцией распределения* непрерывной случайной величины ξ, входящей в систему $\{\xi,\ \ \eta\}$, называется ее функция распределения, связывающая значения одной случайной величины с их вероятностями, при условии, что другая случайная величина η приняла определенное значение.

Таким же образом определяется и *условная плотность вероятности,* обозначаемая как $p_{\eta=y}(x)$. Формулы для подсчета условной плотности вероятности аналогичны выражениям (8.3.2.1) и имеют вид

$$p_{\eta=y}(x) = \frac{p(x,y)}{q(y)} = \frac{p(x,y)}{\int\limits_{-\infty}^{+\infty} p(x,y)\,dx}\ , \qquad q_{\xi=x}(y) = \frac{p(x,y)}{p(x)} = \frac{p(x,y)}{\int\limits_{-\infty}^{+\infty} p(x,y)\,dy}\ . \tag{8.3.2.3}$$

Если распределение случайной величины ξ не зависит от того, какое значение принимает вторая случайная величина η, то говорят о независимости этих величин. Для системы непрерывных случайных величин в этом случае справедливы равенства

$$p_{\eta=y}(x) = p(x)\ , \qquad q_{\xi=x}(y) = q(y)\ ,$$

что в сопоставлении с формулами (8.3.2.2), позволяет дать

Определение 8.3.2.2. Непрерывные случайные величины ξ и η называются *независимыми,* если $p(x,y) = p(x) \cdot q(y)$.

При этом стоит обратить внимание, что степень зависимости случайных величин может отличаться для различных диапазонов их значений. Например, рост взрослого человека слабо зависит от его возраста, в то время как для ребенка эта зависимость весьма существенна.

В заключение покажем, что по схеме, аналогичной рассмотренной в §8.3.1, можно использовать функции и для систем случайных величин. Ограничимся следующим демонстрационным примером.

Пример 8.3.2.1. Пусть в системе, состоящей из двух дискретных случайных величин $\{\xi,\ \eta\}$, могущих принимать значения для $\xi\ :\ x_1 = -1,\ x_2 = 0,\ x_3 = 1,\ x_4 = 2$, и соответственно для $\eta\ :\ y_1 = 0,\ y_2 = 1,\ y_3 = 2$, каждая возможная пара значений имеет одинаковую вероятность $P(\ (\xi = x_i)\cdot(\eta = y_j)\ ) = \dfrac{1}{12}$, а функция определена формулой $\kappa = \xi^2\eta$.

Составим таблицу значений случайной величины κ для всех допустимых значений ξ и η.

Значения: ξ η	x_1	x_2	x_3	x_4
y_1	0	0	0	0
y_2	1	0	1	4
y_3	2	0	2	8

Таким образом κ принимает значения $z_1 = 0,\ z_2 = 1,\ z_3 = 2,\ z_4 = 4,\ z_5 = 8$, а распределение случайной величины κ, в силу равновероятности каждой клетки таблицы, будет иметь вид

Значение κ	0	1	2	4	8
Вероятность события $(\kappa = z_k)$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$

8.3.3. Количественные характеристики для системы случайных величин

Как и в случае одной случайной величины, в наиболее полной степени описывает систему случайных величин распределение вероятности их значений. Однако далеко не всегда эта форма описания системы удобна для использования, а и вовсе часто не известна. В этих случаях можно применять специальные количественные характеристики, аналогичные рассмотренным для одномерных случайных величин. Дадим определения наиболее часто используемых на практике количественных характеристик системы двух случайных величин ξ и η (как дискретных, так и непрерывных), сохранив обозначения предыдущего параграфа.

Определение 8.3.3.1. *Математическим ожиданием* называется число, получаемое в дискретном случае по формулам

$$M\xi = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ji}x_i \ , \qquad M\eta = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ji}y_j \ ,$$

и в непрерывном –

$$M\xi = \int\limits_{-\infty}^{+\infty}\left(\int\limits_{-\infty}^{+\infty} xp(x,y)\,dx\right)\,dy \ ,$$

$$M\eta = \int\limits_{-\infty}^{+\infty}\left(\int\limits_{-\infty}^{+\infty} yp(x,y)\,dy\right)\,dx \ .$$

Аналогично используется

Определение 8.3.3.2. *Дисперсией* называется число, получаемое в дискретном случае по формулам

$$D\xi = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ji}(x_i - M\xi)^2 \ , \quad D\eta = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ji}(y_j - M\eta)^2 \ ,$$

и в непрерывном –

$$D\xi = \int\limits_{-\infty}^{+\infty}\left(\int\limits_{-\infty}^{+\infty}(x - M\xi)^2 p(x,y)\,dx\right)\,dy,$$

$$D\eta = \int\limits_{-\infty}^{+\infty}\left(\int\limits_{-\infty}^{+\infty}(y - M\eta)^2 p(x,y)\,dy\right)\,dx \ .$$

А также *средние квадратичные отклонения*: $\sigma_\xi = \sqrt{D\xi}$ и $\sigma_\eta = \sqrt{D\eta}$.

