МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ»

кафедра высшей математики

В. Ю. Дорофеев

МАТЕМАТИКА и СТАТИСТИКА

Санкт-Петербург

2018

УДК 510.2
Д 694

Рекомендовано научно-методическим советом университета

Дорофеев В. Ю.
Математика и Статистика
В. Ю. Дорофеев. – СПб: Издательство СПбГЭУ, 2018 – 137 с.

Учебное пособие предназначено для студентов гуманитарного факультета по направлению рекламы и связи с общественностью соответствует программе «Математика и Статистика» кафедры высшей математики СПбГЭУ по специальности 42.03.01.

Рецензент: д-р. физ.-мат. н., Ю. В. Павлов

Оглавление

Глава 1

Основы аналитической геометрии

1.1 Расстояние между двумя точками

Будем считать, что мы знаем такие геометрические объекты как точка, прямая, плоскость и пространство, что такое направление и угол.

Проведём на плоскости прямую слева направо, которая, по нашему мнению, как геометрический объект состоит из бесконечного числа точек. Отметим на этой прямой точку. Обозначим её буквой О. Сопоставим этой точке число ноль. Справа от нуля отметим ещё одну точку. Этой точке сопоставим число один. После чего каждой точке справа сопоставим положительное число, соблюдая выбранную пропорцию по отношению к первой точке, а каждой точке слева – отрицательное. Новый объект – геометрическая прямая, на которой отмечены указанным выше образом числа, называется *числовой прямой* или *числовой осью*. Таким образом устанавливается взаимно-однозначное соответствие между точками прямой и всеми действительными числами.

Отметим на выделенной прямой пару точек как геометрические объекты. Одну точку обозначим буквой A, другую – буквой B. Но так как прямая числовая, то мы автоматически попадаем в два числа. Пусть это будут числа x_A и x_B соответственно. Числа, которые сопоставляются геометрической точке на числовой прямой называются *координатами этой точки*. Числовая прямая называется также *координатной* *прямой*, а предложенное выше построение – *системой координат на прямой* (рис. 1).

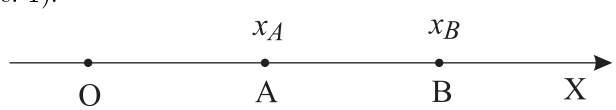


рис. 1.

Распространим это представление на плоскость. Обозначим имеющуюся прямую как прямую OX и назовём её осью абсцисс. Перпендикулярно ей через точку O проведём другую координатную прямую снизу вверх, считая что под точкой O расположены отрицательные числа, а сверху – положительные. Новую прямую обозначим как OY и назовём её осью ординат.

Отметим на плоскости как на геометрическом объекте точку A. Через эту точку проведём две перпендикулярные прямые, параллельные осям координат. Эти прямые пересекут координатные оси в некоторых точках. Точку на оси абсцисс обозначим как x_A, а на оси ординат как y_A. Из геометрических соображений понятно, что каждой точке плоскости будет соответствовать одна упорядоченная пара чисел (x_A, y_A) и наоборот: каждой паре чисел соответствует одна точка плоскости. Таким образом мы установили взаимно однозначное соответствие между точками на плоскости и упорядоченными парами чисел. Эта пара чисел называется координатами точки на плоскости, а таким образом построенные координатные оси – *прямоугольной декартовой системой координат на плоскости* (рис. 2).

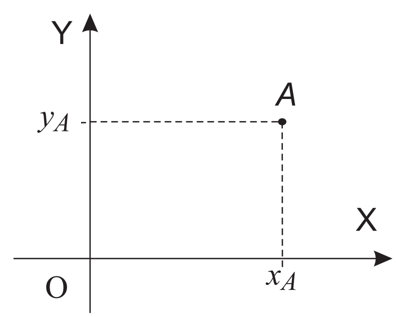
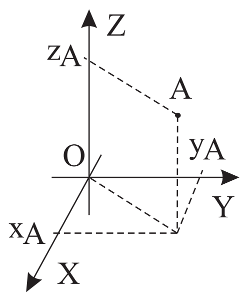


рис. 2.

Наконец, проведя третью координатную прямую OZ через точку O перпендикулярно плоскости XOY так как указано на рис. 3 (ось OX смотрит на нас, ось OY направлена слева направо и ось OZ – снизу вверх), мы получаем прямоугольную систему координат в пространстве. Теперь каждой точке пространства сопоставляется упорядоченная тройка чисел (x, y, z).

Важным моментом построения, проводимого выше, является взаимная однозначность между геометрической и алгебраической конструкциями. Но требование взаимной однозначности не нарушится, если координатные прямые проводить не под прямым углом. Такая конструкция тоже допустима и называется косоугольной декартовой системой координат. На рис. 4 изображена косоугольная система координат на плоскости. Однако она менее удобна и в дальнейшем ею мы пользоваться не будем.



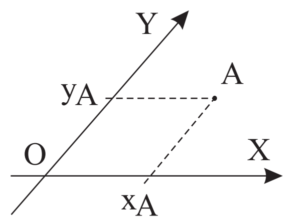


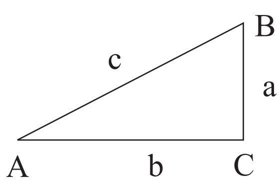
рис. 3. рис. 4.

Всякой паре точек, например A и B, на координатной прямой соответствует пара чисел x_A и x_B соответственно (рис. 1). Определим величину $\rho(A, B)$, которую назовём *расстоянием между точками на прямой*, формулой

$$\rho(A, B) = |x_B - x_A|. \tag{1.1}$$

Для упрощения обозначений будем считать допустимой запись расстояния между двумя точками как $|AB|$, полагая $\rho(A, B) = |AB|$.

В геометрии Евклида доказывается теорема Пифагора: *в прямоугольном треугольнике квадрат гипотенузы равен сумме квадратов катетов:* $c^2 = a^2 + b^2$ (в обозначениях на рис. 5).



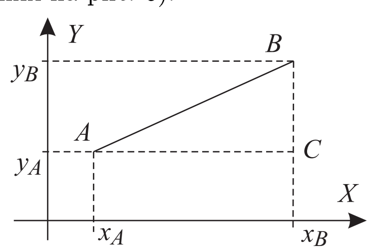


рис. 5. рис. 6.

Из рис. 6 и формулы (1.1) следуют равенства

$$|AC| = |x_B - x_A|, \quad |BC| = |y_B - y_A|. \tag{1.2}$$

Осталось сравнить рисунки 5 и 6 и применить теорему Пифагора. В результате получим искомую формулу:

$$\rho(A,B) = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}. \qquad (1.3)$$

Используя дополнительные построения к рис. 3 и формулу (1.3), нетрудно найти формулу расстояния между двумя точками в пространстве

$$\rho(A,B) = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}, \qquad (1.4)$$

где $A = (x_A, y_A, z_A), B = (x_B, y_B, z_B)$.

Например, если точка A имеет координаты (2, 7), а точка B имеет координаты (–1, 3), то расстояние между точками A и B на плоскости равно

$$|AB| = \sqrt{(-1-2)^2 + (3-7)^2} = \sqrt{9+16} = 5.$$

Аналогично расстояние между точками $A(2,-1,5)$ и $B(1,2,3)$ в пространстве находится так

$$|AB| = \sqrt{(1-2)^2 + (2+1)^2 + (3-5)^2} = \sqrt{1+9+4} = \sqrt{14}.$$

1.2 Уравнение прямой на плоскости

Пусть на плоскости имеется две точки A и B. Из курса евклидовой геометрии известно, что через две точки проходит только одна прямая.

Обозначим эту прямую как (AB) и найдём её уравнение в прямоугольной декартовой системе координат XOY. Для этого через точки A и B проведём прямые, параллельные осям координат. Эти прямые пересекут координатные оси в некоторых точках. Точки на оси абсцисс обозначим как x_A, x_B, а на оси ординат как y_A, y_B.

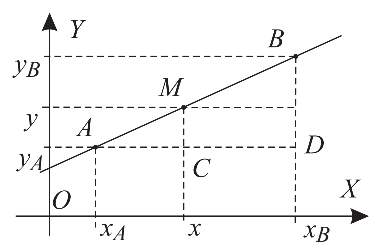


рис. 7.

Тем самым определены координаты точек $A = (x_A, y_A)$ и $B = (x_B, y_B)$ на плоскости XOY, через которые проведём прямую (AB). На этой прямой отметим некоторую точку $M = (x, y)$. Из подобия треугольников ABD и AMC получаем следующее равенство:

$$\frac{x - x_A}{x_B - x_A} = \frac{y - y_A}{y_B - y_A}. \qquad (1.5)$$

Точка M как бы пробегает всю прямую, поэтому её координаты называются *текущими координатами.* Уравнение (1.5) называется *уравнением прямой, проходящей через две точки.* Его можно преобразовать к виду

$$y - y_A = \frac{y_B - y_A}{x_B - x_A}(x - x_A).$$

Вводя обозначения

$$k = \frac{y_B - y_A}{x_B - x_A} \tag{1.6}$$

и перенося постоянную y_A в правую часть равенства, получим

$$y = k(x - x_A) + y_A \tag{1.7}$$

– *уравнением прямой, проходящей через заданную точку (здесь это точка $A(x_A, y_A)$) в заданном направлении.* Если точку A зафиксировать и менять коэффициент k, то уравнение (1.7) имеет смысл *уравнения пучка прямых.* При этом нетрудно убедиться, что коэффициент k имеет *геометрический смысл* тангенса угла наклона прямой. Раскроем в (1.7) скобки и приведём подобные. Уравнение принимает вид

$$y = kx + b \tag{1.8}$$

и называется *уравнением прямой с угловым коэффициентом.*

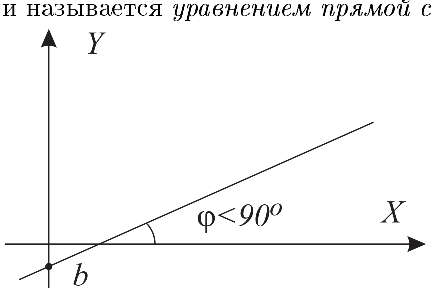


рис. 8.

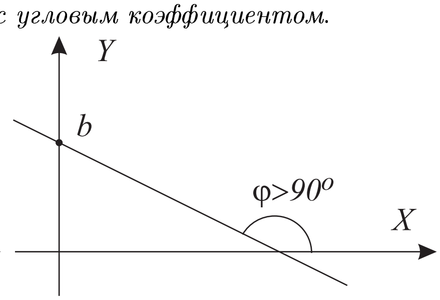


рис. 9.

В этом уравнении *угловой коэффициент* – это число k, а число b – *свободный член* имеет смысл точки пересечения прямой с осью ординат (рис. 8). В соответствии с общим смыслом коэффициента k его знак определяет угол наклона прямой по отношению к оси OX: если коэффициент k положительный, то угол острый (рис. 8); если коэффициент k отрицательный, то угол тупой (рис. 9).

В частном случае $k = 0$ получим уравнение прямой $y = b$, параллельной оси OX. При $b = 0$, получим прямую, проходящую через начало координат, но уравнение прямой с угловым коэффициентом (1.7) не содержит прямой, параллельной оси ординат, так как уравнение этой прямой имеет вид $x = x_0$.

Чтобы учесть это обстоятельство введём ещё один коэффициент перед переменной y, допуская его равным нулю. Таким образом мы приходим к ещё одному уравнению прямой на плоскости:

$$Ax + By + C = 0, \tag{1.9}$$

которое называется *общим уравнением прямой на плоскости.* При разных значениях коэффициентов A, B и свободного члена C оно определяет теперь уже произвольную прямую на плоскости.

Если на плоскости две прямые заданы уравнениями с угловым коэффициентом (1.8):

$l_1: \quad y = k_1x + b_1$

$l_2: \quad y = k_2x + b_2$

то можно ответить на вопросы об их взаимном расположении (рис. 10):

– если $k_1 = k_2$, то прямые параллельны,

– если $k_2 = -1/k_1$, то прямые перпендкулярны.

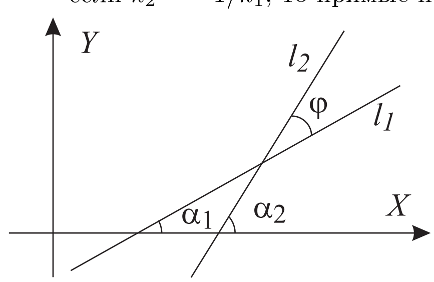


рис. 10.

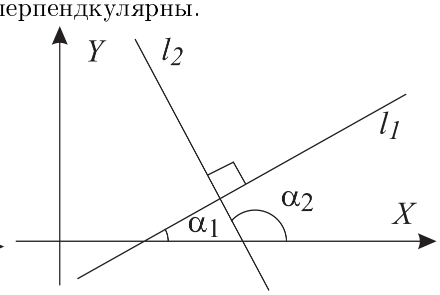


рис. 11.

Например, среди прямых

$l_1: y = 2x + 3, \quad l_2: y = 5x + 9, \quad l_3: y = -\frac{1}{2}x - 4, \quad l_4: y = 2x + 6$

l_1 и l_3 перпендикулярны, а l_1 и l_4 параллельны.

1.3 Решение неравенств на плоскости

Рассмотрим выделение точек плоскости, удовлетворяющих неравенству $y \leq 2x + 1$, как его решение.

Из смысла знака $\leq$ следует, что этому неравенству удовлетворяют точки, для которых выполняется либо знак равенства, либо знак строгого неравенства. В случае знака равенства получим уравнение прямой

$y = 2x - 4$, точки которой, как уже было сказано, принадлежат области решений этого неравенства. Найдём уравнение этой прямой по двум точкам.

В качестве первой координаты возьмём точку $x = 0$ на оси OX и получим $y = -4$. Эту точку можно отметить сразу, так как она совпадает с точкой пересечения прямой и оси OY. В принципе прямую уже можно проводить через эту точку слева направо вверх (см. рис. 12), так как коэффициент при переменной равен положительному числу. С другой стороны, когда на плоскости изображается не одна прямая, то обычно необходимо их более точное взаимное расположение, поэтому найдём вторую точку. Возьмём, например, $x = 3$, тогда $y = 2 \cdot 3 - 4 = 2$, откуда вторая точка имеет координаты $(3, 2)$ (см. рис. 13).

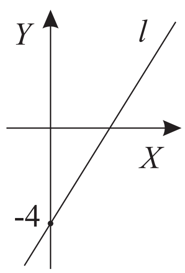


рис. 12.

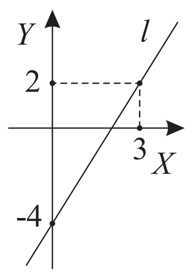


рис. 13.

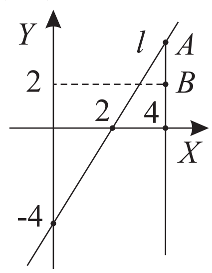


рис. 14.

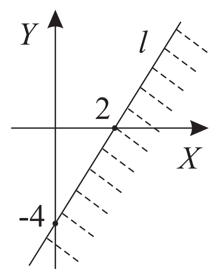


рис. 15.

Укажем область решений неравенства $y < 2x - 4$.

Возьмём $x = 4$, тогда $y = 2 \cdot 2 - 4 = 4$ и на прямой l отметим точку $A(4, 4)$. При $x = 4$ точка $B(4, 3)$, лежащая ниже точки будет удовлетворять неравенству $3 = y < 2 \cdot 4 - 4 = 4$. Из рисунка (14) понятно, что и все точки на луче $[AB)$ также будут являться решением этого неравенства. Таким образом решением неравенства $y \leq 2x + 1$ будут точки плоскости на рис. 14, расположенные ниже прямой $y = 2x + 1$. Эти точки отметим перпендикулярной к прямой $y = 2x + 1$ штриховкой (см. рис. 15)

Формализуем приведённые рассуждеия, задав линейное неравенство несколько иначе

$$Ax + By + C \geqslant 0, \quad Ax + By + C \leqslant 0$$

Множество точек плоскости, координаты которых удовлетворяют неравенству $Ax + By + C \geqslant 0$ или $Ax + By + C \leqslant 0$, образуют *множество решений* соответствующего неравенства.

Теорема 1. *Областью решений неравенства* $Ax+By+C \leqslant 0$ *является полуплоскость, ограниченная прямой* $Ax + By + C = 0$.

Таким образом, при решении задачи о нахождении области решений на плоскости линейных неравенств вида $Ax+By+C \geqslant 0$ или $Ax+By+C \leqslant 0$ необходимо:

1. Построить прямую $Ax + By + C = 0$.

2. Подставить в выражение $Ax + By + C$ координаты любой точки, не принадлежащей прямой, и определить знак этого выражения. (Обычно это точка $O = (0, 0)$). Если рассматриваемое неравенство верно, то областью решения является полуплоскость, содержащая данную точку. Если неравенство не выполняется, то областью решения будет полуплоскость, в которой выбранная точка не лежит.

Если задано строгое неравенство вида $Ax + By + C > 0$ или $Ax + By+C < 0$, то решением очевидно будет являться полуплоскость, ограниченная прямой l, но не содержащей точек прямой l.

Пример. Решить неравенство

$$2x + 3y - 6 \leqslant 0. \tag{1.10}$$

Решение. Заменим неравенство на равенство и перенесём свободный член в правую часть, оставляя в левой части переменные:

$$2x + 3y = 6.$$

Прямую построим по точкам, записывая их координаты в таблицу:

x	3	0
y	0	2

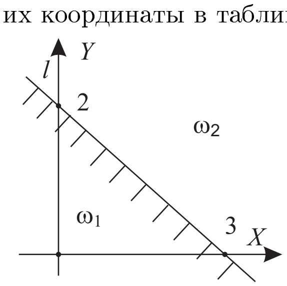


рис. 16.

Подставим координаты точки (0,0) в неравенство (1.10): $2 \cdot 0 + 3 \cdot 0 - 6 \leq 0$ – верно. Следовательно точка $(0, 0)$ принадлежит области решений неравенства, поэтому и вся область ω_1 на рис. 16 также является решением этого неравенства.

Задача решена как отмечена штриховка на рис. 16.

В случае системы линейных неравенств:

$$\begin{cases} A_1x + B_1y + C_1 \leqslant 0 \quad (l_1) \\ A_2x + B_2y + C_2 \leqslant 0 \quad (l_2) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ A_nx + B_ny + C_n \leqslant 0 \quad (l_n) \end{cases}$$

множество решений представляет собой многоугольник, так как решение каждого неравенства – полуплоскость. Полученная фигура будет обладать ещё одним важным свойством – это выпуклый многоугольник (фигура называется *выпуклой*, если для любых двух точек, принадлежащих данной фигуре, ей будет принадлежать и отрезок, соединяющий эти две точки. Например, точки A и B и отрезок $[A, B]$ входят

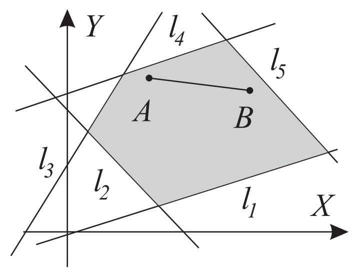


рис. 17.