Эти характеристики обладают свойствами аналогичными одномерному случаю и не требуют самостоятельного рассмотрения. Однако, в случае системы двух (или большего числа) случайных величин также используются новые, специфические для многомерного случая, количественные оценки. Рассмотрим некоторые из них.

Определение 8.3.3.3. *Условным математическим ожиданием* или *регрессией* случайной величины ξ на η называется число, равное в дискретном случае сумме или ряду (или же интегралу – в непрерывном)

$$M_{\eta=y_j}\xi = \sum_{i=1}^{n} x_i \frac{p_{ji}}{q_j} \ , \qquad M_{\eta=y}\xi = \int\limits_{-\infty}^{+\infty} x\, p_{\eta=y}(x)\,dx \ .$$

Аналогично число, равное в дискретном случае сумме или ряду (или же интегралу – в непрерывном)

$$M_{\xi=x_i}\xi = \sum_{j=1}^{m} y_j \frac{p_{ji}}{p_i} \ , \qquad M_{\xi=x}\eta = \int\limits_{-\infty}^{+\infty} y\, q_{\xi=x}(y)\,dy \ ,$$

называется регрессией случайной величины η на ξ.

Нетрудно видеть, что данное определение совпадает с определением математического ожидания одномерной случайной величины, в котором вместо вероятности значений и плотности вероятности использованы соответствующие *условные* распределения, рассчитанные по формулам (8.3.2.1) и (8.3.2.2). А, поскольку математическое ожидание определяется при этом однозначно, то $M_{\eta=y}\xi$ и $M_{\xi=x}\eta$ можно рассматривать как функции, зависящие соответственно от y и x. Эти функции принято называть *функциями регрессии.*

Если функция регрессии $M_{\eta=y}\xi$ постоянна, то говорят, что случайная величина ξ *не коррелирует* со случайной величиной η, иначе эти случайные величины корреляционно зависимы.

Пример 8.3.3.1. Выяснить, коррелируют ли две дискретные случайные величины $\{\xi, \ \eta\}$, вероятности возможных значений которых приведены в следующей таблице

Значения: ξ η	$x_1 = -1$	$x_2 = 0$	$x_3 = 1$
$y_1 = 2$	0.1	0.25	0.15
$y_2 = 3$	0.3	0.1	0.1

Решение. Заметим вначале, что

$$q_1 = P(\eta = 2) = 0.1{+}0.25{+}0.15 = 0.5 \quad \text{и} \quad q_2 = P(\eta = 3) = 0.3{+}0.1{+}0.1 = 0.5 \ .$$

Тогда по определению 8.3.3.3, для условных математических ожиданий

$$M_{\eta=2}\xi = x_1\frac{p_{11}}{q_1} + x_2\frac{p_{12}}{q_1} + x_3\frac{p_{13}}{q_1} = (-1)\cdot\frac{0.1}{0.5} + 0\cdot\frac{0.25}{0.5} + 1\cdot\frac{0.15}{0.5} = 0.1$$

и

$$M_{\eta=3}\xi = x_1\frac{p_{21}}{q_2} + x_2\frac{p_{22}}{q_2} + x_3\frac{p_{23}}{q_2} = (-1)\cdot\frac{0.3}{0.5} + 0\cdot\frac{0.1}{0.5} \ + 1\cdot\frac{0.1}{0.5} = -0.4 \ .$$

Мы получили, что условное математическое ожидание ξ изменилось при изменении значения η. То есть случайная величина ξ находится в корреляционной взаимозависимости со случайной величиной η.

Согласно определению 8.3.2.2 случайные величины ξ и η считаются независимыми, если

$$P\big((\xi = x) \cdot (\eta = y) \big) \;\; = \;\; P\big((\xi = x) \big) \cdot P\big((\eta = y) \big) \; .$$

При этом возникает естественный вопрос: можно ли оценивать степень зависимости случайных величин по их количественным характеристикам, без использования распределений значений этих величин? Ответ таков: в большом числе практически важных случаев это удается сделать путем введения специальных количественных характеристик *системы* случайных величин. Приведем последовательно их определения.

Определение 8.3.3.4. *Ковариацией* или *корреляционным моментом* системы двух случайных величин ξ и η называется число

$$\operatorname{cov}(\xi, \eta) = M\big(\; (\xi - M\xi) \cdot (\eta - M\eta) \; \big) \; .$$

Для дискретных случайных величин

$$\operatorname{cov}(\xi, \eta) = \sum_{i=1}^{n} \sum_{j=1}^{m} (\xi - M\xi)\,(\eta - M\eta)\, p_{ji} \; ,$$

а для непрерывных

$$\operatorname{cov}(\xi, \eta) = \int\limits_{-\infty}^{+\infty} \left(\int\limits_{-\infty}^{+\infty} (x - M\xi)\,(y - M\eta)\, p(x, y)\, dy \right) \, dx \; .$$

Значение ковариации можно также подсчитывать по альтернативной, следующей из свойств математического ожидания (см. §8.1.3), формуле

$$\operatorname{cov}(\xi, \eta) = M\big(\; (\xi - M\xi) \cdot (\eta - M\eta) \; \big) = M(\; \xi \cdot \eta - \eta \cdot M\xi - \xi \cdot M\eta + M\xi \cdot M\eta \;) =$$

$$= M(\xi \cdot \eta) - M\xi \cdot M\eta \; .$$

А поскольку математическое ожидание произведения *независимых* случайных величин равно произведению математических ожиданий сомножителей, то величину ковариации можно попытаться использовать для количественной оценки степени взаимозависимости случайных величин.

Рассмотрим следующие примеры.

Пример 8.3.3.2. Вычислить ковариацию двух дискретных случайных величин $\{\xi,\ \eta\}$, вероятности возможных значений которых приведены в следующей таблице

Значения: ξ η	$x_1 = -1$	$x_2 = 0$	$x_3 = 2$
$y_1 = 1$	0.1	0.25	0.15
$y_2 = 3$	0.3	0.1	0.1

Решение. Чтобы воспользоваться формулой $\operatorname{cov}(\xi,\eta) = M(\xi\cdot\eta) - M\xi\cdot M\eta$, предварительно найдем значения $M\xi$, $M\eta$ и $M(\xi\cdot\eta)$. Поскольку

Значения: ξ	$x_1 = -1$	$x_2 = 0$	$x_3 = 2$
Вероятность $(\xi = x_i)$	0.4	0.35	0.25

и

Значения: η	$y_1 = 1$	$y_2 = 3$
Вероятность $(\eta = y_j)$	0.5	0.5

то

$$M\xi = (-1)\cdot 0.4 + 0\cdot 0.35 + 2\cdot 0.25 = 0.1 \quad \text{и} \quad M\eta = 1\cdot 0.5 + 3\cdot 0.5 = 2\ .$$

С другой стороны, распределение случайной величины $\kappa = \xi \cdot \eta$ имеет вид

Значение κ	-3	-1	0	2	6
Вероятность события $(\kappa = z_k)$	0.3	0.1	0.35	0.15	0.1

Поэтому

$M\kappa = M(\xi \cdot \eta) = (-3) \cdot 0.3 + (-1) \cdot 0.1 + 0 \cdot 0.35 + 2 \cdot 0.15 + 6 \cdot 0.1 = -0.1$.

Наконец, по формуле для ковариации находим

$$\operatorname{cov}(\xi, \eta) = M(\xi \cdot \eta) - M\xi \cdot M\eta = -0.1 - 0.1 \cdot 2 = -0.3 \ .$$

Анализируя формулу $\operatorname{cov}(\xi, \eta) = M(\xi \cdot \eta) - M\xi \cdot M\eta$ нетрудно понять, как при помощи ковариации можно оценить степень зависимости случайных величин ξ и η. Действительно, если эти величины *независимые,* то согласно свойству 3° для математического ожидания (см. § 8.1.3) выполняется равенство $M(\xi \cdot \eta) = M\xi \cdot M\eta$, то есть ковариация равна нулю. Таким образом, условие $\operatorname{cov}(\xi, \eta) = 0$ очевидно *необходимое* для независимости случайных величин ξ и η. Однако, как показывает следующий пример, это условие не является *достаточным.*

Пример 8.3.3.3. Пусть случайная величина ξ имеет распределение вида

Значение ξ	-2	-1	1	2
Вероятность события $(\xi = x_k)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

При этом случайная величина $\eta = \xi^2$. Ее распределение будет

Значение η	1	4
Вероятность события $(\eta = y_k)$	$\frac{1}{2}$	$\frac{1}{2}$

Очевидно, что $M\xi = 0$. Проверим также справедливость $M(\xi \cdot \eta) = 0$. Действительно, распределение случайной величины $\kappa = \xi \cdot \eta$ имеет вид

Значение κ	-8	-4	-2	-1	1	2	4	8
Вероятность события $(\kappa = z_k)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Поэтому

$$M\kappa = M(\xi \cdot \eta) = \frac{(-8) + (-4) + (-2) + (-1) + 1 + 2 + 4 + 8}{8} = 0\ .$$

и, значит, $\operatorname{cov}(\xi, \eta) = 0$. Однако в рассматриваемом случае ξ и η связаны функционально и потому зависимы.

Теперь рассмотрим возможность использования ковариации для оценки *степени зависимости* случайных величин. Сама по себе «малость» числа $M\big((\xi - M\xi) \cdot (\eta - M\eta)\big)$, вообще говоря, не является гарантирующим признаком независимости случайных величин ξ и η, поскольку эта «малость» может также обусловливаться и узкими диапазонами отклонения их значений от математического ожидания. Поэтому для оценки степени зависимости случайных величин предпочтительно использовать нормированную ковариацию, называемую *коэффициентом корреляции* – численную характеристику равную

$$r(\xi, \eta) = \frac{\operatorname{cov}(\xi, \eta)}{\sigma_\xi \cdot \sigma_\eta}\ ,$$

где $\sigma_\xi = \sqrt{D\xi}$ и $\sigma_\eta = \sqrt{D\eta}$ – средние квадратичные отклонения ξ и η.

Перечислим основные свойства коэффициента корреляции.

1°. $-1 \le r(\xi, \eta) \le 1$.

2°. Если ξ и η независимы, то $r(\xi, \eta) = 0$.