в заштрихованную область, которая является решением системы неравенств, на рис. 17.).

Глава 2

Основы векторной алгебры

2.1 Векторы и действия над ними

Отрезок – это часть прямой, заключённой между двумя точками. Если эти точки перенумеровать (*упорядочить*), то про них становится известно, какая первая (начало), а какая вторая (конец). Отрезок, с отмеченными началом и концом, называется *геометрическим вектором* или просто *вектором*. К векторам будем относить и так называемый *нулевой вектор*, вектор у которого начало и конец совпадают. Вектор с началом в точке A и концом в точке B обозначается как $\overrightarrow{AB}$ (при этом буква, соответствующая началу вектора, ставится впереди). Так как обычно нет необходимости указывать начало и конец вектора, то чаще вектор будет обозначаться латинскими буквами со стрелкой – $\vec{a}$. Нулевой вектор обозначается как $\vec{0}$. Расстояние между началом и концом вектора называется *длиной вектора*, которая обозначается как $|\vec{a}|$.

Векторы называются *коллинеарными*, если они расположены на одной прямой или на параллельных прямых. Векторы называются *компланарными*, если существует плоскость, которой они параллельны. Нулевой вектор считается коллинеарным любому вектору, так как он не имеет определённого направления. Длина его, разумеется, равна нулю.

Два вектора называются *равными*, если они коллинеарны, одинаково направлены и имеют равные длины. Договоримся не различать равные векторы. Таким образом под одним и тем же вектором будем понимать множество всех равных векторов.

Для любых двух векторов $\vec{a}$ и $\vec{b}$ определим *операцию сложения век-*

торов по правилу: совместим начало $\vec{b}$ с концом $\vec{a}$. Вектор $\vec{c}$ (рис. 18), начало которого совпадает с началом вектора $\vec{a}$, а конец – с концом вектора $\vec{b}$, называется *суммой* векторов $\vec{a}$ и $\vec{b}$, и обозначается $\vec{a}+\vec{b}$.

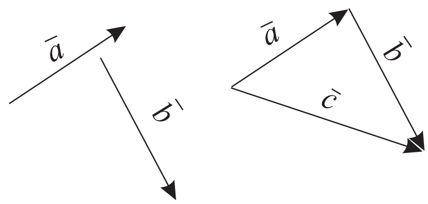


рис. 18.

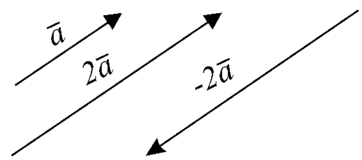


рис. 19.

Определим *операцию умножения вектора на число*: *умножение вектора $\vec{a}$ на вещественное число* λ называется вектор $\vec{b}$, обозначаемый как $\lambda\vec{a}$, удовлетворяющий следующим условиям (рис. 19):

а) вектор $\vec{b}$ коллинеарен вектору $\vec{a}$;

б) векторы $\vec{b}$ и $\vec{a}$ направлены одинаково, если $\lambda > 0$, и противоположно, если $\lambda < 0$. (Если же $\lambda = 0$, то $\vec{b} = \vec{0}$.)

в) $|\vec{b}| = |\lambda| \cdot |\vec{a}|$*;*

Операции сложения векторов и умножения вектора на число называются *линейными операциями над векторами*. Из перечисленных линейных операций над векторами следует, что вектор на прямой получается умножением некоторого числа на некоторый другой ненулевой вектор на этой прямой. Фактически это означает, что

Теорема 2. *Для ненулевых векторов $\vec{a}$ и $\vec{b}$ необходимым и достаточным условием коллинеарности является существование числа λ, что $\vec{a} = \lambda\vec{b}$.*

Если же один из векторов равен нулю, то они заведомо коллинеарны.

Всякие три вектора, лежащие в одной плоскости, либо параллельных плоскостях, называются *компланарными векторами*.

Через любую пару неколлинеарных векторов можно провести плоскость. Действительно, совмещая два вектора своими началами, получим три точки: общая точка начала и две точки концов векторов, которые однозначно определяют плоскость. Заметим, что сумма таких векторов даст вектор в той же плоскости. С другой стороны, можно показать, что всякий вектор этой плоскости, либо ей параллельной, можно представить в виде суммы двух векторов, коллинеарных двум заданным выше неколлинеарным векторам, то есть должна быть справедлива

Теорема 3. *Чтобы три ненулевых вектора $\vec{a}$, $\vec{b}$ и $\vec{c}$ были компланарны необходимо и достаточно существование пары чисел λ, μ, чтобы выполнялось равенство $\vec{c} = \lambda\vec{a} + \mu\vec{b}$.*

Применяя линейные операции, можно составлять суммы векторов, умноженных на числа: $\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n$. Выражения такого вида называются *линейными комбинациями векторов.* Числа, входящие в линейную комбинацию, называются *коэффициентами* (рис. 20, 21).

Сформулированные выше линейные операции позволяют преобразовывать выражения, составленные из линейных комбинаций, по обычным правилам алгебры: можно раскрывать скобки, приводить подобные члены, переносить некоторые члены в другую часть равенства с противоположным знаком и т. д.

Линейные комбинации векторов обладают следующими очевидными геометрическими свойствами:

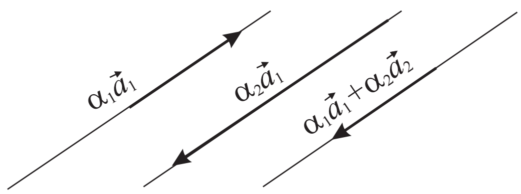


рис. 20.

– *если векторы $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ коллинеарны, то любая их линейная комбинация им коллинеарна.*

В самом деле, если все векторы попарно коллинеарны, то они расположены на одной прямой. Тогда их сумма и умножение на число даёт вектор на той же прямой (рис. 20); – *если векторы $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ компланарны, то любая их линейная комбинация компланарна.*

Это следует из того, что, если вектора расположены в одной плоскости, то как следует из рис. 21 и их линейная комбинация также принадлежит той же плоскости.

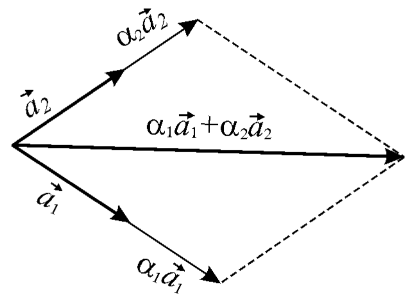


рис. 21.

Линейная комбинация векторов $\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \ldots + \alpha_n\vec{a}_n$ образует вектор. Пусть это будет нулевой вектор. Если при этом все коэффициенты нулевые, то такая линейная комбинация неинтересна (линейная комбинация векторов с нулевыми коффициентаи очевидно равна нулю), так как представляет собой естественное равенство. Значительно интереснее, если не все коэффициенты нулевой комбинации равны нулю. Про векторы, удовлетворяющие такой линейной комбинации говорят как о *линейно зависимых*. Например, если $\vec{a} + 2\vec{b} - 3\vec{c} = \vec{0}$, то система векторов $\{\vec{a}, \vec{b}, \vec{c}\}$ – линейно зависима. Заметим, что тогда $\vec{a} = 3\vec{c} - 2\vec{b}$, то есть вектор $\vec{c}$ выражается через оставшиеся вектора $\{\vec{b}, \vec{c}\}$ исходной системы векторов $\{\vec{a}, \vec{b}, \vec{c}\}$. Нетрудно выразить и другие векторы через оставшиеся, например, $\vec{b} = -0.5\vec{a} + 1.5\vec{c}$. Поэтому для ненулевых векторов можно дать другой определение линейной зависимости: система ненулевых вектораов линейно зависима, если один из векторов этой системы может быть выражен через другие векторы этой системы. Если же в системе есть хотя бы один нулевой вектор, то такая система автоматически становится линейно зависимой. Например, система $\{\vec{a}, \vec{b}, \vec{0}\}$ зависима, так как из можно составить нулевую линейную комбинацию, при этом не все коэффициенты равны нулю $0\vec{a} + 0\vec{b} + 1 \cdot \vec{0} = 0$.

Векторы, единственная нулевая линейная комбинация которых, содержит только нулевые коэффициенты называются *линейно независимыми*. Можно сказать и по-другому, *если векторы не являются линейно зависимыми, то они лиейно независимы*. (На рис. 21 векторы $\vec{a}_1$ и $\vec{a}_2$ линейно независимы.)

2.2 Базис и система координат

Рассмотрим векторы на прямой, на плоскости и в пространстве.

Базисом на прямой называется любой ненулевой вектор на этой прямой.

Базисом на плоскости называются два неколлинеарных вектора на этой плоскости.

Базисом в пространстве называются три некомпланарных вектора.

Отметим, что векторы базиса на плоскости не могут быть нулевыми, иначе они были бы коллинеарны. Точно так же никакие два из векторов базиса в пространстве не коллинеарны – в противном случае все три вектора были бы компланарны.

Если вектор представлен как линейная комбинация векторов, например, $\vec{a} = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2$, то говорят, что он *разложен* по этим векторам.

Если $\vec{e}_1, \vec{e}_2, \vec{e}_3$ – базис в пространстве и $\vec{a} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \alpha_3\vec{e}_3$, то числа $\alpha_1, \alpha_2, \alpha_3$ называются *координатами* (или компонентами) вектора

$\vec{a}$ в данном базисе. Аналогично определяются компоненты вектора на плоскости и на прямой. Координаты вектора пишут в скобках после буквенного обозначения вектора. Например, $\vec{a} = (2, 0, 1)$ означает, что координаты вектора $\vec{a}$ в определенном базисе равны 2, 0 и 1.

Основное значение базиса заключается в следующей теореме:

Теорема 4. *Каждый вектор может быть разложен по базису единственным образом.*

Данная теорема устанавливает важное соответствие между геометрическими векторами и их алгебраическим представлением как упорядоченного набора чисел.

Пусть вектор разложен по базису. В этом случае:

– при умножении вектора на число все его компоненты умножаются на это число,

– при сложении векторов складываются их соответствующие компоненты.

То есть для любого $\vec{a} = \alpha_1\vec{e}_1 + \alpha_2\vec{e}_2 + \alpha_3\vec{e}_3 = (\alpha_1, \alpha_2, \alpha_3)$ и любого числа λ получим $\lambda\vec{a} = (\lambda\alpha_1, \lambda\alpha_2, \lambda\alpha_3)$. Если взять ещё один вектор $\vec{b} = (\beta_1, \beta_2, \beta_3)$, то $\vec{a}+\vec{b} = (\alpha_1, \alpha_2, \alpha_3)+(\beta_1, \beta_2, \beta_3) = (\alpha_1+\beta_1, \alpha_2+\beta_2, \alpha_3+\beta_3)$.

Например, для $\vec{a} = (3, -2, 4), \vec{b} = (2, 1, -3), \lambda = 5$ получим $\lambda\vec{a} = 5 \cdot (3, -2, 4) = (15, -10, 20)$ и $\vec{a} + \vec{b} = (3, -2, 4) + (2, 1, -3) = (5, -1, 1)$.

Фиксируем в пространстве точку O. Выберем в этом пространстве некоторый базис. Ранее введённую деартову систему координат можно определить на языке векторной алгебры: *декартова системой координат* называется совокупность точки и базиса. Точка называется *началом координат*; координатные прямые, проходящие через начало координат в направлении базисных векторов, называются *осями координат*. Первая – осью абсцисс, вторая – осью ординат, третья – осью аппликат. Плоскости, проходящие через оси координат, называют координатными плоскостями.

Обычно вместо общих декартовых систем координат будем пользоваться только прямоугольными декартовыми системами координат с ортонормированным базисом. В этом случае базисные векторы, называемые также *ортами*, имеют специальное обозначение: $\vec{i}, \vec{j}, \vec{k}$. Таким образом, для всякого вектора в пространстве справедливо разложение по ортонормированному базису:

$$\vec{a} = x\vec{i} + y\vec{j} + z\vec{k} = (x, y, z) \tag{2.1}$$

которое единственно по теореме 4, то есть набор чисел x, y, z единственен для каждого вектора $\vec{a}$.

Пусть в пространстве, в котором задана декартова система координат, имеется две точки A и B. Они определяют векторы $\overrightarrow{OA}$ и $\overrightarrow{OB}$. Раскладывая вектора $\overrightarrow{OA}$ и $\overrightarrow{OB}$ по базису декартовой системы координат, получим координаты векторов $\overrightarrow{OA}$ и $\overrightarrow{OB}$. Эти координаты векторов в дальнейшем будем рассматривать как координаты соответствующих точек.

Например, имеется точка $A = (x_A, y_A)$, тогда координаты этой точки определяют вектор $\overrightarrow{OA} = (x_A, y_A)$. Здесь не должно возникать путаницы, так как координатную запись всегда можно понимать и как вектор и как координаты соответствующей точки.

По точкам A и B можно построить вектор $\overrightarrow{AB}$, для которого справедлива следующая

Теорема 5. *Чтобы по точкам A и B найти координаты вектора $\overrightarrow{AB}$ в декартовой системе координат нужно из координат его конца (точки B) вычесть координаты его начала (точки A).*

Из теоремы следует, что если точки A и B на плоскости имеют координаты (x_A, y_A) и (x_B, y_B), то $\overrightarrow{OA} = (x_A, y_A)$, $\overrightarrow{OB} = (x_B, y_B)$ и $\overrightarrow{AB} = (x_B - x_A, y_B - y_A)$. Например, для $A(3, 5)$ и $B(8, -2)$ получим $\overrightarrow{AB} = (5, 7)$.

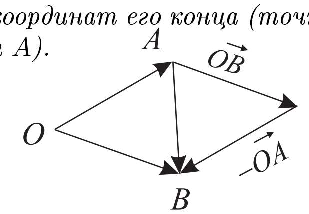


рис. 22.

2.3 Скалярное произведение векторов

Будем понимать под углом между векторами угол между этими векторами, с совмещённым началом, не превосходящий 180^o.

Скалярным произведением двух векторов $\vec{a}$ и $\vec{b}$ называется число, обозначаемое $(\vec{a}, \vec{b}) = \vec{a} \cdot \vec{b}$, равное произведению длин этих векторов на косинус угла между ними. Если хоть один из векторов нулевой, то скалярное произведение оказывается равным нулю.

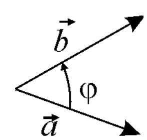


рис. 23.

Таким образом можно записать

$$(\vec{a}, \vec{b}) = |\vec{a}| \cdot |\vec{b}| \cos \varphi, \tag{2.2}$$

где φ – угол между векторами $\vec{a}$ и $\vec{b}$.

Если скалярное произведение равно нулю, то векторы называются *ортогональными*.

Из определения скалярного произведения векторов следует, что

$$|\vec{a}| = \sqrt{(\vec{a}, \vec{a})}. \tag{2.3}$$

Пусть два вектора $\vec{a}$ и $\vec{b}$ разложены по ортогональному базису $\vec{i}, \vec{j}, \vec{k}$ пространства:

$$\vec{a} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k} = (a_1, a_2, a_3),$$

$$\vec{b} = b_1\vec{i} + b_2\vec{j} + b_3\vec{k} = (b_1, b_2, b_3).$$

тогда

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3. \tag{2.4}$$

Например, для $\vec{a} = (2, -5, 3), \vec{b} = (3, -1, -2)$ по формуле (2.4) находим

$$\vec{a} \cdot \vec{b} = 6 + 5 - 6 = 5.$$

Формула вычисления длины вектора в ортонормированном базисе также записывается значительно проще:

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}. \tag{2.5}$$

В частности,

$$|a| = \sqrt{4 + 25 + 9} = \sqrt{38}, \quad |b| = \sqrt{14}.$$

2.4 Векторный анализ на плоскости и в пространстве

Можно показать, что если записать уравнение прямой на плоскости как общее уравнение $Ax + By + C = 0$, то коэффициенты A и B имеют смысл координат вектора, перпендикулярного этой прямой. Этот вектор называется ***вектором нормали*** и обозначает как $\vec{N}$ как на рис. 24.

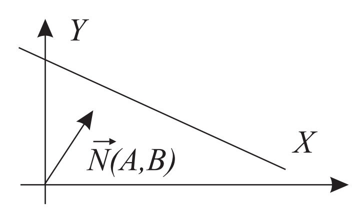


рис. 24.

Например, для прямой $2x+3y=6$ нормалью будет вектор $\vec{N}=(2,3)$. Уравнения прямой с одиниковыми векторами нормали, но различными свободными членами отвечают параллельным прямым.

Перенесём свободый член общего уравнения прямой в правую часть:

$$Ax+By=C$$

Нетрудно доказать, что *вектор нормали будет показывать направление параллельного перемещения прямой, при котором постоянная C увеличивается*, что видно из рис. 25.

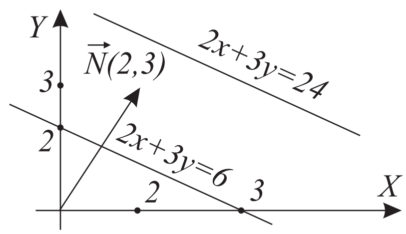


рис. 25.

Уравнение плоскости ω в пространстве имеет вид (см. рис. 26.)

$$\omega : Ax+By+Cz+D=0 \tag{2.6}$$

и называется *общим уравнением плоскости в пространстве.* Коэффициенты A,B,C являются координатами вектора, ортогонального плоскости $\vec{N}=(A,B,C)$ (нормаль к плоскости).

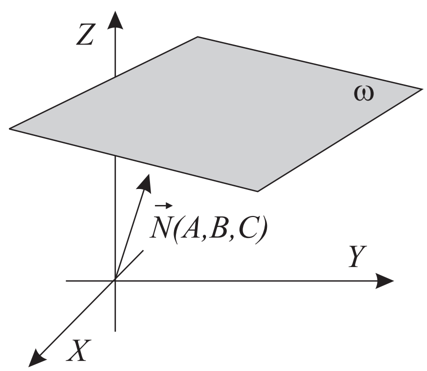


рис. 26.

Например, вектор $\vec{N}=(3,-2,4)$ ортогонален как плоскости $3x-2y+4z-5=0$, так и плоскости $3x-2y+4z+8=0$.

Глава 3

Линейные пространства

3.1 Введение

Рассмотрим фирму, состоящую из двух предприятий. Финансовый результат её деятельности за год представим в виде таблицы:

предприятие	прибыль фирмы по кварталам			
	1	2	3	4
I	18	22	16	14
II	12	18	14	16

Если нас интересуют данные за конкретный квартал, например, за второй, то, обозначая прибыль фирмы (прибыль фирмы в целом разбита на прибыль по отдельным предприятиям) u_2, их можно представить в виде записи:

u_2 =(прибыль I-го пр., прибыль II-го пр.)=(22,18).

Аналогично $u_3 = (16, 14)$ и т. д. (Числа в скобках, отделённые запятыми назовём компонентами.) Прибыль фирмы, полученная во втором полугодии и записанная как v_2, равна: $(30, 30)$. Эта же запись может быть получена как сумма записей $u_3 + u_4$, если договориться при сложении записей складывать соответствующие компоненты: $v_2 = u_3 + u_4$.

Пусть в течении всего второго квартала прибыль фирмы распределена равномерно. Тогда в первой половине второго квартала прибыль фирмы, разбитая по предприятиям, как это следует из таблицы, составила $w = (11, 9)$. Это же число может быть получено умножением u_2 на $1/2$, если договориться при умножении записей на число умножать отдельные компоненты записи, то есть: $w = \frac{1}{2} \cdot u_2 = (\frac{1}{2} \cdot 22, \frac{1}{2} \cdot 18) = (11, 9)$.

Из приведённых правил действия с записями прибыли видим, что они очень напоминают правила действий с геометрическими векторами, представленными в координатной форме. Важным является и то, что в результате арифметических операций над записями были получены записи того же смысла: прибыль + прибыль = прибыль, число × прибыль = прибыль. Данные аналогии наводят на мысль использовать аппарат векторной алгебры, разработанный для геометрических векторов, в экономических задачах. Оказывается, что такой подход возможен, при этом операции введённые для геометрических векторов будут иметь смысл и для экономических записей, которые также можно называть векторами (чтобы различать геометрические вектора и экономические записи над такими векторами стрелки ставить не будем). В случае геометрических векторов имелось ещё одно важное ограничение: не очень понятно, что из себя представляет геометрический вектор с четырьмя компонентами, но экономическая запись с любым числом компонент вполне осмысленна. Таким образом необходимо обобщить аппарат геометрических векторов на произвольное число компонент. Математический аппарат, полученный обобщением аппарата геометрических векторов – это аппарат линейной алгебры. В настоящее время линейная алгебра – это важнейший раздел математики, для которого геометрические вектора представляют собой очень частную, но наглядную конструкцию некоторых понятий современной линейной алгебры.

3.2 n-мерное линейное пространство

Упорядоченное множество элементов вида $x = (x_1, x_2, \ldots, x_n)$, где все $x_i \in R$, над которыми введены операции сложения и умножения на число по правилу:

1) $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$,

2) $x \cdot \lambda = \lambda \cdot x = \lambda \cdot (x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$,

называется *n-мерным вещественным линейным координатным пространством или просто – пространством* R^n.

Числа $x_i, i = 1, \ldots, n$ называются *координатами вектора* x. В дальнейшем под *линейным пространством* будем понимать именно пространство R^n. Так как не всегда необходимо знать число координат пространства, то вместо R^n будем писать L.

Вектор, все координаты которого нулевые, называется нулевым вектором $0 = (0, 0, \ldots, 0)$.

Пусть $x = (x_1, x_2, \ldots, x_n)$ – некоторый n-мерный вектор. Тогда вектор $\tilde{x} = (-x_1, -x_2, \ldots, -x_n)$ называется противоположным вектором вектору x. Понятно, что $x + \tilde{x} = 0$.

Таким образом линейное пространство – это совокупность элементов $x, y, \ldots$ вида $x = (x_1, x_2, \ldots, x_n), y = (y_1, y_2, \ldots, y_n), \ldots$, называемых векторами, для которых определены операция сложения по правилу 1) и умножения на число по правилу 2).

Любое конечное множество векторов назовём *системой векторов* (если это понятно по контексту, то – просто *система*) и будем обозначать латинской буквой S. Если известно, что система S содержит ровно n векторов, то систему записывают как S_n. Например, $S_n = \{a_1, a_2, \ldots, a_n\}$ – система векторов $a_1, a_2, \ldots, a_n$. Заметим, что это неупорядоченный набор векторов. Среди различных систем векторов будем выделять большие системы и меньшие. Так, будем считать, что *система векторов S_n больше системы векторов S_k*, если $n > k$. Любую часть системы векторов S_n назовём её *подсистемой*.

Линейной комбинацией системы векторов или просто векторов $a_1, a_2, \ldots, a_n$ называется сумма вида

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n, \tag{3.1}$$

где $\lambda_1, \lambda_2, \ldots, \lambda_n$ – произвольные вещественные числа. Линейная комбинация называется *тривиальной*, если все $\lambda_1, \lambda_2, \ldots, \lambda_n$ равны нулю. В противном случае линейная комбинация называется *нетривиальной*. Например, линейная комбинация вида $5a_1 + 3a_2 - 0a_3$ – нетривиальная, а линейная комбинация $0a_1 + 0a_2 - 0a_3$ – тривиальная.

Система векторов называется *линейно независимой*, если не существует нетривиальной линейной комбинации векторов этой системы, равной нулевому элементу, в противном случае говорят, что система *линейно зависима*. В частности, три компланарных геометрических вектора $\vec{a}, \vec{b}, \vec{c}$ будут линейно зависимы, так как в силу компланарности один из них должен быть представим в виде линейной комбинации двух других, но тогда очевидно имеется нетривиальная линейная комбинация этих векторов. Например, если $\vec{c} = \alpha\vec{a} + \beta\vec{b}$, то имеется нетривиальная, равная нулю, линейная комбинация $\alpha\vec{a} + \beta\vec{b} - 1 \cdot \vec{c} = 0$. В том случае, когда три геометрических вектора некомпланарные – эти векторы линейно независимы, так как в противном случае один из них может быть выражен в виде линейной комбинации других векторов, откуда следовало бы, как было только что показано, что они компланарны.

Следовательно введённое ранее понятие линейной зависимости и независимости для геометрических векторов является частным и наглядным случаем понятия линейной зависимости и линейной независимости векторов линейного пространства.

Максимальное число линейно независимых векторов системы называется ***рангом этой системы***. Пусть, например, в системе $S_n = \{a_1, a_2, \ldots, a_n\}$ только $k, k \leqslant n$, векторов линейно независимы, а любое большее число векторов системы линейно зависимы. Тогда пишут $\mathrm{rang}\ S_n = k$.

Максимальная линейно-независимая система векторов $a_1, a_2, \ldots, a_n$ пространства называется ***базисом*** этого пространства. Число векторов в базисе пространства L называется ***размерностью пространства*** и обозначается как $\dim L$. Если пространство состоит только из нулевого вектора, то его размерность по определению равна нулю. Так, если базис пространства L состоит из n векторов, то пишут $\dim L = n$.

Вместо всего пространства R^n часто полезно рассматривать его часть. Например, дана совокупность векторов вида $(x_1, x_2, 0)$. У всех таких векторов третья координата равна нулю. С другой стороны, множество векторов (x_1, x_2) образует пространство R^2. Векторы вида $(x_1, x_2, 0)$ обладают важной особенностью: их сумма и умножение на число даёт опять вектор такого же вида $(x_1', x_2', 0)$, поэтому про них говорят, что они образуют подпространство. Более строго: ***подпространством*** L' пространства L называется множество M векторов пространства L, если любая сумма векторов этого множества и умножение вектора этого множества на число также является элементом этого множества.

В случае геометрических векторов в трёхмерном пространстве, понятно, что базисом являются любые три некомпланарных вектора. В том случае, если эти векторы заданы на плоскости базисом являются любые два неколлинеарных вектора. При этом набор всех векторов плоскости трёхмерного пространства образует двумерное подпространство трёхмерного геометрического пространства.

Базис в геометрическом пространстве является частным случаем базиса произвольного линейного пространства. Поэтому вполне естественно, что в линейном пространстве имеется аналог Теоремы 3:

Теорема 6. *Любой вектор n-мерного пространства может быть представлен в виде линейной комбинации векторов базиса этого пространства и это представление единственно.*

В соответствии с Теоремой 6 любой вектор x линейного пространства L может быть единственным образом представлен в виде

$$x = x_1 a_1 + x_2 a_2 + \ldots + x_n a_n, \tag{3.2}$$

где $a_1, a_2, \ldots, a_n$ – базис пространства L. Такое разложение называется *разложением вектора x по базису* $a_1, a_2, \ldots, a_n$, а числа $x_1, x_2, \ldots x_n$ – *координатами* вектора x относительно базиса $a_1, a_2, \ldots, a_n$.

Линейная комбинация системы векторов $S = \{a_1, a_2, \ldots, a_k\}$

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_k a_k, \tag{3.3}$$

где $\lambda_1, \lambda_2, \ldots, \lambda_k$ – произвольные вещественные числа, образует некоторую совокупность векторов, обозначаемую как $L(S)$, называется *линейной оболочкой, образованной векторами* $a_1, a_2, \ldots, a_k$. Линейная оболочка также является пространством, но в общем случае подпространством исходного пространства L. Если это базисные векторы, то из Теоремы 6 следует важная

Теорема 7. *Линейное пространство является линейной оболочкой своего базиса.*

Ранее была показана взаимная однозначность вектора как геометрического объекта геометрии Евклида в трёхмерном пространстве и упорядоченной тройки чисел-координат. Казалось бы лучше и не говорить об этой тройке чисел, а держать в голове его физический образ в виде стрелки. Но вот мы имеем вектор-строку из n чисел. Как её представить в пространстве? Ведь нужно пространство n измерений. Видимо никак. С другой стороны, рассматривая n-мерный вектор-столбец или n-мерную вектор строку, можно пользоваться физическим образом трёх измерений так, как будто это пространство большего числа измерений.

3.3 Скалярное произведение n-мерных векторов

Угол между векторами можно определить для двумерного или трёхмерного пространств. В случае пространства большей размерности понятие угла в привычном нам смысле евклидовой геометрии теряется. Возникает вопрос – имеет ли смысл угол в пространстве большем, чем три измерения? На самом деле должен иметь. Действительно, через любые два вектора можно провести двумерную плоскость, а в двумерном пространстве хорошо понятно, что такое угол.

Скалярным произведением $(x, y) = x \cdot y$ двух n-мерных векторов

$x = (x_1, x_2 \dots x_n)$ и $y = (y_1, y_2 \dots y_n)$ называется число

$$x \cdot y = x_1 \cdot y_1 + x_2 \cdot y_2 + \ldots + x_n \cdot y_n. \tag{3.4}$$

Длиной n-мерного вектора $a = (a_1, a_2, \ldots, a_n)$, обозначаемой как $||a||$ (норма вектора a) или $|a| = ||a||$, называется число

$$||a|| = \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}. \tag{3.5}$$

Косинусом угла между двумя ненулевыми n-мерными векторами a и b называется число

$$\cos\varphi = \frac{a \cdot b}{||a|| \cdot ||b||}. \tag{3.6}$$

Пример 1. Найти скалярное произведение, длины векторов $a = (2, 5, -4, 7, 8)$ и $b = (3, -8, 5, -9, 1)$ и угол между ними.

$$a \cdot b = (2, 5, -4, 7, 8) \cdot (3, -8, 5, -9, 1) =$$

$$= 2 \cdot 3 + 5 \cdot (-8) - 4 \cdot 5 + 7 \cdot (-9) + 8 \cdot 1 = -109,$$

$$||a|| = \sqrt{2^2 + 5^2 + 4^2 + 7^2 + 8^2} = \sqrt{158}, \quad ||b|| = \sqrt{180},$$

$$\cos\varphi = \frac{-109}{\sqrt{158 \cdot 180}} = -0.646,$$

то $\varphi = \arccos(-0.646) = 2.27$ радиан или 130.241 градусов.

Пример 2. Найти угол между $c = (2, 3, -5, 4)$ и $d = (-3, 2, 4, 5)$.

$$c \cdot d = (2, 3, -5, 4) \cdot (-3, 2, 4, 5) = 2 \cdot (-3) + 3 \cdot 2 - 5 \cdot 4 + 4 \cdot 5 = 0.$$

Следовательно эти векторы ортогональны.

Глава 4

Матрицы и определители

4.1 Матрицы и действия над ними

Пусть имеется таблица элементов, содержащая m строк и n столбцов (табл. 1):

Таблица 1.

a_{11}	a_{12}	$\dots$	a_{1n}
a_{21}	a_{22}	$\dots$	a_{2n}
$\dots$	$\dots$	$\dots$	$\dots$
a_{m1}	a_{m2}	$\dots$	a_{mn}

Матрицей размера $m \times n$ называется прямоугольная таблица чисел, обозначаемая как $A = (a_{ij})$, где i – номер строки ($i = 1, 2, \dots, m$) и j – номер столбца ($j = 1, 2, \dots, n$). Над матрицами определены следующие операции:

1) умножение на число, когда на число умножаются все элементы матрицы: $k(a_{ij}) = (ka_{ij})$;

2) если матрицы имеют одинаковое число строк и столбцов, то при сложении матриц складываются их соответствующие элементы: $(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$;

3) если число столбцов матрицы (a_{ij}) *равно числу строк матрицы* (b_{ij})*, то при умножении матрицы* A *на матрицу* B *получается матрица* (c_{ij})*, элементы которой* c_{ij} *равны* $a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$. (Такой способ умножения называется ещё умножением «строки на столбец».)

Две матрицы называются *равными*, если их соответствующие элементы равны.

Матрицы записывают в следующем виде:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{4.1}$$

Поясним операции над матрицами, указанные в определении, на примерах. Пусть

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & -2 & 0 \end{pmatrix}, B = \begin{pmatrix} -1 & 6 & -4 \\ 4 & -2 & -8 \end{pmatrix}, C = \begin{pmatrix} 2 & 1 & 0 \\ -3 & -1 & 5 \\ 1 & 2 & -2 \end{pmatrix}, \text{ то}$$

$$A + B = \begin{pmatrix} 1 & 2 & 1 \\ 3 & -2 & 0 \end{pmatrix} + \begin{pmatrix} -1 & 6 & -4 \\ 4 & -2 & -8 \end{pmatrix} = \begin{pmatrix} 0 & 8 & -3 \\ 7 & -4 & -8 \end{pmatrix}$$

$$3 \cdot B = 3 \cdot \begin{pmatrix} -1 & 6 & -4 \\ 4 & -2 & -8 \end{pmatrix} = \begin{pmatrix} -3 & 18 & -12 \\ 12 & -6 & -24 \end{pmatrix}$$

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

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

$$= \begin{pmatrix} -3 & 1 & 8 \\ 12 & 5 & -10 \end{pmatrix}.$$

Заметим, что для матриц важен порядок умножения и $A \cdot B \neq B \cdot A$. Если же матрицы A и B таковы, что существуют оба произведения ($A \cdot B$ и $B \cdot A$), причем $A \cdot B = B \cdot A$, то матрицы A и B называются перестановочными или коммутирующими.

Нулевой матрицей называется матрица, состоящая из одних нулей и обозначается как O. Матрицей, *противоположной* матрице A, называется матрица $-1 \cdot A = -A$.

Существует следующая классификация матриц:

– матрица, у которой все элементы выше (ниже) главной диагонали равны нулю, называется *нижнетреугольной* (*верхнетреугольной*) или просто *треугольной*;

– матрица, у которой равное число строк и столбцов называют *квадратной*. Квадратная матрица размера $n \times n$ называется также квадратной матрицей порядка n;

– квадратная матрица, у которой все элементы, кроме главной диагонали равны нулю, называется *диагональной*.
– диагональная матрица, у которой все элементы главной диагонали равны между собой, называется *скалярной* матрицей;
– скалярная матрица, у которой все элементы главной диагонали равны единице, называется *единичной* матрицей и обозначается как E;
– матрица $1 \times n$ называется *вектор-строкой*, а матрица $n \times 1$ – *вектор-столбцом*.

Элементы матрицы, имеющие одинаковые индексы, образуют *главную диагональ* матрицы. Для матрицы A в (4.1) – это элементы a_{11}, $a_{22}, \ldots, a_{nn} (n \leqslant m)$.

В отличии от умножения деление для матриц не определено и его заменяют умножением на обратную матрицу. Матрицу B называют *обратной* матрице A, если $A \cdot B = B \cdot A = E$. Обратная матрица для матрицы A обозначается как A^{-1}. Из определения следует, что обратная матрица существует только у квадратной. Но даже у квадратной матрицы обратная существует не всегда. Если её нет, то про такую матрицу говорят, что она *выраждена* или *особенная*, если обратная матрица существует, то она называется *невырожденной* или *неособенной*.

Например матрица A_1 обратной не имеет, а матрицы A_2, A_3 имеют обратную:

$$A_1 = \begin{pmatrix} 5 & 3 \\ 10 & 6 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix}. \tag{4.2}$$

Покажем, что $A_2^{-1} = A_3$. Для этого проверим, что произведение $A_2 \cdot A_3$ равно единичной матрице:

$$\begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & -3 \\ -3 & 5 \end{pmatrix} = \begin{pmatrix} 5 \cdot 2 - 3 \cdot 3 & 5 \cdot (-3) + 3 \cdot 5 \\ 3 \cdot 2 - 2 \cdot 3 & 3 \cdot (-3) + 2 \cdot 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.3}$$

– верно.

Матрицу A размера $(m \times n)$ можно записать либо в виде m вектор-строк A_i, либо в виде n вектор-столбцов A^j:

$$A = \begin{pmatrix} A_1 \\ \ldots \\ A_m \end{pmatrix} = (\, A^1 A^2 \ldots A^n \,). \tag{4.4}$$

Определение произведения матриц $A = \begin{pmatrix} A_1 \\ \ldots \\ A_m \end{pmatrix}$ и $B = (\, B^1 \ldots B^n \,)$

можно представить как произведение вектор-строк матрицы A на вектор-столбцы матрицы B:

$$AB = \begin{pmatrix} A_1B^1 & A_1B^2 & \dots & A_1B^m \\ A_2B^1 & A_2B^2 & \dots & A_2B^m \\ \dots & \dots & \dots & \\ A_nB^1 & A_nB^2 & \dots & A_nB^m \end{pmatrix},$$

где под выражением A_iB^k понимаетя скалярное произведение вида

$$A_iB^j = (a_1, \dots, a_k) \cdot (b^1, \dots, b^k) = a_1b^1 + \dots + a_kb^k$$

(оно определено только тогда, когда число элементов вектор-строки A_i равно числу элементов вектор-столбца B^j, поэтому произведение матриц определено не для всех матриц, а только для тех, у которых число столбцов (матрицы, стоящей слева) равно числу строк матрицы, стоящей справа).

Пусть $A = (a_{ij})$ – некоторая матрица. Матрица, в которой строки заменены на столбцы называется *транспонированной матрицей* и обозначается как $A' = (a_{ji})$. Например,

$$A = \begin{pmatrix} 4 & 6 & -7 & 5 \\ 3 & 1 & 9 & 0 \\ 6 & 2 & -5 & 7 \end{pmatrix}, \qquad A' = \begin{pmatrix} 4 & 3 & 6 \\ 6 & 1 & 2 \\ -7 & 9 & -5 \\ 5 & 0 & 7 \end{pmatrix}.$$

По существу вводится некоторая операция по изменению исходной матрицы. Эта операция отлична от изучаемых ранее и называется *транспонированием.*

Для векторов линейных пространств вводились понятия линейной комбинации векторов, их линейной зависимости и независимости. Так как вектор-строки и вектор-столбцы матриц эквивалентны векторам, то понятия линейной комбинации, линейной зависимости и линейной независимости автоматически распространяются на новые объекты: вектор-строки и вектор-столбцы матрицы.