3°. Если $r(\xi, \eta) \ne 0$, то ξ и η зависимы.

4°. Если $\eta = a \cdot \xi + b$, то

$$r(\xi, \eta) = \left[\begin{array}{rl} -1, & \text{если } a < 0\ , \\ 0, & \text{если } a = 0\ , \\ 1, & \text{если } a > 0\ , \end{array}\right.$$

то есть коэффициент корреляции является мерой степени *линейной зависимости* случайных величин, что и иллюстрирует пример 8.3.3.3.

Пример 8.3.3.4. Вычислить коэффициент корреляции двух дискретных случайных величин $r(\xi,\ \eta)$, вероятности возможных значений которых приведены в условии примера 8.3.3.2.

Решение. Чтобы найти коэффициент корреляции, помимо найденной в примере 8.3.3.2 ковариации, нужны значения средних квадратичных отклонений случайных величин ξ и η. Поскольку $\sigma_\xi = \sqrt{D\xi}$ и $\sigma_\eta = \sqrt{D\eta}$, найдем сначала дисперсии. В нашем случае

Значения: ξ^2	$x_1 = 0$	$x_2 = 1$	$x_3 = 4$
Вероятность $(\xi^2 = x_i)$	0.35	0.4	0.25

то

$$M(\xi^2) = 0{\cdot}0.35{+}1{\cdot}0.4{+}4{\cdot}0.25 = 1.4 \ \text{ и } \ D\xi = M(\xi^2){-}(M\xi)^2 = 1.4{-}(0.1)^2 = 1.39\ .$$

Аналогично по распределению

Значения: η^2	$y_1 = 1$	$y_2 = 9$
Вероятность $(\eta^2 = y_j)$	0.5	0.5

находим, что

$$M(\eta^2) = 1 \cdot 0.5 + 9 \cdot 0.5 = 5 \quad \text{и} \quad D\eta = M(\eta^2) - (M\eta)^2 = 5 - 2^2 = 4\ .$$

Наконец, находим коэффициент корреляции по формуле

$$r(\xi,\eta) = \frac{\operatorname{cov}(\xi,\eta)}{\sigma_\xi \cdot \sigma_\eta} = \frac{\operatorname{cov}(\xi,\eta)}{\sqrt{D\xi}\cdot\sqrt{D\eta}} = \frac{-\ 0.3}{\sqrt{1.39}\cdot\sqrt{1}} \approx -0.25\ .$$

Для системы двух случайных величин ξ и η *совместной* количественной характеристикой разброса и степени взамозависимости их значений является *ковариационная матрица,* задаваемая следующим образом

$$\|K\| = \left\| \begin{array}{cc} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{array} \right\| = M \left(\left\| \begin{array}{c} \xi - M\xi \\ \eta - M\eta \end{array} \right\| \| \, \xi - M\xi \quad \eta - M\eta \, \| \right) =$$

$$= \left\| \begin{array}{cc} M(\xi - M\xi)^2 & M((\xi - M\xi)\cdot(\eta - M\eta)) \\ M((\eta - M\eta)\cdot(\xi - M\xi)) & M(\eta - M\eta)^2 \end{array} \right\| =$$

$$= \left\| \begin{array}{cc} D\xi & \operatorname{cov}(\xi,\eta) \\ \operatorname{cov}(\xi,\eta) & D\eta \end{array} \right\| .$$

Если же требуется совместно оценить степени разброса и взамозависимости значений системы двух случайных величин, то можно использовать *корреляционную матрицу,* элементами которой служат числа

$$\rho_{11} = \frac{\kappa_{11}}{\sigma_\xi \cdot \sigma_\xi} = 1 \ , \qquad \rho_{12} = \frac{\kappa_{12}}{\sigma_\xi \cdot \sigma_\eta} = r(\xi,\eta) \ ,$$

$$\rho_{21} = \frac{\kappa_{21}}{\sigma_\xi \cdot \sigma_\eta} = r(\xi,\eta) \ , \quad \rho_{22} = \frac{\kappa_{22}}{\sigma_\eta \cdot \sigma_\eta} = 1 \ .$$

Значит, эта матрица имеет вид

$$\|R\| = \left\| \begin{array}{cc} \rho_{11} & \rho_{12} \\ \rho_{211} & \rho_{22} \end{array} \right\| = \left\| \begin{array}{cc} 1 & r(\xi,\eta) \\ r(\xi,\eta) & 1 \end{array} \right\| ,$$

где $r(\xi,\eta)$ – коэффициент корреляции случайных величин ξ и η.