Максимальное число линейно независимых вектор-строк матрицы называется *рангом* матрицы, что пишут как $\operatorname{rang} A$. Заметим, что если все элементы матрицы равны нулю, то ранг такой матрицы, как следует из определения, равен нулю.

4.2 Определитель матрицы

Произвольное число назовём *определителем первого порядка.*

Рассмотрим четыре числа $a_{11}, a_{12}, a_{21}, a_{22}$. Из них можно составить квадратную матрицу:

$$A = (a_{ik}) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \tag{4.5}$$

Сопоставим этой матрице A число $|A| = \Delta(A)$ по правилу:

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}. \tag{4.6}$$

Это число называется *определителем второго порядка.* Так как это число можно рассматривать как сопоставление числа квадратной матрице, то его называют *определителем квадратной матрицы A.*

Эти же четыре числа, записанные в виде определителя второго порядка можно рассматривать как два двумерных вектора на плоскости. Можно показать, что модуль определителя этих векторов будет равен площади параллелограмма, построенного на этих векторах.

Для произвольной квадратной матрицы n-го порядка можно также сопоставить число по определённому правилу, но более сложному. Но всякий раз модуль определителя будет равен объёму n-мерного параллелограмма, построенного на этих векторах.

Для представления о том, что такое определитель часто достаточно понимание именно того, что его модуль равен объёму. В частности, отсюда следует, что определитель трёх вектор пространства, принадлежащих одной плоскости равен нулю, так как соответствующий лбъём равен нулю. Оказывается верно и обратное: если определитель трёх векторов трёхмерного пространства равен нулю, то эти векторы находятся в одной плоскости, то есть они линено зависимы. Таким образом определитель играет роль критерия являются ли векторы линейно зависимы или нет. В этом смысле нам не столь важен знак определителя, так достаточно знания равен ли или нет он нулю. Знак определителя обусловлен тем, в каком порядке записаны вектор-столбцы определителя, и если он не ноль, вектор-столбцы будут образовывать либо правый, либо левый базис в пространстве.

Интересным обстоятельством является то, что

Теорема 8. *Определитель при транспонировании не меняется.*

Сформулируем в заключении теорему, выражающую основное свойство определителя.

Теорема 9. *Для того, чтобы определитель равнялся нулю, необходимо и достаточно, чтобы его строки или столбцы были линейно зависимы.*

4.3 Обратная матрица

Всякой квадратной матрице можно сопоставить определитель. Но не у всякой квадратной матрицы имеется обратная матрица. На примере, действительных чисел для числа 4 обратным числом будет число $4^{-1} = 1/4$, так их произведение равно единице. Но у числа 0 нет обратного. Для матриц можно показать, что обратная имеется только в том случае, если её определитель отличен от нуля. Таким образом определитель квадратной матрицы является критерием того, есть ли у неё обратная. Например, матрица второго порядка A вида

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

не имеет обратной, так как представляет собой два коллинеарных вектора $\vec{a} = (2, 3)$ и $\vec{b} = (4, 6)$, что $\vec{b} = 2\vec{a}$. Тем самым её определитель равен нулю, как нулевая площадь на двух коллинеарных векторах.

Для произвольной невырожденной квадратной матрицы (то есть той, у которой ненулевой определитель) можно указать формулу ей обратной:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{4.7}$$

Действительно, пусть

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

тогда

$$A^{-1} = \frac{1}{22} \begin{pmatrix} 7 & 3 \\ 2 & 4 \end{pmatrix}$$

в чём нетрудно убедиться непосредственной проверкой.

В общем случае справедлива

Теорема 10. *Всякая квадратная матрица, определитель которой отличен от нуля имеет единственную обратную.*

Так же как и числа матрицы можно умножать, но здесь важен порядок умножения. Например, $A \cdot X = B$, где A и X – матрицы, для которых определено умножение. Если нам известны матрицы A и B, и матрица A имеет обратную, то можно показать, что $X = A^{-1}B$. При этом умножение $B \cdot A^{-1}$ в общем случае даст другую матрицу, не являющуся решением матричного уравнения $AX = B$.

Например,
$$\begin{pmatrix} 5 & -7 \\ -2 & 3 \end{pmatrix} \cdot X = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix} \tag{4.8}$$

Тогда из формулы (4.7) получим
$$A^{-1} = \begin{pmatrix} 3 & 7 \\ 2 & 5 \end{pmatrix}$$
откуда
$$X = A^{-1}B = \begin{pmatrix} 3 & 7 \\ 2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 23 & 31 \\ 16 & 22 \end{pmatrix}$$

Заметим, что умножая матрицы в обратном порядке, получили бы другой ответ:
$$Y = BA^{-1} = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 3 & 7 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 11 & 26 \\ 14 & 34 \end{pmatrix}$$

При этом матрица X является решением уравнения (4.8), а матрица Y – нет:
$$A \cdot X = \begin{pmatrix} 5 & -7 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 23 & 31 \\ 16 & 22 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix} = B$$
$$A \cdot Y = \begin{pmatrix} 5 & -7 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 11 & 26 \\ 14 & 34 \end{pmatrix} = \begin{pmatrix} -43 & -108 \\ 20 & 50 \end{pmatrix} \neq B$$

Глава 5

Система линейных уравнений и её решения

Набор чисел, которые объединены знаками «$+, -, :, \cdot$» назовём *числовым выражением* или просто выражением. Внутри выражения допустимы знаки скобки «(» или «)», определяющие порядок действий «$+, -, :$ $, \cdot$» над числами. Например, $2 + 3 - 5 \cdot 7$ – числовое выражение. Из двух выражений можно составить *равенство*. Например, $2 \cdot 6 = 12$ или $2 \cdot 6 = 2 \cdot 5 + 2$, которое либо верно либо неверно. Равенство считается верным, если в его левой и правой частях после упрощений стоят одинаковые числа, иначе равенства считаются неверными. (Например, равенство $2 \cdot 6 = 2 \cdot 5 - 2$ является неверным.)

Числовое выражение может содержать буквы и числа. Считается, что буквы обозначают некоторые числа. Например, в выражении $3x - 3y + 4$ под буквами x и y понимаются числа. Аналогично числовые равенства также могут содержать не только числа, но и буквы. Например, $6x - y = z + 2$. Буквы, под которыми понимаются некоторые числа называются *неизвестными*.

Из школьного курса известно, что с числовыми равенствами можно производить преобразования, которые не меняют их смысл (то есть оставляют их верными, если они были верными или не верными, если они были не верными):

– к левой и правой частям равенства можно прибавлять одно и то же число;

– левую и правую части неравенств можно умножать на одно и то же ненулевое число.

Числовое равенство, содержащее буквенные обозначения чисел, называется *уравнением.* Например, $3x = 12$ или $2x + 4y - 8 = 0$.

Для уравнения возникает естественный вопрос: какие числа необходимо подставить вместо неизвестных, чтобы числовое равенство стало верным? Найденные значения неизвестных будем называть *решением.*

Рассмотрим уравнение с одной неизвестной: $ax = b$, где a и b числа. Если $a = 0, b \neq 0$, то в соответствии с правилом умножения на число 0 приходим к выводу, что такое уравнение решений не имеет. Если $a = 0, b = 0$, то такое уравнение имеет бесконечно много решений – любое $x \in R$. Наконец пусть $a \neq 0$, тогда умножая левую и правую части уравнения на $1/a$ получим $x = b/a$. Так как такое преобразование не меняет смысл равенства, то исходное уравнение имеет только одно решение $x = b/a$.

Рассмотрим систему m линейных уравнений с n неизвестными:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{5.1}$$

где a_{ij}, b_i – известные действительные числа, $i, j = 1, 2, \ldots, m, n$, а x_j – переменные (неизвестные величины).

Решением системы линейных уравнений (5.1) называется упорядоченная совокупность чисел $(x_1^*, x_2^*, \ldots, x_n^*) = x^*$, при подстановке которых в (5.1) вместо неизвестных $(x_1, x_2, \ldots, x_n) = x$ каждое уравнение системы становится верным числовым равенством.

Система уравнений называется *совместной*, если она имеет хотя бы одно решение. В противном случае система называется *несовместной.*

Совместная система называется *определённой*, если она имеет единственное решение. В противном случае система называется *неопределённой.*

Система уравнений называется *однородной*, если все её свободные члены равны нулю.

Теорема 11. *Система однородных уравнений всегда имеет решение.*

Доказательство. Подставим в систему однородных уравнений $x = (0, 0, \ldots, 0)$. Очевидно получим верные равенства. Теорема доказана.

Решение $x = (0, 0, \ldots, 0)$ называется *тривиальным.*

Таким образом в случае однородных уравнений одно решение есть всегда. Для нахождения всех решений произвольной системы уравнений, в том числе и однородной потребуется преобразовать исходную систему.

Определим *вектор-столбец неизвестных* x, *вектор-столбец свободных членов* b и *основную матрицу* A данной системы уравнений, составленную из коэффициентов при неизвестных, формулами:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{5.2}$$

Тогда рассматриваемую систему линейных уравнений можно записать в *матричном виде* $Ax = b$ или:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} \tag{5.3}$$

Так как система (5.1) следует из правила равенства матриц, в данном случае двух вектор-столбцов.

Кроме матрицы коэффициентов A вводят также расширенную матрицу A^r, приписывая к основной матрице справа столбец свободных членов:

$$A^r = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right) \tag{5.4}$$

Если в системе линейных уравнений число неизвестных равно числу строк, то основная матрица этой системы – квадратная. Из теории матриц известно, что когда ранг квадратной матрицы равен числу её строк, то у матрицы коэффициентов системы имеется обратная матрица. Следовательно система имеет единственное решение $x = A^{-1}b$. Но остаётся вопрос о решении системы (в том числе его единственности) в случае произвольной основной матрицы.

Введём ещё одну форму записи системе линейных уравнений (5.1). Представляя основную матрицу A в виде вектор-столбцов $A^1, A^2, \dots, A^n$, запишем система линейных уравнений (5.1) в *векторном виде:*

$$x_1A^1 + x_2A^2 + \dots + x_nA^n = b, \tag{5.5}$$

где

$$A^1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, A^2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}, \dots, A^n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}.$$

Из (5.5) следует: система линейных уравнений $m \times n$ может быть представлена в виде разложения вектора $b \in R^m$ – вектора свободных членов (5.1) по n векторам $A^1, A^2, \dots, A^n \in R^m$ – вектор-столбцам матрицы коэффициентов (5.1), при этом коэффициентами разложения (5.5) оказываются переменные (5.1). Например, для системы

$$\begin{cases} x + 3y = 2 \\ 2x + y = 2 \end{cases} \qquad \vec{A}^1 x + \vec{A}^2 y = \vec{b} \tag{5.6}$$

её векторная запись означает: можно ли вектор $\vec{b} = (2, 2)$ представить в виде суммы двух векторов $\vec{A}^1 = (1, 2)$ и $\vec{A}^2 = (3, 1)$? Нарисовав векторы $\vec{A}^1$ и $\vec{A}^2$ на плоскости нетрудно убедиться в существовании решения, его даже можно указать, глядя на рисунок: $0.8\vec{A}^1 + 0.4\vec{A}^2 = 0.8(1, 2) + 0.4(3, 1) = (2, 2)$, что совпадает с решением (5.6), полученного методом подстановки $x = 0.8, y = 0.4$. Заметим, что решение оказалось единственным, как одна точка на плоскости, определяемая двумя координатами.

Рассмотрим другую систему линейных уравнений

$$\begin{cases} 2x + 2y = 2 \\ x + y = 2 \end{cases} \qquad \begin{cases} 2(x + y) = 2 \\ x + y = 2 \end{cases} \tag{5.7}$$

Понятно, что эта система не имеет решений, так как левая часть верхнего уравнения есть нижняя часть, умноженная на два, а правая часть не изменилась (выражение $x+y$ не может одновременно равняться и одном и двум). Векторная форма записи (5.7) также демонстрирует отсутствие решений

$$x \begin{pmatrix} 2 \\ 1 \end{pmatrix} + y \begin{pmatrix} 2 \\ 1 \end{pmatrix} = (x + y) \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \tag{5.8}$$

Действительно, из рисунка видно, что умножение вектора $\vec{a} = (2, 1)$ на число даст вектор на прямой, проходящей через этот вектор, при этом никогда не получить вектор $\vec{b} = (2, 2)$. То есть вектор $\vec{b}$ должен быть равен вектору $\vec{a}$, умноженному на некоторое число, или, как видно из (5.7), вектор $\vec{b}$ должен являться линейной комбинацией векторов $\vec{A}^1$ и $\vec{A}^2$. Иными словами, максимальное число линейно независимых вектор-столбцов матрицы A должно быть равно числу линейно независимых вектор-столбцов расширенной матрицы. Вспомним, максимальное число линейно-независимых векторов системы векторов называется рангом системы. Наконец, оказывается верно и обратное, что в итоге составляет содержание основной теоремы теории решения систем линейных уравнений

Теорема 12. *(Теорема Кронекера-Капелли.) Система линейных уравнений имеет решения тогда и только тогда, когда ранг основной матрицы системы равен рангу расширенной матрицы этой системы.*

Глава 6

Элементы математического анализа

6.1 Множества вещественных чисел

В курсе математического анализа удобно использовать ряд обозначений. Договоримся, что $\cap$ - пересечение, $\cup$ - объединение, $\emptyset$ - пустое множество, $\forall$ - любой, $\exists$ - существует, $|$ - такой что, $]$ – пусть.

Натуральные числа это числа счёта – $1, 2, \ldots, n, \ldots$. Обозначаются как N.

Целые числа – натуральные числа, им противоположные и ноль. Обозначаются как Z.

Рациональные числа – числа, представимые в виде дроби $\dfrac{m}{n}$, где m – целое число и n – натуральное. Обозначаются как Q.

Действительные числа – числа, представимые в виде всевозможных *десятичных дробей*, то есть в виде $a_0, a_1 a_2 \ldots, a_n \ldots$, где a_0 – целое число, а $a_1, a_2, \ldots, a_n, \ldots$ – число $0, 1, \ldots, 9$. Обозначаются дествительные числа как $R = (-\infty, +\infty)$.

Действительные числа, не являющиеся рациональными, называются *иррациональными.*

Из определения десятичных дробей следует, что $N, Z, Q \subset R$.

Множество действительных чисел $\{x\}$ называется *ограниченным сверху (снизу)*, если существует вещественное число M (m), что для всех $x \in \{x\}$ справедливо равенство: $x \leq M$ $(x \geq m)$. При этом число M (m)

называется *верхней гранью (нижней гранью)*.

Множество, ограниченное и сверху и снизу называется *ограниченным*. Заметим, что из определения следует, что множество имеет бесконечно много верхних и нижних граней.

Наименьшая из всех верхних граней множества $\{x\}$ называется *точной верхней гранью* этого множества и обозначается как $\bar{x} = \sup\{x\}$.

Наибольшая из всех нижних граней множества $\{x\}$ называется *точной нижней гранью* этого множества и обозначается как $\underline{x} = \inf\{x\}$.

Оказывается, что вещественные числа обладают замечательным свойством, выраженном в следующей теореме

Теорема 13. *(О существовании точной грани ограниченного множества.) Если множество вещественных чисел содержит хотя бы один элемент и ограничено сверху (снизу), то существует вещественное число $\bar{x}$, которое является точной верхней гранью (точной верхней гранью) этого множества.*

Например, множество $(-2; 4]$ ограничено числами 4 сверху и –2 снизу. При этом верхней гранью являются числа 5, 6, 12.5 и т. д., а число 4 является точной верхней гранью. Аналогично нижней гранью являются числа -12, -7, -3.7 и т. д., а число –2 является точной нижней гранью.

Как видно из примера точная верхняя (нижняя) грань не обязательно является элементом этого множества.

Пусть R – множество действительных чисел. Назовём его подмножество $[a; b] = \{\forall x : x \in R, a \le x \le b\}$ – *отрезком*, подмножество $(a; b) = \{\forall x : x \in R, a < x < b\}$ – *интервалом*, подмножества $(a; b] = \{\forall x : x \in R, a < x \le b\}$ и $[a; b) = \{x : x \in R, a \le x < b\}$ – *полуинтервалами*. Отрезки, интервалы и полуинтервалы называются *промежутками* на числовой оси. Точки a и b называются *концами* этих промежутков, а точки x такие, что $a < x < b$, – их *внутренними точками*

ε-окрестность точки a называется интервал $(a-\varepsilon, a+\varepsilon)$, который в дальнейшем будем обозначать как $V_\varepsilon(a)$ или $U_\varepsilon(a)$. В том случае, если точка a не принадлежит ε-окрестности точки a, такое множество будем называть проколотой ε-окрестностью $V_\varepsilon(a)$ и обозначать как $\dot{V}_\varepsilon(a)$. Таким образом $\dot{V}_\varepsilon(a) = V_\varepsilon(a) \backslash x$.

К множеству вещественных чисел полезно добавлять ещё два абстрактных числа: плюс и минус бесконечность. В этом случае говорят о расширенной числовой оси: $\bar{R} = [-\infty, +\infty]$. Тогда E-окрестностью «числа» $+\infty$ $(-\infty)$ понимается интервал $(E; +\infty)$ (соответственно $(-\infty; -E)$), при этом обязательно $E > 0$.

Важнейшим свойством множества действительных чисел, следующим непосредственно из определения их окрестностей, является то, что у двух различных точек числовой прямой *всегда имеются непересекающиеся окрестности*.

6.2 Функция и её предел

Пусть X и Y – некоторые множества. Говорят, что на множестве X задана функция со значениями в множестве Y, если имеется правило, по которому каждому элементу $x \in X$ указан ровно один элемент $y \in Y$.

Это соответствие записывается в виде $y = f(x)$. При этом x называют независимой переменной или аргументом функции, а y – значением функции. Множество X называют *областью определения функции* f и записывается как $D(f)$, а множество $E(f) = \{\forall y \in Y | \exists x \in X : y = f(x)\}$ – *множеством значений функции* f.

Соответствие $f : X \to Y$ может быть задано различным образом:

– в виде таблицы;

– с помощью описания;

– графически;

– аналитически.

Например, табличное задание функции может иметь вид:

x	-2	1	2	3	5	8	12	17
$f(x)$	-2	5	-4	7	13	12	3	7

Описание может быть, например, таким (*функция Дирихле*):

$$D(x) = \begin{cases} 1, & x \quad - \quad \text{иррациональное} \\ 0, & x \quad - \quad \text{рациональное.} \end{cases}$$

Можно предложить ещё одно описание функции: функция всякому чётному числу сопоставляет -1, а нечётному – ноль. На самом деле эту же функцию можно записать аналитически так: $f(2n) = -1, f(2n-1) = 0, n \in N$.

Приведённый пример является важным частным видом функции, когда областью её определения является множество натуральных чисел. Таким образом последовательность является частным видом функции, когда областью определения функции является множество натуральных чисел. Из «обычной функции» можно получить последовательность, ограничив её область определения натуральными числами. Например, если функция $f(x)$ имеет вид $f(x) = \frac{1}{x}$, то соответствующая последовательность приобретает вид $a_n = \frac{1}{n}$.