Кроме того, на практике для оценки степени взаимозависимости значений системы случайных величин используются также *упрощенные* числовые характеристики, образуемые из элементов ковариационной матрицы

$$\operatorname{tr}\|K\| = \kappa_{11} + \kappa_{22} \quad \text{и} \quad \det\|K\| = \kappa_{11}\kappa_{22} - \kappa_{12}\kappa_{21} \ ,$$

последняя из которых называется *обобщенной дисперсией.*

В заключение отметим, что, хотя любые количественные характеристики систем случайных величин находятся по их распределениям, сами распределения иногда (скажем, когда тип распределения известен) удобно записывать с помощью количественных характеристик. Например, *совместная плотность вероятности двух (вообще говоря, зависимых) случайных величин* ξ *и* η*, распределенных по нормальному закону* дается формулой

$$p(x,y) = \frac{1}{2\pi\sigma_\xi\sigma_\eta\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-M\xi)^2}{\sigma_\xi^2} - 2\rho\frac{(x-M\xi)(y-M\eta)}{\sigma_\xi\sigma_\eta} + \frac{(y-M\eta)}{\sigma_\eta^2}\right.}$$

где $\rho = r(\xi,\eta)$. Графическое представление зависимости $z = p(x,y)$ показано на рис.4.25.

Проверьте самостоятельно, что для *независимых* нормально распределенных случайных величин, то есть при $\rho = 0$, данная формула упрощается и приобретает соответствующий определению 8.2.1.2 и формуле (8.2.1.5) вид

$$p(x,y) = p(x)\cdot p(y)\ .$$

8.4. Закон больших чисел и центральная предельная теорема

8.4.1. Закон больших чисел

Очевидно, что результаты некоторой пары одинаковых независимых испытаний являющихся случайными событиями, могут значительно отличаться друг от друга. Вместе с тем, с древнейших времен было известно, что для *больших* совокупностей таких испытаний могут наблюдаться устойчивые статистические закономерности. Например, при многократном последовательном взвешивании тела среднее арифметическое значений веса практически перестает изменяться с ростом числа взвешиваний. Или же, относительная частота выпадения «герба вверх» в серии последовательных подбрасываний симметричной монеты имеет устойчивую тенденцию приближения к $\frac{1}{2}$ при увеличении числа бросаний. Иначе говоря, случайные отклонения, присущие каждому отдельному испытанию, в среднем взаимно компенсируются, погашаются для больших объемов этих испытаний.

Эти наблюдаемые свойства устойчивости (в тех случаях, когда они имеют *реальное физическое или статистическое обоснование*) позволяют делать теоретическое заключение о *существовании вероятностей* тех или иных случайных событий. Практическая же суть этой устойчивости, которую принято называть *законом больших чисел* состоит в том, что средний результат большого числа однородных испытаний со случайным исходом, перестает быть случайным, и его можно предсказать с высокой степенью достоверности.

Формально термин «закон больших чисел» обозначает набор теорем в курсе теории вероятностей, обосновывающих (при выполнении определенных условий) факт приближения средних характеристик большой по объему совокупности значений случайных величин к некоторым константам. Рассмотрим основные из этих теорем.

Пусть имеется случайная величина ξ с математическим ожиданием $M\xi$ и дисперсией $D\xi$. Проведем серию из n испытаний, каждое из которых дает некоторое значение ξ_k , $k = [1, n]$. Для системы независимых и очевидно одинаково распределенных случайных величин ξ_k введем новую случайную величину η по формуле

$$\eta = \frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} = \frac{1}{n}\sum_{k=1}^{n} \xi_k \ .$$

Тогда будет иметь место

Теорема 8.4.1.1. **Теорема Чебышева.**

Среднее арифметическое наблюденных значений случайной величины сходится по вероятности к ее математическому ожиданию, то есть,

$$\forall\, \varepsilon > 0 : \quad \lim_{n\to+\infty} P\left(\left| \frac{\xi_1 + \xi_2 + \ldots + \xi_n}{n} - M\xi \right| < \varepsilon \right) = 1 \ .$$

Доказательство. Убедимся сначала в справедливости так называемого *неравенства Чебышева,* имеющего вид

$$\forall\, \alpha > 0 : \quad P\Big(\mid \eta - M\eta \mid \geq \alpha \Big) \leq \frac{D\eta}{\alpha^2}$$

для некоторой дискретной случайной величины ξ имеющей распределение вида

Значение ξ	x_1	x_2	$\dots$	x_N	$\dots$
Вероятность события $(\xi = x_k)$	p_1	p_2	$\dots$	p_N	$\dots$

Для некоторого конкретного α вероятность события, заключающегося в том, что значение случайной величины ξ отклонится от ее математического ожидания $M\xi$ не меньше, чем на α очевидно равна

$$P(\,|\,\xi - M\xi\,| \geq \alpha\,) \quad = \sum_{|x_k - M\xi| \geq \alpha} p_k \ ,$$

где суммирование в правой части выполняется только по тем k, для которых $|x_k - M\xi| \geq \alpha$.

Поскольку величина дисперсии случайной величины ξ равна

$$D\xi = M(\xi - M\xi)^2 = \sum_{k=1}^{n} |x_k - M\xi|^2 p_k \ ,$$

то путем отбрасывания части неотрицательных слагаемых в этой сумме получаем оценку

$$D\xi \geq \sum_{|x_k - M\xi| \geq \alpha} |x_k - M\xi|^2 p_k \geq \sum_{|x_k - M\xi| \geq \alpha} \alpha^2 p_k = \alpha^2 \cdot P(\,|\,\xi - M\xi\,| \geq \alpha\,) \ .$$

Откуда и следует неравенство Чебышева.

Найдем теперь $M\eta$ и $D\eta$, воспользовавшись свойством 2° для математического ожидания и свойством 3° для дисперсии (см. §8.1.3).

$$M\eta = M\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n}\,M\left(\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n}\sum_{k=1}^{n} M\xi_k = \frac{1}{n}\cdot n \cdot M\xi = M\xi \ .$$

$$D\eta = D\left(\frac{1}{n}\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n^2}\,D\left(\sum_{k=1}^{n}\xi_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} D\xi_k = \frac{1}{n^2}\cdot n \cdot D\xi = \frac{D\xi}{n} \ .$$

Таким образом при $n \to +\infty$ математическое ожидание $M\eta = M\xi$, а дисперсия $D\eta \to 0$. Наконец, воспользовавшись неравенством Чебышева, получим для случайной величины η

$$\forall\, \alpha > 0: \quad P\Bigl(\ |\ \eta - M\xi\ |\ \geq \alpha\Bigr) \leq \frac{D\xi}{n \cdot \alpha^2}\,.$$

И поскольку за счет увеличения n вероятность

$$P\Bigl(\ |\ \eta - M\xi\ |\ \geq \alpha\Bigr)$$

может быть сделана сколь угодно малой, то мы приходим к заключению о справедливости утверждения теоремы.

По поводу неравенства Чебышева следует сделать два замечания. Во-первых, оно справедливо и для непрерывных случайных величин. Доказательство аналогично дискретному случаю, только вместо вероятностей следует использовать плотность вероятности, а суммирование заменить интегрированием.

Во-вторых, необходимо понимать, что это неравенство дает оценку, которая остается верной при *любых* видах распределения случайной величины. Однако, эта оценка может оказаться достаточно грубой. Например, применив неравенство Чебышева для нормально распределенной случайной величины ξ, при значении $\alpha = 3\sigma_\xi$, мы получим

$$P\Bigl(\ |\ \eta - M\xi\ |\ \geq 3\sigma_\xi\Bigr) \leq \frac{D\xi}{9\sigma_\xi^2} = \frac{1}{9}\,,$$

в то время как для нормального распределения вероятность отклонения от математического ожидания на величину «больше, чем три сигма» равна приблизительно 0,003.

Теорема Чебышева допускает обобщение как для системы независимых случайных величин с различными распределениями, так и для системы зависимых случайных величин (так называемая *теорема Маркова*) . К практически важным следствиям теоремы Чебышева относятся утверждения:

1°. В схеме испытаний Бернулли относительная частота «успеха» – события A, при большом числе испытаний дает приближенную оценку вероятности $p = P(A)$.

2°. Среднее арифметическое значений некоторого количественного признака при случайной выборке значений из генеральной совокупности может быть использовано в качестве оценки математического ожидания этого признака.

8.4.2. Центральная предельная теорема

Закон больших чисел позволяет оценивать значения количественных характеристик случайных величин. Однако также достаточно часто возникает необходимость оценки предельного представления их функции распределения. Решение этой задачи основывается на использовании *центральной предельной теоремы.* Основным условием применимости этой теоремы является возможность представления исследуемой случайной величины в виде суммы большого числа независимых (или слабо зависимых) случайных величин, каждая из которых сравнительно мало влияет на общую сумму.

Будет правильнее сказать, что центральная предельная теорема представляет собой набор нескольких, близких по формулировкам утверждений. Рассмотрим основную из них.

Для системы независимых и одинаково распределенных случайных величин ξ_k с математическим ожиданием a и дисперсией σ^2 введем новую случайную величину η

$$\eta_n = \xi_1 + \xi_2 + \ldots + \xi_n = \sum_{k=1}^{n} \xi_k \ .$$

Тогда будет справедлива

Теорема 8.4.2.1. **Центральная предельная теорема.**

При неограниченном росте числа слагаемых n закон распределения случайной величины η_n будет неограничено приближаться к *нормальному закону распределения* с параметрами a и σ^2.

Напомним, что формулы *нормального* и *стандартного нормального* распределения соответственно имеют вид

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\dfrac{(t-a)^2}{2\sigma^2}} \, dt \ , \qquad F^*(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\dfrac{t^2}{2}} \, dt \ .$$

Центральная предельная теорема может использоваться и для систем дискретных случайных величин. Например, пусть η число появления события A в схеме Бернулли, состоящей из n независимых испытаний, в каждом из которых событие A происходит с вероятностью p и не происходит с вероятностью $q = 1 - p$. Тогда справедлива оценочная *формула Муавра-Лапласа*

$$P\left(\alpha < \frac{\eta - np}{\sqrt{npq}} < \beta\right) = F^*(\beta) - F^*(\alpha)\ .$$

Пример 8.4.2.1. Найти вероятность того, что число выпадений «герба вверх» в серии из 100 бросаний симметричной монеты окажется в пределах от 60 до 70.

Решение. В данном случае $n = 100$ и $p = q = \dfrac{1}{2}$. Тогда при $60 \le \eta \le 70$ параметр

$$\xi = \frac{\eta - np}{\sqrt{npq}} = \frac{\eta - 50}{5}$$

будет удовлетворять условию $2 \le \xi \le 4$, и, согласно формуле Муавра-Лапласа, оценка искомой вероятности составит

$$P(60 \le \eta \le 70) = P(2 \le \xi \le 4) = F^*(4) - F^*(2) = \frac{1}{\sqrt{2\pi}} \int\limits_2^4 e^{-\dfrac{x^2}{2}}\, dx \approx 0.023\ .$$

Глава 9.