При аналитическом задании функции, когда множества X и Y – это числовые множества действительных чисел R, аналитической записи задания функции обычно можно сопоставить её графическое изображение. Так, на плоскости, в прямоугольной декартовой системе координат, *графиком* функции $f(x)$ называют множество точек $M(x, y = f(x))$. Примером графического изображения функции является прямая (1.8), изучаемая в курсе аналитической геометрии. Заметим, что кривые второго порядка, приведённые в той же главе, функциями не являются (попытайтесь сами ответить на вопрос почему это так). Но иногда графически изобразить функцию оказывается невозможным. Например, невозможно графически изобразить функцию Дирихле. Конечно попытаться можно, но без комментария к рисунку ничего понятно не будет.

Про функции говорят, что она *ограничена на множестве* X, если ограничено множество её значений на этом множестве.

Важным понятием для функции является определение её предела в некоторой точке. При этом точка не обязательно принадлежит области определения функции. В дальнейшем будет полезно использовать несколько различных определений, про которые нетрудно доказать, что они эквивалентны

(1) *Определение предела на языке $\varepsilon - \delta$ (читается как «эпсилон-дельта»)*: говорят, что число b является пределом функция $f(x)$ в точке a, если для любого $\delta > 0$, существует $\varepsilon > 0$, что для всех x, таких что $|x - a| < \delta, x \neq a$ выполняется неравенство $|f(x) - b| < \varepsilon$.

(2) *Определение предела на языке окрестностей:*

говорят, что *функция* $f(x)$ *в точке* a *имеет предел* b, если для $\forall U_\varepsilon(b)$ $\exists \dot{V}_\delta(a)$ $|$ $\forall x \in \dot{V}_\delta(a) \Rightarrow f(x) \in U_\varepsilon(b)$ (читается так: для любой ε-окрестности точки b существует проколотая δ-окрестность точки a такая, что для всех x из проколотой δ-окрестности точки a следует, что $f(x)$ принадлежит ε-окрестности точки b). (см. рис. 27)

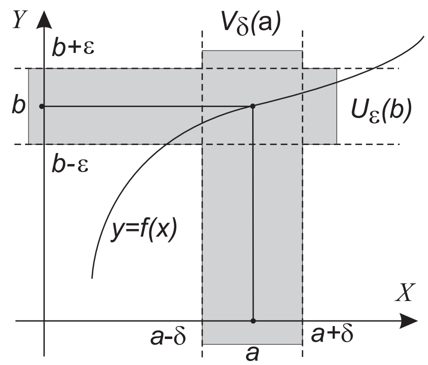


рис. 27.

Пример 1. Пользуясь определением, найдём предел функции $f(x) = 3x+2, D(f) = R$ в точке $a = 1$ из области определения функции. Так как $3 \cdot 1 + 2 = 5$, то возьмём произвольную полоску шириной $2\varepsilon > 0$ вокруг числа 5, тогда из рисунка видим, что если x входят в полоску 2δ, то $f(x)$ входит в полоску 2ε вокруг числа 5. Очевидно, это справедливо для любой 2ε-полоски. Следовательно в точке $x = 1$ функция $f(x) = 3x + 2$ имеет предел равный 5 (см. рис. 28).

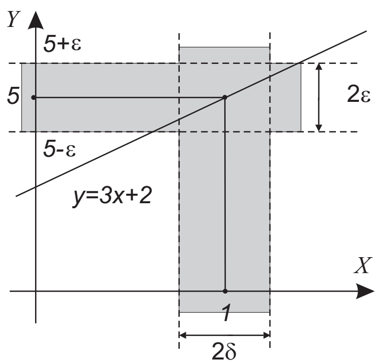


рис. 28.

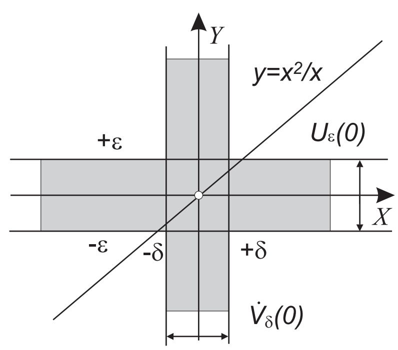


рис. 29.

2. $g(x) = \dfrac{x^2}{x}, D(f) = R \setminus \{0\}$. Заметим, что $f(x) = x$, если $x \neq 0$. В этом примере $\lim\limits_{x \to 0} f(x) = 0$. Доказать это можно так же, как в примере 1 по рис. 29.

Определение предела можно распространить на случай, когда $a = \infty$

и $a = -\infty$. Например, $\lim\limits_{x\to+\infty} f(x) = b$, означает, что для любой полоски шириной 2ε вокруг числа b, найдётся такое число $E > 0$, что для всех $x \in D(f)$, удовлетворяющих условию $x > E, f(x)$ принадлежало бы полоске вокруг числа b.

Для функции $f(x) = 1/x$ при $x \to \infty$ возьмём произвольную полоску вокруг числа 0 шириной 2ε, и в качестве E любое число, не меньшее $\dfrac{1}{\varepsilon}$, получим (рис. 30) $|1/x - 0| < 1/E \le \varepsilon$).

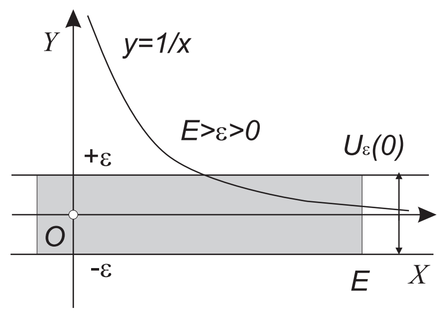


рис. 30.

Нетрудно убедиться в том, предел постоянной $f(x) = C$ равен постоянной в любой точке, то есть

$$\lim_{x\to a} C = C. \tag{6.1}$$

(Аккуратное доказательство этого факта с помощью рисунка, по аналогии с рис. 29 предлагается сделать самостоятельно.)

До сих пор в качестве предела функции выступало обязательно некоторое число из множества действительных чисел. Но на расширенной числовой оси рассматривают ещё один вид предела – бесконечность. При этом рассматривается отдельно предел равный плюс бесконечности, минус бесконечности и бесконечности. Выпишем все три определения предела на языке $\varepsilon - \delta$.

1. Говорят, что $+\infty$ является пределом функция $f(x)$ в точке a, если для любого $E > 0$, существует $\delta > 0$, что для всех x, таких, что $|x-a| < \delta, x \neq a$ выполняется неравенство $f(x) > E$.

2. Говорят, что $-\infty$ является пределом функция $f(x)$ в точке a, если для любого $E > 0$, существует $\delta > 0$, что для всех x, таких, что $|x-a| < \delta, x \neq a$ выполняется неравенство $f(x) < -E$.

3. Говорят, что ∞ является пределом функция $f(x)$ в точке a, если для любого $E > 0$, существует $\delta > 0$, что для всех x, таких, что $|x-a| < \delta, x \neq a$ выполняется неравенство $|f(x)| > E$.

Например, функция $f(x) = 1/x^2$ имеет предел, равный $+\infty$ при $x \to 0$, что нетрудно увидеть из графика. Кроме того $\lim\limits_{x\to+\infty} x = +\infty$, $\lim\limits_{x\to-\infty} x = -\infty$, $\lim\limits_{x\to\infty} x^2 = \infty$.

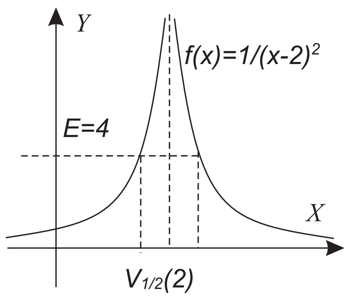


рис. 31.

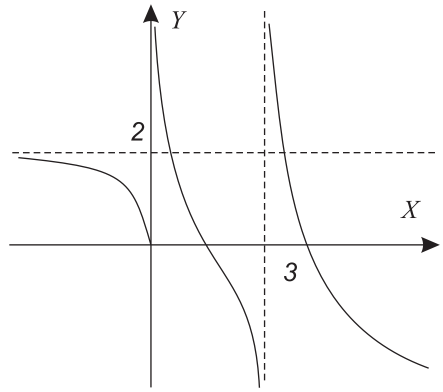


рис. 32.

На рис. 31 приведён пример функции $f(x) = 1/(x-2)^2$, предел которой равен $+\infty$ в точке 2. Действительно, для $E = 4$ выбираем малую окрестность, показанную на рисунке $\dot{V}_{\delta=\frac{1}{2}}(2) = (2-\frac{1}{2}; 2+\frac{1}{2})$, внутри которой для всех точек, за исключением точки $x = 2$ будет $f(x) > 4$. Аналогично можно найти нужную окрестность точки $x = 2$ и для любого другого числа E, поэтому действительно $\lim\limits_{x\to 2} = +\infty$.

Из определения предела следует важная

Теорема 14. *(О единственности предела.) Если предел в точке существует, то он только один.*

Чтобы лучше понять поведение функции вблизи точки вводят понятие *одностороннего предела* функции. В частности, предел слева функции $f(x)$ в точке a равный b на языке ε–δ записывается как $\lim\limits_{x\to a-0} f(x) = f(a-0) = b$ и означает, что для любого $\varepsilon > 0$, существует $\delta > 0$, что для всех x, таких что $0 < a - x < \delta$ выполняется неравенство $|f(x) - b| < \varepsilon$. Аналогично, предел справа функции $f(x)$ в точке a равный b на языке $\varepsilon - \delta$ записывается как $\lim\limits_{x\to a+0} f(x) = f(a+0) = b$ и означает, что для любого $\varepsilon > 0$, существует $\delta > 0$, что для всех x, таких что $0 < x - a < \delta$ выполняется неравенство $|f(x) - b| < \varepsilon$.

На рис. 32 приведён пример функции, для которой:

1. $\lim\limits_{x\to-\infty} f(x) = 2$,
2. $\lim\limits_{x\to+\infty} f(x) = -\infty$,
3. $\lim\limits_{x\to 0-} f(x) = 0$,
4. $\lim\limits_{x\to 0+} f(x) = \infty$,

5. $\lim\limits_{x\to 0} f(x) =$ не существует,
6. $\lim\limits_{x\to 3-0} f(x) = -\infty$,
7. $\lim\limits_{x\to 3+0} f(x) = +\infty$,
8. $\lim\limits_{x\to 3} f(x) = \infty$.

Дополнительно можно сказать, что при $x \to +\infty, x \to 3$ у функции, заданной графически на рис. 32 нет сходимости, а в точке $x = 0$ у функции не предела (Это следует из того, что на языке последовательностей для последовательности слева найдём преле равный нулю, а для последовательности справа предел будет равен бесконечности. Следовательно нашли два разных предела. Тогда по теореме о единственности предела приходим к выводу, что предела нет.)

Нахождение предела упрощается, если использовать его свойства.

Пусть функции $f(x)$ и $g(x)$ в точке a имеют конечные пределы, тогда выполняются следующие свойства.

1. $\lim\limits_{x\to a}(f(x)+g(x)) = \lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x)$
2. $\lim\limits_{x\to a}(f(x)-g(x)) = \lim\limits_{x\to a} f(x) - \lim\limits_{x\to a} g(x)$
3. $\lim\limits_{x\to a}(f(x)\cdot g(x)) = \lim\limits_{x\to a} f(x) \cdot \lim\limits_{x\to a} g(x) \quad (\lim\limits_{x\to a}(kf(x)) = k\lim\limits_{x\to a} f(x))$
4. $\lim\limits_{x\to a}(f(x)/g(x)) = \lim\limits_{x\to a} f(x)/\lim\limits_{x\to a} g(x) \quad (\lim\limits_{x\to a} g(x) \neq 0)$

Заметим, что требование сходимости в точке отдельно взятых функций $f(x)$ и $g(x)$ обязательно. Например, отдельно нет предела на бесконечности для функций $f(x) = x + 1/x$ и $g(x) = x - 1/x$, но зато их разность имеет предел на бесконечности (используем свойство 3 пределов и определение предела для функции $h(x) = 1/x$:

$$\lim_{x\to\infty}((x+1/x)-(x-1/x)) = \lim_{x\to\infty} 2/x = 0.$$

(Пример использования выписанных свойств предел будет приведён ниже.)

6.3 Бесконечно малые функции

Функция называется *бесконечно малой в точке*, если её предел в этой точке равен нулю.

Примеры бесконечно малых функций в точке $x = 0$:

$$f(x) = x^2, \sin x, x^4, x.$$

Видно, что $f(0) = 0$.

Примеры бесконечно малых функций в точке $x = x_0$:

$$f(x) = (x - x_0)^2, \sin(x - x_0), (x - x_0)^4, x - x_0,$$

так как $f(x_0) = 0$.

Примеры бесконечно малых функций на бесконечности:

$$f(x) = 1/x^2, 1/x^4, 1/x.$$

Важной характеристикой бесконечно малой, определяющей её важное значение является

Теорема 15. *Разность функции и её предельного значения в точке есть бесконечно малая в этой точке.*

Бесконечно малые обычно обозначают греческими буквами $\alpha(x, x_0)$. Здесь переменная x_0 указывает точку, в которой функция является бесконечно малой. Например, $\alpha(x, 3) = (x-3)^2$ – бесконечно малая в точке $x = 3$.

Бесконечно малые функции обладают важным свойством:

Сумма, разность и произведение бескончно малых в одной точке, а также произведение бесконечно малой на постоянную – также бесконечно малая в той же точке.

Например, функции $\alpha(x) = (x-5)^2, \beta(x) = (x-5)^4$. Тогда бесконечно малыми в точке $x = 5$ будут $\gamma_1 = \alpha(x, 5) + \beta(x, 5) = (x - 5)^2 + (x - 5)^4$, $\gamma_2 = 4\alpha(x, 5) = 4(x - 5)^2$, $\gamma_3 = \alpha(x, 5) \cdot \beta(x, 5) = (x - 5)^6$,

6.4 Непрерывность функции

Про функцию, предел которой в точке равен значению функции в этой точке говорят как о *функции непрерывной в этой точке.* Для непрерывной функции $f(a - 0) = f(a) = f(a + 0)$.

Если функция непрерывна в каждой точке множества D, то говорят, что функция непрерывна на множестве D.

Среди всего многообразия функций выделяется специальное множество, про которое говорят как об *элементарных функциях.* Это такие функции как $\sin x, \cos x, \operatorname{tg} x, \operatorname{ctg} x, \arcsin x, \arccos x, \operatorname{arctg} x, \operatorname{arcctg} x, \ln x$, $a^x (a > 0), x^\alpha (x > 0), x^n (n \in N), f(x) = Constant$.

Имеется важная теорема, отражающая основное свойство элементарных функций

Теорема 16. *Все элементарные функции непрерывны на своей области определения.*

Таким образом, например, функции $f(x) = x^2$ и $f(x) = 2^x$ – непрерывные функции для всех значений x, поэтому для вычисления их предела в любой точке достаточная найти их значения в этой точке. Например,

$\lim\limits_{x\to 3} x^2 = 3^2 = 9$ или $\lim\limits_{x\to 0} 2^x = e^0 = 1$.

Если в некоторой точке нарушается непрерывность функции, то говорят о разрыве функции в этой точке. При этом выделяют важный частный случай: если пределы слева и справа функции в точке равны, а значение функции этой точке либо не определено, либо отлично от значения односторонних пределов, то говорят об *устранимом разрыве* в этой точке. Например, функция $f(x) = x^2/x$ в точке $x = 0$ имеет устранимый разрыв и может быть доопределено требованием $f(0) = 0$.

6.5 Теоремы о непрерывных функциях

Функцию как пару значений $(x, f(x))$ удобно изображать в системе координат XOY в виде графика $y = f(x)$. Как уже указывалось это не всегда удаётся сделать. Например, это не так для функции Дирихле.

Нетрудно показать, что эта функция разрывна в каждой своей точке. С другой стороны, хотелось бы найти класс таких функций, которые можно было бы представить в виде графика, точнее в виде некоторой линии в системе координат XOY. Оказывается, что именно непрерывные функции являются наиболее подходящими для представления их в виде такой линии. Существенной помощью понимания того, что такое непрерывная функция являются нижеследующие теоремы.

I Теорема Вейерштрасса. *Если функция $f(x)$ непрерывна на отрезке [a,b], то она ограничена на этом отрезке.*

II Теорема Вейерштрасса. *Если функция $f(x)$ непрерывна на отрезке [a,b], то она достигает на этом отрезке своих точной верхней и нижней граней.*

Теорема о множестве значений функции. *Пусть функция $f(x)$ непрерывна на отрезке [a,b]. Пусть M – её максимальное значение и m – её минимальное значение. Тогда для всякого μ между наибольшим и наименьшим значениями функции, то есть для $m \le \mu \le M$, найдётся $x_0 \in [a, b]$, что $f(x_0) = \mu$.*

Теорема о нуле функции. *Пусть функция $f(x)$ непрерывна на отрезке [a,b] и на этом отрезке принимает положительные и отрицательные значения, то на этом отрезке найдётся точка, в которой значение функции равно нулю.*

Доказательство. По II теореме Вейерштрасса функция достигает на этом отрезке своё наибольшее и наименьшее значения. По условию на этом отрезке она принимает положительное и отрицательное значения, поэтому её наибольшее значение – положительно, а её наименьшее значение – отрицательно. Тогда по Теореме о множестве значений функции на этом отрезке найдётся точка, в которой значение функции равно нулю. Теорема доказана.

На самом деле непрерывность функции может нарушаться в отдельных точках. Это не сильно её портит (в каком-то смысле). Например, ступенчатая функция на рис. 33 стоимости покупок (ось OY) в зависимости от дохода (ось OX) отражает тот факт, что в обществе имеются социальные группы, для которых более высокий доход инициирует покупки более дорогих товаров в отличии от групп с меньшим доходам, ищущим аналогичный товар по более низкой цене.

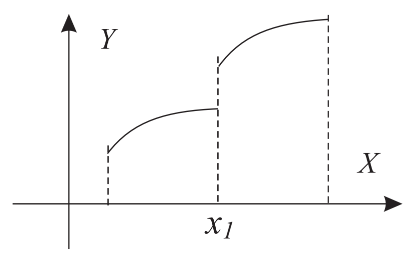


рис. 33.

Из приведённого графика видим, что разрыв в точке x_1 соответствует *скачку функции в этой точке.*

В заключении заметим, что именно кусочная непрерывность практически обеспечивает возможность построения графика функции.

6.6 Производная функции

Рассмотрим функцию $f(x)$, определённую на множестве D. Пусть $x \in D$. Придадим аргументу x приращение Δx, такое, что полученное число $x + \Delta x \in D$. Тогда $\Delta f(x) = f(x + \Delta x) - f(x)$ – *приращение функции* в точке x, вызванное приращением аргумента Δx.

Замечание. Мы используем термин «приращение» в смысле «изменение», а не «увеличение» Если, например, $\Delta x < 0$, то значение аргумента уменьшится в результате приращения Δx.