Элементы математической статистики

Математической статистикой называется раздел математики, изучающий закономерности массовых свойств, процессов и явлений, получаемые путем наблюдений. К основным задачам математической статистики относятся разработка методологии сбора и систематизации результатов наблюдений, равно как и средств их анализа и оценки достоверности. Инструментальной основой математической статистики является теория вероятностей и, в первую очередь, законы больших чисел и предельные теоремы.

Термин *массовое явление* означает, что предметом исследования является множество объектов или явлений, обладающих некоторым общим свойством (или набором свойств), подобным для всех элементов этого множества, называемого *генеральной совокупностью.* В ряде случаев проведение наблюдений для генеральной совокупности оказывается либо затруднительным, либо вовсе невозможным. Тогда соответствующие исследования проводятся на одном или нескольких подмножествах генеральной совокупности, которые принято называть *выборочными совокупностями* или просто *выборками. Объемом выборки,* как генеральной, так и выборочной, называют число элементов, ее составляющих.

Следует отметить, что в общем случае статистические результаты, получаемые для некоторой выборки, могут не отражать соответствующие свойства генеральной совокупности. Корректность переноса результатов наблюдений с выборочной совокупности на генеральную требует обоснования в каждом конкретном случае. Если выборка правильно отражает соотношения в генеральной совокупности, то ее называют *представительной* или *репрезентативной.* На практике *необходимым (но не достаточным) условием* репрезентативности выборки является ее *случайность.*

9.1. Статистические оценки случайных величин

Достаточно часто исследуемый признак или свойство элементов генеральной совокупности описывается некоторой числовой характеристикой, которая представляет собой дискретную или непрерывную случайную величину ξ, тогда центральной задачей математической статистики является изучение (по предположению, существующего) закона распределения этой случайной величины.

Пусть из генеральной совокупности извлекается выборка объема n, и пусть при этом значение x_k наблюдалось n_k раз. Очевидно, что

$$n = \sum_{k=1}^{n} n_k \; .$$

Дадим определения основных терминов, используемых в математической статистике при изучении свойств случайных числовых характеристик.

Определение 9.1.1. Наблюдаемые значения случайной величины ξ называются *вариантами.* Последовательность вариант, записанных в порядке возрастания, называется *вариационным рядом.* Число n_k называется *частотой* или *частостью* наблюдения варианты x_k. Соответственно отношение $p_k^* = n_k/n$ называется *относительной частотой* этой варианты.

Определение 9.1.2. *Статистическим распределением* выборки называется таблица (или правило) сопоставления значений вариант и соответствующих им частот(или относительных частот.)

Определение 9.1.3. *Эмпирической функцией распределения* выборки называется функция $F(x)$, значение которой равно относительной частоте наблюдений значений ξ, удовлетворяющих условию $\xi < x$. Иначе говоря,

$$F(x) = \sum_{x_k < x} \frac{n_k}{n} = \frac{1}{n} \sum_{x_k < x} n_k \ .$$

На практике обычно делается предположение (высказывается гипотеза) о виде закона распределения для генеральной совокупности наблюдаемой случайной величины, значения некоторых параметров которого считаются изначально неизвестными. Например, предполагается, что ξ распределена по нормальному закону вида

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\dfrac{(t-a)^2}{2\sigma^2}}\, dt \ ,$$

с неизвестными *a priori* значениями a и σ. Затем по выборкам из генеральной совокупности оцениваются значения этих параметров, которые рассматриваются как некоторые новые случайные величины.

Пусть θ^* есть случайная величина, оценивающая значения некоторого параметра θ в законе распределения случайной величины ξ, а θ_n^* – его конкретная оценка, построенная по результатам n случайных выборок. В этом случае можно попытаться найти значение θ по числовым характеристикам случайной величины θ^*, Однако, что гарантировать успех подобной процедуры в общем случае нельзя. Поэтому в математической статистике используются специальные понятия, характеризующие *степень качества* данной оценки.

Определение 9.1.4. Статистическая оценка θ^* параметра θ называется *несмещенной,* если $M\theta^* = \theta$. Иначе, говорят о *смещенной* оценке.

Определение 9.1.5. Статистическая оценка называется *эффективной,* если при фиксированном объеме выборки, она имеет наименьшую дисперсию.

Определение 9.1.6. Статистическая оценка называется *состоятельной,* если ее значение при неограниченном увеличении n – объема выборки, стремится по вероятности к оцениваемому параметру, то есть если

$$\forall\, \varepsilon > 0 \quad \lim_{n\to\infty} P\Big(\ |\ \theta_n^* - \theta\ | < \varepsilon\ \Big) = 1\ .$$

В частности, если дисперсия оценки стремится к нулю при $n \to \infty$, то такая оценка является состоятельной.