Если существует предел отношения приращения функции $\Delta f(x)$ к приращению аргумента Δx при $\Delta x \to 0$, то он называется производной функции в точке x и обозначается как $f'(x)$. Таким образом, по определению

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Нахождение производной называют дифференцированием.

Вводя обозначения $\Delta y = f(x + \Delta x) - f(x)$, получим ещё одну форму записи определения производной: $f'(x) = \lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x}$.

Пример 1. Найдём производную постоянной функции $y = C$.

$f(x) = C$ и $f(x + \Delta x) = C$,

$$C' = \lim_{\Delta x \to 0} \frac{C - C}{\Delta x} = \lim_{\Delta x \to 0} \frac{0}{\Delta x} = 0.$$

Таким образом $(C)' = 0$.

Пример 2. Найдём производную функции $y = x^2$.

$f(x) = x^2$, $f(x + \Delta x) = (x + \Delta x)^2$,

$$(x^2)' = \lim_{\Delta x \to 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \to 0} \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x} =$$

$$= \lim_{\Delta x \to 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x} = \lim_{\Delta x \to 0} (2x + \Delta x) = 2x.$$

Таким образом $(x^2)' = 2x$.

Производные функций можно вычислять из определения, т. е. вычисляя пределы, однако такой способ – процесс достаточно сложный, поэтому пользуются свойствами и таблицей производных элементарных функций, которые выводятся на основании определения производной.

Следующие свойства справедливы при условии, что все производные, указанные в них, существуют.

1-2. Производная суммы, разности двух функций равна сумме, разности этих функций соответственно.

$$(f(x) \pm g(x))' = f'(x) \pm g'(x).$$

3. Постоянная выносится за знак производной:

$$(C \cdot f(x))' = C \cdot f'(x).$$

4. Производные произведения и частного двух функций находятся по формулам:

$$(f \cdot g)' = f' \cdot g + f \cdot g' \qquad \left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2}.$$

Таблица производных.

N	$f(x)$	$f'(x)$	N	$f(x)$	$f'(x)$	N	$f(x)$	$f'(x)$
1	x^α	$\alpha x^{\alpha-1}$	5	e^x	e^x	9	$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
2	$\ln x$	$\frac{1}{x}$	6	a^x	$a^x \ln x$	10	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
3	$\sin x$	$\cos x$	7	$\operatorname{tg} x$	$\frac{1}{\cos^2 x}$	11	$\operatorname{arctg} x$	$\frac{1}{1+x^2}$
4	$\cos x$	$-\sin x$	8	$\operatorname{ctg} x$	$-\frac{1}{\sin^2 x}$	12	$\operatorname{arcctg} x$	$-\frac{1}{1+x^2}$

Например, найдём $f'(x)$, если $f(x) = 2 + x^2 - \sqrt{x} + 7x$.

$$f'(x) = (2 + x^2 - \sqrt{x} + 7x)' = (2)' + (x^2)' + (-x^{\frac{1}{2}})' + (7x^1)' =$$

$$= 0 + 2x^{2-1} + (-1) \cdot \frac{1}{2} \cdot x^{\frac{1}{2}-1} + 7x^{1-1} = 2x - \frac{1}{2\sqrt{x}} + 7.$$

6.7 Экстремум функции

Представим функцию $f(x)$ в виде графика $y = f(x)$ (рис. 34).

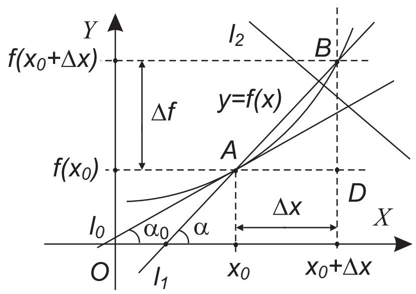


рис. 34.

Через две точки A и B графика функции проведём секущую l_1, которая пересекает ось OX под углом α. Очевидно, что $\mathrm{tg}\alpha = \dfrac{|BD|}{|AD|} = \dfrac{\Delta f}{\Delta x}$. Если устремить точку B к точке A, то полученная прямая, которая каждый раз будет проходить через две точки A и B, называется *касательной*.

Для касательной справедливо определение как *предельного положения секущей*. Исходя из способа её построения, видим, что касательная пересекает график функции только в одной точке, но обратное неверно, то есть не всякая прямая, пересекающая график функции в одной точке, является касательной. Например, прямая l_2 (на рис. 34) касательной не является.

Получим уравнение касательной: известно, что уравнение прямой, проходящей через одну точку имеет вид: $y = k(x - x_0) + y_0$. Из геометрического смысла коэффициента k (1.6) следует, что он равен тангенсу угла наклона прямой: $k = \mathrm{tg}\alpha_0 = \lim\limits_{\Delta x \to 0} \dfrac{\Delta f}{\Delta x} = f'(x_0)$, а так как $y_0 = f(x_0)$, то уравнение касательной принимает вид:

$$y_{кас.} = f'(x_0) \cdot (x - x_0) + f(x_0).$$

Этот вывод проясняет *геометрический смысл производной*: производная функции в точке x_0 равна тангенсу угла наклона касательной к графику в точке x_0, взятому в положительном направлении с осью абсцисс, т. е. угловому коэффициенту касательной.

Из геометрического смысла производной следует

Теорема 17. *Если функция дифференцируема в точке c и $f'(c) > 0$ (или < 0), то эта функция возрастает (или убывает) в этой точке.*

Доказательство. В самом деле, производная имеет смысл тангенса угла наклона касательной с осью OX, поэтому положительность производной означает положительность тангенса угла наклона, то есть угол с осью OX у касательной острый ($< 90^o$), то, как видно из рис. 34, в этой точке функция возрастает, если производная отрицательна, то касательная составляет тупой угол с осью абсцисс ($> 90^o$), то в этой точке функция убывает.

Говорят, что функция $f(x)$ имеет локальный минимум (максимум) в точке, если существует окрестность этой точки, в которой значение функции минимально (максимально).

Теорема Ферма. *Если функция дифференцируема в точке с и имеет в этой точке локальный экстремум, то* $f'(c) = 0$.

Доказательство. Так как $f(x)$ имеет локальный экстремум, следовательно по предыдущему рассуждению она в этой точке не может иметь ни положительную ни отрицательную производную, поэтому $f'(c) = 0$.

Из определения локального экстремума следует, что наибольшее (наименьшее) значение функции на отрезке не обязательно совпадает с локальным экстремумом. В частности, справедливо

Теорема 18. *Наибольшее (наименьшее) значение функция на промежутке принимает либо в точке локального экстремума, либо на концах этого промежутка, либо в точках, где функция не дифференцируема.*

Действительно, на примере рис. 35 видим, что в точке a функция принимает наименьшее значение, хотя локальный минимум находится в точке c, а наибольшее значение совпадает с точкой локального максимума – точкой d, при этом наименьшее значение функции находится в точке a.

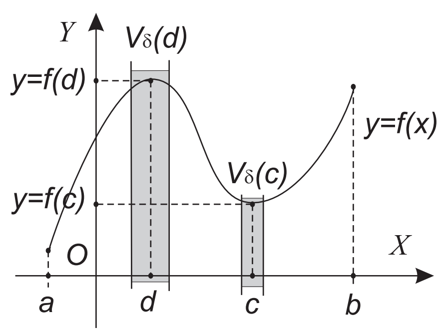


рис. 35.

Говорят, что дифференцируемая функция возрастает (убывает) в точке a, если существует такая окрестность этой точки, что для всех $x > a$ из этой окрестности $f(x) > f(a)(f(x) < f(a))$, а для всех $x < a$ следует, что $f(x) < f(a)(f(x) > f(x))$. Заметим, что о возрастании или о убывании функции в точке имеет смысл говорить только по отношению к дифференцируемой в этой точке функции, так как для такой функции можно выделить окрестность, во всех точках котой функция будет *монотонной*, то есть либо только возрастать, либо только убывать.

Дифференциалом функции $f(x)$ обозначается величина $df(x)$, равную $f'(x)\Delta x$. Но так как $x' = 1$, то $\Delta x = dx$, поэтому формулу для дифференциала функции записывают так:

$$df(x) = f'(x)dx.$$

Для примера найдём дифференциал функции $f(x) = 4x^5$.

Так как $f'(x) = 20x^4$, то $df(x) = 20x^4dx$.

На практике дифференциал используют для замены функции сложного вида на более простую линейную, если известно, что аргумент функции меняется незначительно. Это обусловлено тем, что как видно из рисунка 34 для малых изменений аргумента функции

$$df(x) \sim \Delta f(x) = f(x) - f(x_0)$$

Например, упростим с помощью дифференциала функцию $f(x) = \sqrt{x}$ для малых изменений x вблизи $x = 100$.

Так как $f'(x) = (\sqrt{x})' = \dfrac{1}{2\sqrt{x}}$, то $f'(100) = \dfrac{1}{20}$.

Так как $f(x_0) = f(100) = \sqrt{100} = 10$ и $\Delta x = x - x_0 = x - 100$, то

$$f(x) - f(x_0) = f(x) - 10 \sim f'(x_0)\Delta x = \frac{1}{20}(x - 100)$$

Следовательно вблизи $x_0 = 100$ функция $f(x)$ можно заменить на функцию $h(x)$:

$$f(x) = \sqrt{x} \sim h(x) = \frac{1}{20}x + 5.$$

В частности, получим

$f(100) = \sqrt{100} = 10, \quad h(100) = 10$ – точное равенство.

$f(99) = \sqrt{99} = 9.94987, \quad h(99) = 9.95$ – ошибка равна 0.00013. То есть одной сотой процента!

$f(101) = \sqrt{100} = 10.0498756, \quad h(100) = 10.05$ – ошибка равна 0.0001244!

6.8 Теоремы о дифференцируемых функциях

Дифференцируемость функции на интервале, как ранее было сказано, означает наличие конечной производной в каждой точке этого интервала.

Все обычные функции: степенные, показательные, логарифмические, дробные и т. п. непрерывны и дифференцируемы на своей области определения.

Дифференцируемость накладывает дополнительные требования к функции. Например, функция $y = |x|$ не дифференцируема при $x = 0$, хотя график этой функции изобразить можно (рис. 36), так как она непрерывна. То, что функция $y = |x|$ не дифференцируема в нуле понятно из геометрического смысла производной: так как производная соответствует тангенсу угла наклона, то из графика видим, что справа производная равна единице, а слева производная равна -1. Таким образом дифференцируемость на интервале означает её гладкость, и чем большее число раз функция дифференцируема, тем более она гладкая функция.

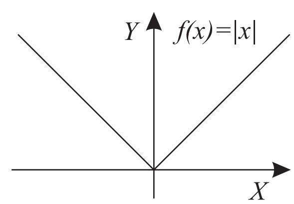


рис. 36.

Например, функция $y = x|x|$ является и непрерывной и дифференцируемой на всей числовой оси. Но в нуле эта функция не имеет второй производной, так как $f''(0-) = -2, f''(0+) = 2$.

Имеются полезные теоремы для функций, которые дифференцируемы по крайней мере один раз.

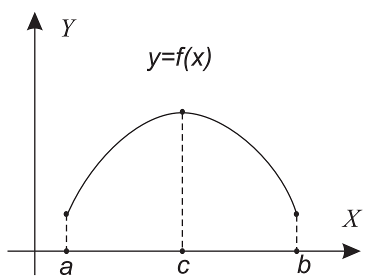


рис. 37.

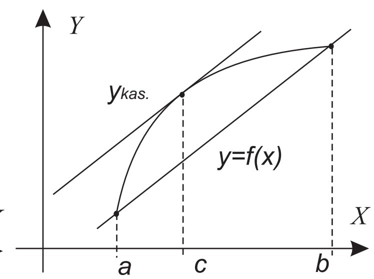


рис. 38.

Теорема Ролля. *Пусть функция непрерывна на отрезке* $[a, b]$ *и дифференцируема на интервале* (a, b)*. Пусть на концах отрезка функция принимает одинаковые значения. Тогда на интервале* (a, b) *найдётся точка, в которой значение производная функции равна нулю.*

Доказательство. Если функция постоянна на отрезке, то производная в любой внутренней точке равна нулю. Если она не постоянна, то по II Теореме Вейерштрасса имеются точка, в которой функция принимает наибольшее и наименьшее значения, хотя бы одно из которых отлично от значения на концах (см. рис. 37). Но тогда по Теореме Ферма производная в этой точке равна нулю. Теорема доказана.

Теорема Лагранжа. *Пусть функция непрерывна на отрезке* $[a, b]$ *и дифференцируема на интервале* (a, b)*. Тогда на интервале* (a, b) *найдётся точка* $c \in (a, b)$*, в которой выполняется равенство*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Доказательство. Действительно, из графика на рисунке (см. рис. 38) видно, что секущая, проходящая через точки a и b, имеет тангенс угла наклона, равный правой части в формуле Лагранжа. Двигая секущую вверх (см. картинку), добираемся до точки c, в которой секущая превращается в касательную. Вспоминая геометрический смысл производной как тангенса угла наклона касательной, получим доказательство теоремы.

Доказанная теорема уже позволяет сказать о виде функции, обладающей указанным ниже свойством.

Теорема о постоянстве дифференцируемой функции, имеющей на интервале равную нулю производную. *Если в каждой точке некоторого интервала* (a, b) *производная функции равна нулю, то это функция постоянна на всём интервале.*

Доказательство. Действительно, по теореме Лагранжа на любом отрезке заданного интервала $f(b) = f(a) = Constant$. Теорема доказана.

Из теоремы Лагранжа можно вывести ещё одну важную теорему о виде функции.

Формула Тейлора с остаточным членом в форме Лагранжа. *Пусть функция* $f(x)$ *на интервале* (a, b) n *раз дифференцируема. Тогда между любыми двумя точками* x, x_0 *интервала* (a, b) *найдётся точка* c*, что выполняется равенство*

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots$$

$$+\frac{f^{(n-1)}(x_0)}{(n-1)!}(x - x_0)^{n-1} + \frac{f^{(n)}(c)}{n!}(x - x_0)^n \qquad (*)$$

В том случае, когда n производная функции $f(x)$ ограничена, а для этого по I теореме Вейерштрасса достаточно считать n производную непрерывной функцией, последний член является бесконечно малой n-го порядка в точке x_0. Тогда формула $(*)$ может быть записана следующим образом

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots$$
$$+ \frac{f^{(n-1)}(x_0)}{(n-1)!}(x - x_0)^{n-1} + o((x - x_0)^n)$$

и называется *формулой Тейлора в форме Пеано.*

Если функция $f(x)$ бесконечное число раз дифференцируема, то в формуле $(*)$ формально можно не останавливаться на некотором n, полагая что $n \to \infty$, тогда формула $(*)$ принимает вид

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \dots$$

Полученное разложение функции $f(x)$ называют *рядом Тейлора.*

В качестве точки x_0 часто выбирают ноль. Тогда формула ряд Тейлора принимает вид

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \dots$$

и называется *рядом Маклорена.*

Хорошо известны разложения в ряд Маклорена некоторых популярных функций

$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$

$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$

$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + \dots$

Важное значение формулы Тейлора состоит в том, практически любую функию можно представить в виде степенного ряда со сколь угодной точностью. При этом для малого значения переменного аргумета x относительно некоторого фиксированного значения x_0 в зависимости от требуемой точности можно ограничиться только вплоть до линейного поведения функции:

$$f(x) \sim f(x_0) + f'(x_0)(x - x_0) \tag{6.2}$$

Правая часть (6.2) имеет смысл касательной к графику функции.

6.9 Первообразная и неопределённый интеграл

Рассмотрим функцию $f(x)$, определённую на числовом множестве X. Функция $F(x)$ называется *первообразной функции* $f(x)$ на множестве X, если для всех $x \in X$ выполняется равенство:

$$F'(x) = f(x).$$

Совокупность всех первообразных называется *неопределённым интегралом*. Неопределённый интеграл обозначается как $\int f(x)dx$, где $f(x)dx$ – *подынтегральное выражение*, а $f(x)$ – *подынтегральная функция*. Связь неопределённого интеграла и первообразной даётся формулой:

$$\int f(x)dx = F(x) + C,$$

где $F(x)$ – первообразная для $f(x)$, а C – произвольная постоянная.

Например, функция $F(x) = \sin x$ является первообразной функции $f(x) = \cos x$ на $X \subset (-\infty, \infty)$, так как $(\sin x)' = \cos x$ при любом $x \in X$.

Для вычисления неопределённых интегралов используют таблицу первообразных от элементарных функций:

$f(x)$	$F(x)$	$f(x)$	$F(x)$	$f(x)$	$F(x)$
0	C	$\sin x$	$-\cos x$	$\frac{1}{1+x^2}$	$\mathrm{arctg}\, x$
1	x	$\cos x$	$\sin x$	$1/\sqrt{1-x^2}$	$\arcsin x$
$1/x$	$\ln x$	$1/\cos^2 x$	$\mathrm{tg}\, x$	e^x	e^x
$x^\alpha, \alpha \neq -1$	$\frac{x^{\alpha+1}}{\alpha+1}$	$-1/\sin^2 x$	$\mathrm{ctg}\, x$	a^x	$a^x/\ln a$

и свойства неопределённых интегралов:

1. $\left(\int f(x)dx\right)' = f(x)$.
2. $\int kf(x)dx = k\int f(x)dx, \qquad k \neq 0.$
3. $\int (f(x) \pm g(x))dx = \int f(x)dx \pm \int g(x)dx.$

Пример. Вычислить интеграл: $\int (3\sin x + 5 - 2x^4 + \frac{1}{x} - \frac{2}{\sqrt{1-x^2}})dx$

Решение. Применив свойства 2) и 3), получим:

$$\int (3\sin x + 5 - 2x^4 + \frac{1}{x} - \frac{2}{\sqrt{1-x^2}})dx =$$

$$= 3\int \sin x + 5\int dx - 2\int x^4 dx + \int \frac{1}{x}dx - \int \frac{2}{\sqrt{1-x^2}}dx.$$

Далее, используя формулы таблицы первообразных элементарных функций, находим

$$\int \sin x dx = -\cos x + C_1, \quad \int dx = x + C_2,$$

$$\int x^4 dx = \frac{x^5}{5} + C_3, \quad \int \frac{1}{x}dx = \ln|x| + C_4, \quad \int \frac{1}{\sqrt{1-x^2}}dx = \arcsin x + C_5.$$

Таким образом,

$$\int (3\sin x + 5 - 2x^4 + \frac{1}{x} - \frac{2}{\sqrt{1-x^2}})dx = -3\cos x + C_1 + 5x + C_2 - 2\frac{x^5}{5} + C_3 +$$

$$+ \ln|x| + C_4 - 2\arcsin x + C_5.$$

Обозначим $C = 3C_1 + C_2 - 2C_3 + C_4 - 2C_5$. Окончательно получаем:

$$\int (3\sin x + 5 - 2x^4 + \frac{1}{x} - \frac{2}{\sqrt{1-x^2}})dx =$$

$$= -3\cos x + 5x - 2\frac{x^5}{5} + \ln|x| - 2\arcsin x + C.$$

Проверим, что найденный интеграл является первообразной для заданной функции.