9.2. Выборочные, точечные и интервальные оценки

Определение 9.2.1. *Выборочной средней* оценкой случайной величиной ξ при выборке объема n называется число

$$\overline{M\xi} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{k=1}^{n} x_k \ .$$

Определение 9.2.2. *Выборочной дисперсией* случайной величиной ξ при выборке объема n называется среднее арифметическое квадратов отклонений наблюдаемых значений от выборочной средней, то есть число

$$\overline{D\xi} = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{M\xi} \right)^2 \ .$$

Определение 9.2.2. *Выборочным средним квадратичным отклонением* случайной величиной ξ при выборке объема n называется число $\overline{\sigma_\xi} = \sqrt{\overline{D\xi}}$.

Выборочная средняя $\overline{M\xi}$ является несмещенной оценкой математического ожидания. Действительно, пусть дана выборка $\{x_k,\ x_2,\ \ldots\ x_n\}$. Будем рассматривать эти выборочные значения как реализации n случайных величин, одинаково распределенных по закону распределения генеральной совокупности. Значит эти новые случайные величины имеют одинаковые математические ожидания $M\xi$ и дисперсии $D\xi$. Поэтому математическое ожидание выборочной средней будет равно

$$M\left(\overline{M\xi}\right) = M\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) = n \cdot \frac{M\xi}{n} = M\xi \ ,$$

что и доказывает несмещенность выборочной средней.

Теперь рассмотрим выборочную дисперсию

$$\overline{D\xi} = \frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{M\xi} \right)^2 .$$

Ее математическое ожидание можно представить в виде

$$M\left(\overline{D\xi}\right) = M\left(\frac{1}{n}\sum_{k=1}^{n}\left(x_k - \overline{M\xi} \right)^2\right) =$$

$$= M\left(\frac{1}{n}\sum_{k=1}^{n}\left((x_k - M\xi) - (\overline{M\xi} - M\xi) \right)^2\right) =$$

$$= M\left(\frac{1}{n}\sum_{k=1}^{n}\left(x_k - M\xi \right)^2\right) - M\left(\frac{2}{n}\sum_{k=1}^{n}\left((x_k - M\xi)\cdot(\overline{M\xi} - M\xi) \right)\right) +$$

$$+ M\left(\frac{1}{n}\sum_{k=1}^{n}\left(\overline{M\xi} - M\xi \right)^2\right) .$$

Рассмотрим теперь каждое слагаемое в последней формуле по отдельности.

$$M\left(\frac{1}{n}\sum_{k=1}^{n}\left(x_k - M\xi \right)^2\right) = \frac{1}{n}\sum_{k=1}^{n} M\left(x_k - M\xi \right)^2 = \frac{n\cdot D\xi}{n} = D\xi .$$

Аналогично

$$M\left(\frac{2}{n}\sum_{k=1}^{n}(x_k - M\xi)\cdot(\overline{M\xi} - M\xi)\right) = M\left(\frac{2(\overline{M\xi} - M\xi)}{n}\sum_{k=1}^{n}\left(x_k - M\xi \right)\right) =$$

$$= \frac{2(\overline{M\xi} - M\xi)}{n}\sum_{k=1}^{n} M(x_k - M\xi) = 2(\overline{M\xi} - M\xi)^2 =$$

$$= 2\overline{D\xi} = 2D\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) = 2n\cdot\frac{D\xi}{n^2} = 2\frac{D\xi}{n} .$$

Наконец, использовав те же преобразования, получим

$$M\left(\frac{1}{n}\sum_{k=1}^{n}\left(\overline{M\xi} - M\xi \right)^2\right) = \frac{D\xi}{n} .$$

В итоге
$$M\left(\overline{D\xi}\right) = D\xi - \frac{D\xi}{n} = \frac{n-1}{n} \cdot D\xi \; .$$
То есть математическое ожидание выборочной дисперсии не равно дисперсии генеральной совокупности и данная оценка является смещенной. Хотя стоит отметить, что данное смещение определяется множителем $\dfrac{n-1}{n}$, мало отличающимся от единицы при больших n.

При практических расчетах удобнее пользоваться равенством
$$\overline{D\xi} = \overline{\xi^2} - \left(\overline{\xi}\right)^2 \; ,$$
которое означает, что выборочная дисперсия равна среднему арифметическому квадратов значений выборки минус квадрат средней выборочной.

Помимо выборочных в математической статистике используются и другие типы оценок.

Определение 9.2.3. *Точечной* называют оценку, которая задается одним числом.

Поскольку при выборке малого объема точечная оценка может существенно отличаться от оцениваемого параметра, то в таких случаях целесообразно использовать не точечные, а *интервальные* оценки, то есть пары чисел, являющихся границами промежутка числовой оси, которому принадлежит значение параметра.

Определение 9.2.4. *Доверительной вероятностью* или *надежностью* оценки θ^* параметра θ называется вероятность
$$\gamma = P(\; |\; \theta - \theta^*| < \delta \;) \; .$$
Доверительным интервалом с надежностью γ в этом случае называется интервал $(\theta^* - \delta, \, \theta^* + \delta)$.

Параметры доверительных интервалов для значений случайных величин обычно находятся при помощи их законов распределения.