$(-3\cos x + 5x - 2\frac{x^5}{5} + \ln|x| - 2\arcsin x + C)' = 3\sin x + 5 - 2x^4 + \frac{1}{x} - \frac{2}{\sqrt{1-x^2}}$ – верно.

6.10 Определённый интеграл

Пусть функция $f(x)$ определена на отрезке $[a,b]$ (рис. 39). Разобьём $[a,b]$ на n произвольных частей точками $a = x_0 < x_1 < x_2 < \ldots < x_n = b$. В каждом из полученных отрезков $[x_{i-1}, x_i]$ выберем произвольную точку $\xi_i \in [x_{i-1}, x_i]$ и составим *интегральную сумму*

$$S_n = f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \ldots + f(\xi_n)(x_n - x_{n-1}).$$

Обозначим $\lambda_n = \max\limits_{1 \leqslant i \leqslant n} |x_i - x_{i-1}|$, т. е. λ_n – длина максимального из полученных в результате разбиения отрезков. Будем увеличивать число n таким образом, чтобы длина максимального промежутка стремилась к нулю: $\lambda_n \underset{n\to\infty}{\to} 0$. Тогда получим последовательность интегральных сумм S_n. Если существует предел S_n при $n \to \infty$ и $\lambda_n \to 0$, не зависящий от способа разбиения отрезка $[a,b]$ и способа выбора точек ξ_i, то он называется *определённым интегралом* функции $f(x)$ на $[a,b]$, а про функцию $f(x)$ говорят, что она *интегрируема*. Обозначается он как $\int\limits_a^b f(x)dx$, т. е.

$$\int\limits_a^b f(x)dx = \lim_{\lambda_n \to 0} S_n = \lim_{\lambda_n \to 0} \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).$$

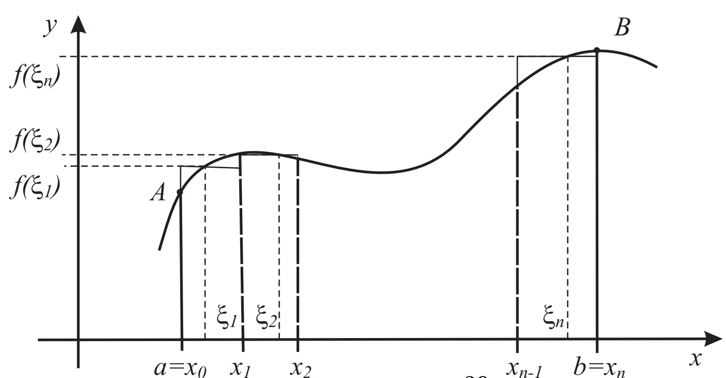


рис. 39.

Определённый интеграл вычисляется по *формуле Ньютона-Лейбница*:

$$\int_a^b f(x)dx = F(x)|_a^b = F(b) - F(a),$$

где $F(x)$ – первообразная функции $f(x)$. Если $f(x) \geqslant 0$, при всех $x \in [a, b]$, то определённый интеграл равен площади криволинейной трапеции (см. рис. 39).

Пример 1. Вычислить интеграл $\int_0^2 (3x^2 - 1)dx$.

Воспользовавшись правилами интегрирования и таблицей интегралов, найдём одну из первообразных функции $f(x) = 3x^2 - 1$:
$F(x) = \frac{3x^3}{3} - x = x^3 - x$. Тогда

$$\int_0^2 (3x^2 - 1)dx = (x^3 - x)|_0^2 = 8 - 2 - (0 - 0) = 6.$$

Пример 2. Найти площадь фигуры, ограниченной параболами $y = 2x^2 - 7x + 6$ и $y = x^2 - 4x + 10$.

Решение. Найдём точки пересечения парабол. Для этого приравняем значения обоих квадратичных функций, так как в точках пересечения одинаковы как абсциссы, так и ординаты:

$$2x^2 - 7x + 6 = x^2 - 4x + 10, x_1 = -1, x_2 = 4.$$

Откуда искомая площадь (заштрихованная область на рис. **??**) равна

6.11 Функции многих переменных

Функции вида $f(x) = x^2, f(x) = \sin x, \ldots x \in D(f)$, при этом $D(f)$ является подмножеством действительных чисел, называется функцией одной переменной. Но в качестве области определения можно взять, например точки на плоскости. Так как точки на плоскости определяются парой чисел $x = (x_1, x_2)$, то такую функцию называют функцией двух переменных. Аналогично вводится функции трёх и т. д. переменных. Соответствие, при котором n-числам, рассматриваемым как одна точка в n-мерном пространстве R^n, сопоставляется только одно число называют *функцией многих или нескольких переменных.*

Ранее (2.6) выписывалось уравнение плоскости в пространстве в виде $Ax+By+Cz+D=0$. При $C \neq 0$, получим $z = (-A/C)x+(-B/C)y+(D/C) = ax+by+c = z(x,y)$, то есть плоскость – это геометрическое представление линейной функции двух переменных.

В качестве другого примера можно привести функцию $f(x_1, x_2) = x_1^2 + x_2^2$, графиком которой является фигура в пространстве виде «чаши» (рис. 40). Обычно переменным x_1, x_2 сопоставляют переменные x, y соответственно, а значение функции обозначают переменной z (рис. 40). Областью определения функции $D(f)$ является плоскость XOY, а её множеством значений $E(f)$ – неотрицательные числа оси OZ.

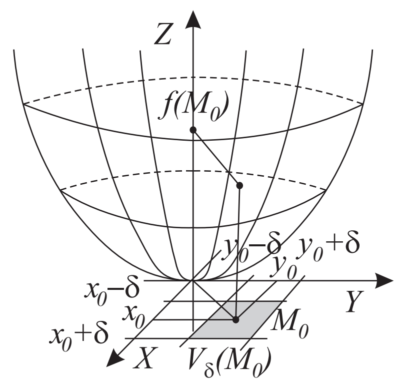


рис. 40.

δ-окрестностью точки $M_0 = (x_0, y_0)$ на плоскости является множество точек области, удовлетворяющие условию:

$$V_\delta(M_0) = \{x, y \in R : |x - x_0| < \delta, |y - y_0| < \delta\}$$

а её проколотой δ-окрестностью (рис. 40) – δ-окрестность без точки M_0:

$$\dot{V}_\delta(M_0) = V_\delta(M_0) \backslash M_0$$

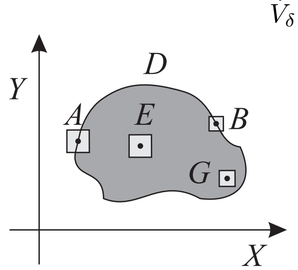


рис. 41.

Точка называется *внутренней точкой множества* M, если существует окрестность этой точки целиком входящая в область. Точка называется *граничное точкой* множества, если всякая окрестность этой точки содержит как точки этого множества, так и точки, не входящие в это множество.

На рис. 40 точки A и B – граничные точки, а точки E и G – внутренние точки области D.

Пределом функции двух переменных $f(M) = f(x, y)$ в точке $M_0 = (x_0, y_0)$ будет число b, если для любой ε-окрестности числа b ($U_\varepsilon(b)$) найдётся проколотая δ-окрестность точки M_0 ($\dot{V}_\delta(M_0)$), что для всех точек из $V_\delta(M_0)$ значение функции $f(x, y)$ находится в $U_\varepsilon(b)$.

Предел функции многих переменных обладает всеми теми же свойствами, что и предел функции одной переменной.

Говорят, что функция многих переменных непрерывна в точке, если её предел в этой точке равен значению функции в этой точке.

Непрерывны функции многих переменных обладают всеми теми же свойствами, что и непрерывные функции одной переменной.

Также как и для функции одной переменной для функции многих переменных справедлива

Теорема 19. *Всякая элементарная функция многих переменных непрерывна на своей области определения.*

Так как обычно приходится иметь дело с функциями одной переменной, то приходим к выводу, что практически все используемые функции непрерывна. Знание этого факта позволяет легко вычислить предел функции. Например, предел функции $f(x, y) = x^2 + y^2$ в точке $(3, -2)$ равен

$$\lim_{x \to 3, y \to -2} (x^2 + y^2) = 3^2 + (-2)^2 = 13$$

Понятие наибольшего и наименьшего значения функции многих переменных, её локального экстремума, аналогично соответствующим определениям функции одной переменной.

Для отрезка на числовой оси вводится понятие его концевой точки. В случае же области в n-мерном пространстве вводят понятие её границы, как множества граничных точек: точка называется граничной области D, если любая её окрестность содержит как точки области D, так и точки не входящие в область D.

Несколько сложнее обстоит дело с введением производной функции многих переменных, если мы, по-прежнему, под производной понимаем эффективность изменения функции: для различных переменных эффективность оказаться различной. Например, пусть прибыль фирмы z от рекламной кампании двух её отделов выражается формулой:

$$z = 3x + 2y,$$

где x, y – соответствующие финансовые вложения в отдел (в рублях). Откуда нетрудно видеть, что на один рубль, вложенный в первую фирму будет получен доход в 3 рубля (достаточно положить $x = 1, y = 0$), а на один рубль во вторую компанию – два рубля (достаточно положить $x = 0, y = 1$). По этой причине вводят понятие частной производной:

Частной производной функции многих переменных $f(x_1, x_2, \ldots, x_n)$ по переменной x_i называется предел вида

$$\lim_{\Delta x_i \to 0} \frac{f(x_1, \ldots, x_i + \Delta x_i, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{\Delta x_i}$$

который обозначается как $f'_{x_i}(x_1, \ldots, x_i, \ldots, x_n)$.

Из определения следует, что у функции n-переменных имеется n-различных частных производных. Например, для двух переменных вводятся две частные производные:

$$f'_x = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

$$f'_y = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

Если функция имеет частные производные в некоторой точке, то говорят что она дифференцируема в этой точке.

Так как частные производные определяются при постоянстве всех переменных, кроме той по которой берётся производная, то все свойства дифференцирования функции одной переменной автоматически переносятся на частные производные функции многих переменных.

Выпишем для примера несколько примеров взятия производных:

$$(x^3 + y^5 + 7)'_x = 3x^2, \quad (x^3 + y^5 + 7)'_y = 5y^4$$

$$(6x^3 - 2y^5 - 9)'_x = 18x^2, \quad (6x^3 - 2y^5 - 9)'_y = -10y^4$$

(При взятии частной производной надо помнить, что производная от суммы равна сумме производных, а производная от постоянной равна нулю.)

Частная производная по переменной x от функции двух перемнных $f(x, y)$ обозначается либо как f'_x либо как $\partial f / \partial x$. Во всяком случае получится новая функция опять двух переменных, которую обозначим, например, как $h(x, y) = f'_x$. Функцию $h(x, y)$ может быть опять продифференцирована, например, по переменной x. В результате получим $h'_x = f''_{xx}$, но функцию $h(x, y)$ можно продифференцировать по переменной y и получить $h'_y = f''_{xy}$, которую называют смешанной производной. Например, для $f(x, y) = 3x^5 y^8 - 2x^4$ получим $f'_x = 15x^4 y^8 - 8x^3$ и $f''_{xx} = 60x^3 y^8 - 24x^2, f''_{xy} = 120x^4 y^7$. При этом возникает вопрос о важности порядка дифференцирования по переменным – оказывается, что для непрерывных функций это не важно, то есть для функций, все частные вторые производные которых непрерывны, порядок дифференцирования не важен, то есть $f''_{xy} = f''_{yx}$.

Для точек локального экстремума функции, как точек наименьшего или наибольшего значения дифференцируемой функции в некоторой окрестности этой точки справедлива

Теорема 20. *В точках локального экстремума дифференцируемой функции все частные производные функции равны нулю.*

Ещё одна теорема позволяет выяснить какого рода экстремум – минимум или максимум.

Теорема 21. *Если частные производные в точке M_0 дважды дифференцируемой функции равны нулю и*

$$f''_{xx}(M_0) \cdot f''_{yy}(M_0) - (f''_{xy}(M_0))^2 > 0 \tag{6.3}$$

то в точке M_0 функция принимает максимум, если $f''_{xx}(M_0) < 0$ и минимум, если $f''_{xx}(M_0) > 0$.

Выражение в (6.3) является определителем, так как

$$f''_{xx}(M_0) \cdot f''_{yy}(M_0) - (f''_{xy}(M_0))^2 = \begin{vmatrix} f''_{xx}(M_0) & f''_{xy}(M_0) \\ f''_{yx}(M_0) & f''_{yy}(M_0) \end{vmatrix} \tag{6.4}$$

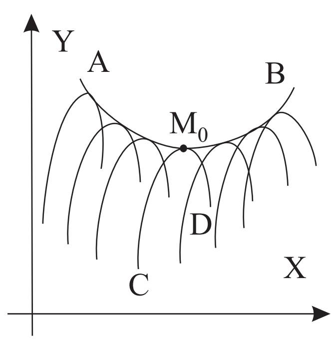


рис. 42.

Возникает вопрос, а что если в (6.4) определитель не положительный?

В том случае, когда он равен нулю, требуются дополнительные исследования, а в случае отрицательного значения имеет место «седловая точка» M_0, изображённая на рис. 42.: двигаясь вниз по кривой AB в точке M_0 получим минимум, а поднимаясь по кривой CD, в неё же получим максимум.

Пример. Найти минимальное значение функции функции

$$z(x, y) = x^2 + 4x + y^2 - 6x - 8y$$

Решение. Найдём частные производные

$$z'_x = 2x - 6, \quad z'_y = 2y - 8$$

Приравняем частные производные нулю, составим из них систему уравнений и решим её:

$$\begin{cases} 2x - 6 = 0 \\ 2y - 8 = 0 \end{cases} \quad \begin{cases} x = 3 \\ y = 4 \end{cases}$$

Таким образом точкой локального экстремума является точка $(3, 4)$. Найдем вторые производные

$$z''_{xx} = 2, \quad z''_{yy} = 2, \quad z''_{xy} = 0$$

Воспользуемся критерием (6.3): $2 \cdot 2 - 0^2 = 4, z''_{xx}(3, 4) > 0$. Таким образом это минимальное значение функции, которое равно

$$z(3, 4) = 3^2 + 4 \cdot 3 + 4^2 - 6 \cdot 4 - 8 \cdot 4 = -19$$

Для функции многих переменных наибольшее и наименьшее значения функции в области находят с помощью

Теорема 22. *Наибольшее и наименьшее значения функции многих переменных находятся либо в точках, где функция не дифференцируема, либо в точках локального экстрема, либо на границе области.*

Справедливость этой теоремы понятна из графиков функции одной переменной.

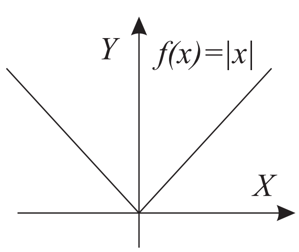


рис. 43.

На рис. 43 нарисован график функции $y = |x|$. В точке $x = 0$ эта функция не дифференцироуема, то есть у нею нет конечной производной. Действительно, справа от нуля угол наклона прямой составляет с осью OX угол 45^o. Но, как известно производная в точке равна тангенсу угла наклона в этой точке, поэтому при приближении к началу координат справа производная оказывется равной тангенсу 45^o, то есть плюс единице. При приближении к началу координат слева производная оказывется равной тангенсу -45^o, то есть минус единице. Так как производная в точке – это предел тангенса угла наклона в этой точке, независимо от того, с которой стороны приближаемся к этой точке, и так как, если предел существует, то он только один, то приходим к выводу, что в нуле производной функции $f(x) = |x|$ нет. Но, как видно из рисунка 43, именно в этой точке функция достигает своего минимального значения – это пример того, что функция достигает экстремума в точке, в которой она не дифференцируема. На рис. 44 изображена функция, которая достигает своего наименьшего значения на границе отрезка $[a, b]$, на котором задана функция. При этом минимальное значение функции на рис. 44 равно $f(a)$, а максимальное значение равно $f(b)$.

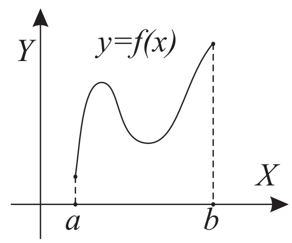


рис. 44.

Из рисунка также видно, что локальный экстремум достигается во внутренних точках отрезка, но они не являются ни максимумом, ни минимумом. В концевых точках производная не определена, так как для её определения необходима возможность подойти к точке как справа, так и слева. Но можно определить односторонний предел, как предел приближения к точке с одной стороны. В этом случае односторонние пределы в точках a и b, как видно из рис. 44, положительны.

Глава 7

Теория вероятностей

Введение

Задачи и проблемы, которые оказали существенное влияние на зарождение и первоначальное развитие теории вероятностей, возникли при обработке статистических данных и результатов наблюдений в различных областях человеческой деятельности: из практики страховых обществ, в связи с задачами из азартных игр и др. Первые статистические данные собирались ещё в древнем Египте, Греции и Риме. Это был подсчёт населения, количество ежегодно собираемого хлеба, падежей скота и т. п. В $\overline{\underline{\text{XIV}}}$ веке возникают первые морские страховые общества в Италии и Нидерландах, в которых производили подсчёт шансов, так как при большем риске собиралась большая страховая премия. Страховые премии при морских перевозках составляли 12 – 15 процентов, а внутри страны 6% – 8% от стоимости перевозимых грузов.

Большую роль в создании теории вероятностей сыграли азартные игры, которые существовали ещё в глубокой древности. Первоначально в качестве игральных костей обычно использовали кости животных - астрагалы, которые при бросании могли падать на четыре стороны. Бросая астрагалы, замечали какая сторона оказывалась сверху и её нумеровали.

Несмотря на столь раннее возникновение интуитивных представлений о вероятности, как наука теория вероятностей стала развиваться только в наше время. В частности, современный аксиоматический подход, которому придерживается изложенный здесь материал, был разработан А. Н. Колмогоровым в середине XX века.

7.1 Пространство элементарных событий

Первоопределяющим понятием теории вероятностей является *событие*. Также как, например, в геометрии основными, но неопределяемыми понятиями являются точка и прямая, событие независимым образом определить невозможно. Единственный путь – это описание того, что является событием.

На практике, в результате реального или мыслимого опыта, такого, как, например, бросание монеты сто раз, бросание трех игральных костей, сдача колоды карт, игра в рулетку, наблюдение продолжительности жизни человека, выбор наудачу некоторой группы людей и подсчет среди них числа левшей, исследование колебаний числа занятых междугородных линий на телефонной станции или числа телефонных вызовов, изучение результатов выборочного контроля качества промышленной продукции ... мы имеем дело с *возможными исходами рассматриваемого опыта или наблюдения*.

При бросании монеты не обязательно выпадает герб или решетка; монета может куда-нибудь закатиться или встать на ребро. Тем не менее мы условимся рассматривать герб и решетку как единственно возможные исходы бросания монеты. Это соглашение упрощает теорию и не сказывается на возможностях его применения. Идеализация подобного рода проводится постоянно. Невозможно безошибочно измерить продолжительность жизни конкретного человека; однако в теоретических исследованиях целесообразно считать эти величины точными числами. Но при этом возникает вопрос: какие числа могут и какие не могут представлять продолжительность жизни человека? Существует ли максимальный возраст, сверх которого жизнь невозможна, или для возрастов возможны любые значения? Мы, конечно, не решимся допустить, что человек может дожить до 1000 лет, и тем не менее обычная практика страхового дела не принимает никакой границы для продолжительности жизни. В соответствии с формулами, на которых основаны современные таблицы смертности, доля людей, доживающих до 1000 лет, имеет величину порядка единицы, деленной на $10^{10^{36}}$. Это утверждение лишено смысла с точки зрения биологии или социологии, но если его рассматривать исключительно с точки зрения статистики, то оно не противоречит опыту. В течение столетия рождается менее чем 10^{10}

людей, и чтобы статистически опровергнуть приведенное выше утверждение, потребовалось бы более чем $10^{10^{35}}$ столетий, что превышает возраст земного шара более чем в $10^{10^{34}}$ раз. Очевидно, приведённые числа совместимы с нашим представлением о невозможности. Можно было бы подумать, что их употребление является полным абсурдом; в действительности оно совершенно безвредно и приводит к упрощению многих формул. Кроме того, если бы мы решили всерьез исключить возможность дожить до 1000 лет, то мы должны были бы допустить существование максимального возраста. Но предположение, что можно дожить до x лет, но нельзя прожить x лет и две секунды может оказаться не только не привлекательным, но и ошибочным. Любая теория обязательно предполагает некоторую идеализацию. Желание же построить последовательную абстрактную модель опыта упирается в определение того, что представляют собой возможные исходы идеализированного опыта.

Договоримся результаты опытов или наблюдений называть *событиями.* Так мы будем говорить о событии, которое состоит в том, что из пяти брошенных монет более трех выпали гербом вверх. Мы будем различать *составные* события и *элементарные* события. Например, сказать, что сумма очков, выпавших при бросании двух игральных костей, равна шести, все равно что сказать, что произошло событие "(1, 5), или (2, 4), или (3, 3), или (4, 2), или (5, 1) и это перечисление разлагает событие "сумма очков равна шести"на пять элементарных событий. В качестве другого примера рассмотрим возраст человека – x. Каждое частное значение представляет *элементарное* событие, тогда как утверждение о том, что данному человеку пошел шестой десяток, описывает *составное* событие "$50 < \ < 60$". Итак, каждое составное событие может быть разложено на *элементарные* события; иначе говоря, составное событие есть совокупность *элементарных событий* ω. Совокупность всех элементарных событий будем называть *пространством элементарных событий* Ω.

В дальнейшем составные события будем называть просто событиями.

Рассмотрим ряд примеров, поясняющих выбор множества Ω.

1. Подбрасывание монеты один раз.

Возможными исходами в этом опыте будут: выпадение монеты гербом вверх (можно обозначить это событие или ω_1), выпадение решётки

(или ω_2). Таким образом, при описании этого опыта мы полагаем

$$\Omega = \{,\}, \quad \text{или} \quad \Omega = \{\omega_1, \omega_2\}.$$

В этом примере пространство элементарных событий состоит из двух различных событий.

2. Подбрасывание игральной кости один раз.

В этом опыте естественно выбрать $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, где ω_i обозначен исход опыта, заключающийся в выпадении i очков. Имеем шесть исключающих друг друга исходов.

3. Подбрасывание монеты 3 раза.

В этом случае

$$\Omega = \{\text{ГГГ, ГГР, ГРГ, РГГ, РРГ, РГР, ГРР, РРР}\}.$$

В этом примере пространство элементарных событий состоит из восьми различных событий.

4. Работа телефонной станции.

Предположим, что мы наблюдаем работу телефонной станции в течение часа и нас интересует число поступивших вызовов. Если телефонная станция обслуживает незначительное количество абонентов, то за время наблюдения может не поступить ни одного вызова, может поступить один вызов, два вызова и т. д. Ясно; что число вызовов будет всегда конечно. Однако разумно установить верхнюю границу числа вызовов довольно затруднительно. Проще не ограничивать возможное число вызовов и считать возможными исходами 0 и все натуральные числа: $\Omega = \{0, 1, 2, \ldots\}$. Это предположение проще, чем довольно искусственный подбор верхней границы числа вызовов, хотя предположение о возможности любого числа вызовов кажется абсурдным.

В этом примере пространство элементарных событий состоит из счётного числа различных событий.

Над событиями определим операции сложения и умножения:

Суммой двух событий A и B назовем событие $A + B$ (или $A \cup B$), состоящее из всех элементарных событий, принадлежащих по крайней мере одному из событий A или B. Можно сказать, что в реальном опыте событие, соответствующее $A + B$, состоит в том, что произошло по крайней мере одно из событий A или B.

Произведением AB (или $A \cap B$) называется событие, состоящее из элементарных событий, принадлежащих и A и B. Событие AB происходит тогда и только тогда, когда происходит и A и B.

Разностью $A \setminus B$ называется событие, состоящее из элементов множества A, не принадлежащих B (рис. 1, б). Событие $A \setminus B$ состоит в том, что A произошло, а B не произошло.

Событие Ω назовем *достоверным*, пустое множество $\emptyset$ назовем *невозможным* событием. Событие $\bar{A} = \Omega \setminus A$ называется *противоположным* событию A (см. рис. 1, г). Событие $\bar{A}$ означает, что A не произошло. Очевидно, что $\Omega \setminus \Omega = \emptyset$.

События A и B *несовместны*, если $AB = \emptyset$. Тот факт, что A является подмножеством B, будем записывать как $A \subset B$ (или $B \supset A$). Это значит, что из наступления события A следует наступление события B. Принадлежность элемента множеству обозначается символом $\in$. Например, $\omega \in \Omega$.

7.2 Вероятность

Числовая функция P, определенная на алгебре событий U, называется вероятностью, если выполнены следующие условия:

1. $P(A) \geq 0$ *для любого* $A \in U$.
2. $P(\Omega) = 1$.
3. *(аксиома конечной аддитивности). Если A и B несовместны, то*

$$P(A + B) = P(A) + P(B). \tag{7.1}$$

Тройку (Ω, U, P), в которой P удовлетворяет аксиомам 1 – 3, называют *вероятностным пространством*.

Понятие вероятностного пространства содержит лишь самые общие требования, предъявляемые к математической модели случайного явления, и не определяет вероятность однозначно. Дальнейшая конкретизация определения проводится применительно к рассматриваемой реальной задаче. Пусть, например, брошена один раз монетка. В этом случае $\Omega = \{\omega_\Gamma, \omega_P\}$. $U = \{\omega_\Gamma, \omega_P, \emptyset\}$, при этом ω_Γ и ω_P несовместны. Поэтому

$$P(\omega_\Gamma + \omega_P) = P(\omega_\Gamma) + P(\omega_P) = 1.$$

Определим $p_1, p_2 \geq 0$ как любые числа, для которых $p_1 + p_2 = 1$. И пусть $P(\omega_\Gamma) = p_1$ и $P(\omega_P) = p_2$.

Нетрудно проверить, что функция $P(A)$, таким образом определённая, удовлетворяет аксиомам 1 – 3. При различных наборах чисел (p_1, p_2)

будем получать разные определения вероятности. Очевидно, что посредством математических рассуждений из этого множества определений выбрать нужное мы не сможем.

Если монета симметрична, то представляется естественным дополнительное предположение о равновероятности выпадения различных граней. Тогда заключаем, что $p_1 = p_2 = 1/2$ и, следовательно, вероятность $P(A)$ определяется однозначно. Окончательное заключение о качестве выбранной модели может быть сделано после экспериментальной проверки.

Если монета имеет сложную форму, то в подходящей модели $p_1 \neq p_2 \neq 1/2$.

Из аксиом 1 – 3 следует ряд свойств вероятности.

Свойство 1.

$$P(\bar{A}) = 1 - P(A). \tag{7.2}$$

В самом деле, из определения противоположного события $\bar{A}$ следует, во-первых, что $\bar{A}+A = \Omega$ и во-вторых, что $\bar{A}$ и A несовместны. Тогда на основании аксиомы конечной аддитивности $P(\Omega) = P(A+\bar{A}) = P(A) + P(\bar{A}) = 1$. Откуда и следует утверждение.

Следствие.

$$P(\emptyset) = 0. \tag{7.3}$$

Следствие 1. *Для любых A и B справедливо*

$$P(A+B) \leq P(A) + P(B). \tag{7.4}$$

Следствие 2. *Если $A \subset B$, то*

$$P(A) \leq P(B). \tag{7.5}$$

7.3 Статистическое определение вероятности

Пусть в результате n испытаний некоторое событие A произошло k раз. Тогда *статистической* (или *стохастической*) вероятностью наступления события A назовём число

$$P(A) = \frac{k}{n}. \tag{7.6}$$

Например, монетка подбрасывается 500 раз. При этом герб выпал в 240 случаях. Тогда вероятность выпадения герба равна $240/500 = 0.48$.

Понятно, что в другом испытании эта вероятность окажется другим числом, так как представляется очень сомнительным, что при следующем подбрасывании мы опять получим 240 выпадений герба. Но зато представляется убедительной мысль, что результат буде близок к 0.48. Таким образом это число является хотя и не точным, но оценочным. В качестве самостоятельного задания можно подбросить монетку 500 раз и найти свою статистическую вероятность.

Можно убедиться, что предложенная формула удовлетворяет всем аксиомам теории вероятности.

Статистическая вероятность имеет важное значение, имея прямое отношение к наблюдению, так как может играть роль критерия правильности выбранного подхода к нахождению вероятности, найденной каким-либо другим методом, исследуемого события.

Можно показать, что статистическая вероятность обладает статистической устойчивостью, то есть с ростом числа испытаний n оно приближается к вполне определённому числу. Тогда важной задачей становится нахождение этого числа. Первая мысль того как найти это число – это сделать как можно больше испытаний. Но не всегда имеется возможность их проведения. Например, необходимо когда необходимо решиться на коммерческую сделку, но её успех имеет неопределённость.

Такого рода соображения, и не только, требуют других способов определения вероятности – теоретических, на основе модели. При этом определение вероятности должно быть таким, что в случае возможного большого повторения испытаний статистическая вероятность давала бы близкое число. На роль такой теоретической вероятности подходит классическая вероятность.

7.4 Классическое определение вероятности

Пусть $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ – пространство элементарных событий, а U – алгебра событий, содержащая все 2^s подмножеств $A = \{\omega_1, \omega_2, \ldots, \omega_k\}$ множества Ω. В *классическом определении вероятности* полагают, что все события равновозможны, то есть имеют одинаковую вероятность $P(\omega_i) = 1/n, i = 1, 2, \ldots, n$, поэтому вероятность $P(A)$ события $A = \{\omega_1, \omega_2, \ldots, \omega_k\}$ равна отношению числа элементарных событий ω_i, входящих в A, к общему числу элементарных событий в Ω:

$$P(A) = \frac{k(A)}{n(\Omega)}. \tag{7.7}$$

Такое определение вероятности удовлетворяет всем аксиомам теории вероятности.

В самом деле, выполнение условий 1 и 2 очевидно, так как, с одной стороны, дробь m/n не может быть отрицательной, а с другой – достоверному событию Ω благоприятствуют все n возможных результатов испытания, поэтому по формуле (7.7) получим, что $P(\Omega) = \frac{n}{n} = 1$.

Пример. На полке 5 книг, из них 2 – в переплёте и 3 – без переплёта. Наудачу взяли 2 книги. Какова вероятность того, что одна взятая книга в переплёте, а другая – без переплёта?

Решение. Обозначим книгу в переплёте как , а – без переплёта, тогда можно построить достоверное событие $\Omega = \{\underbrace{{}_{1}{}_{1}, {}_{1}{}_{2}, {}_{1}{}_{3}, {}_{2}{}_{1}, {}_{2}{}_{2}, {}_{2}{}_{3}}_{A}, {}_{1}{}_{2}, {}_{1}{}_{3}, {}_{1}{}_{2},$

${}_{2}{}_{3}\}$. Следовательно $\begin{cases} k(A) = 6 \\ n(\Omega) = 10 \end{cases}$. По классическому определению вероятности (7.7) получаем: $P(A) = \dfrac{6}{10} = \dfrac{3}{5} = \dfrac{1}{2}$. Заметим, что число благоприятных исходов и общее число исходов можно было найти по формуле гипергеометрического распределения: $n(\Omega) = C_5^2 = 10, k(A) = C_2^1 C_3^1 = 6$ без перечисления всех элементарных исходов.

Классическое определение вероятности служит хорошей математической моделью тех случайных явлений, для которых исходы опыта в каком-либо смысле симметричны и поэтому представляется естественным предположение об их равновозможности. Обычно такое предположение оправдано в задачах из области азартных игр, лотерей и т. д. Это объясняется тем, что при изготовлении игральных костей, карт и организации лотерей заботятся о соблюдении равновозможности различных исходов. Такие же требования предъявляются к организации выборочного контроля и выборочных статистических исследований.

7.5 Условная вероятность двух событий

При изучении реальных случайных явлении иногда возникает или искусственно создаётся ситуация, когда мы получаем дополнительную информацию о возможных исходах опыта, что может привести к изменению первоначальной вероятности $P(A)$. В этих случаях обозначим через $P_B(A) = P(A|B)$ – *условную вероятность события A при условии, что событие B уже произошло.* Например, пусть A – это возникновение несчастного случая (страхового события) с человеком, а B – то, что этот человек болен. Тогда $P(A)$ – вероятность возникновения страхового события, $P(B)$ – вероятность того, что человек болен, а $P_B(A)$ – вероятность возникновения страхового события, если известно, что человек болен. Очевидно, что $P(A)$ не будет совпадать с $P_B(A)$, так как в зависимости от болезни вероятность наступления страхового события в практике страховых компаний принимается различной.

Аксиоматически определяется, что для любых событий A и B из алгебры событий U, для которых $P(B) > 0$, справедлива формула *условной вероятности*

$$P_B(A) = \frac{P(AB)}{P(B)}, \tag{7.8}$$

которая является четвёртой аксиомой теории вероятности. Формулу (7.8) невозможно вывести из теоретических соображений, так же как ранее невозможно было определить событие, но её выполнение можно проверить для каждой вероятностной схемы.

Пример. Бросаются кости и пусть известно, что выпавшее число чётное. Чему равна вероятность того, что это число равно двум?

Решение. Пусть A означает событие, состоящее в том, что выпавшее число чётное, а B – это число равно двум. Тогда AB – число чётное и равно двум, поэтому $P(AB) = 1/6$ и $P(A) = 1/2$. По формуле (1) получаем $P_{\text{A=«чётное число»}}(B =<< 2 >>) = \dfrac{1/6}{1/2} = \dfrac{1}{3}$, так как $P(AB) = 1/6$ и $P(A) = 1/2$.

Следствие. *(Из определения условной вероятности) Для произвольных событий A и B, имеющих ненулевую вероятность, справедлива формула:*

$$P(AB) = P(A)P_A(B) = P(B)P_B(A). \tag{7.9}$$

Для доказательства (7.9) надо левую и правую часть (7.8) умножить на знаменатель правой части (7.8) и учесть равенство $AB = BA$.

Понятие условной вероятности позволяет естесственным образом определить независимость событий. Равенство $P_B(A) = P(A)$, согласуется с интуитивным представлением о независимости события A от B, однако за определение независимости двух событий A и B принимается более симметричное условие. События A и B называются *стохастически независимыми* или просто *независимыми*, если

$$P(AB) = P(A)P(B). \tag{7.10}$$

Это определение распространяется и на случай $P(B) = 0$, когда $P_B(A)$ не определена.

Термин *статистическая независимость* является синонимом термина *стохастическая независимость*. На практике обычно бывает ясно, что некоторые события должны быть независимыми, иначе вероятностная модель будет абсурдной. Однако, существуют ситуации, в которых стохастическая независимость может быть установлена только путём вычислений.

Пример. Подбросим одновременно две монеты. Какова вероятность того, что на обоих сторонах выпадет герб?

Решение. По формуле классической вероятности $P(\text{«герб»}) = 1/2$. Из теоретического рассмотрения опыта следует, что вероятность выпадения, например, «герба» на одной из монет не может зависеть от вероятности выпадения какой-либо из сторон на другой монете, поэтому события: «выпадение одной из сторон» на одной монете и «выпадение одной из сторон» на другой монете – являются независимыми. Следовательно $P(\text{«герб»} \cdot \text{«герб»}) = 1/2 \cdot 1/2 = 1/4$. Таким образом вероятность того, что на обоих монетках выпадет «герб» равна $1/4$.

7.6 Формула полной вероятности. Формула Байеса

Пусть $\Omega = H_1 + H_2 + \ldots + H_n$ так, что H_i попарно несовместны, то есть $H_i H_k = 0$, когда $i \neq k$. Про такие события говорят, что они образуют *полную группу событий* и обозначают как $\{H_i\}_1^n$.

Теорема 23. *Формула полной вероятности). Пусть A – произвольное событие и пусть $\{H_i\}_1^n$ – полная группа событий, тогда*

$$P(A) = \sum_{i=1}^{n} P(H_i) P_{H_i}(A). \tag{7.11}$$

В формуле полной вероятности используется полная группа событий и их вероятности, но часто новая информация изменяет вероятности полной группы событий. Чтобы её пересчитать пользуются формулой Байесса.

$$P_A(H_k) = \frac{P(H_k) P_{H_k}(A)}{\sum\limits_{i=1}^{n} P(H_i) P_{H_i}(A)}. \tag{7.12}$$

Пример. На фабрике изделия производится на двух линиях: на первой линии производится 30 % всех изделий, а на второй – 70 %. Брак на первой линии равен 2%, а на второй – 1%. Какова вероятность брака на фабрике? Какова вероятность того, что бракованное изделие изготовлено на первой линии?

Решение. Пусть событие A обозначает брак,

H_1 - изделие изготовлено на первой линии и

H_2 - изделие изготовлено на второй линии, тогда условие задачи запишется следующим образом:

$$\begin{array}{ll} P(H_1) = 0.3 & P_{H_1}(A) = 0.02 \\ P(H_2) = 0.7 & P_{H_2}(A) = 0.01. \end{array}$$

По формуле полной вероятности находим, что вероятность брака равна: $P(A) = P(H_1)P_{H_1}(A)+P(H_2)P_{H_2}(A) = 0.3\cdot0.02+0.7\cdot0.01 = 0.013.$

Тогда вероятность того, что бракованное изделие изготовлено на первой линии, равна $P_A(H_1) = \dfrac{0.3 \cdot 0.02}{0.013} = \dfrac{6}{13}$.

7.7 Случайные величины

Наблюдения за случайным явлением мы увидим некоторое его проявление, под которым понимаем случайное событие. В качестве случайного события может быть и неожиданный большой выигрыш или радостная встреча с кем-то из друзей. При этом не всякое событие явно иденцифицируется числом. Например, выигрыш в денежную лотерею это число – величина выигрыша, а вот встреча – это просто большая удача. Хотя, конечно же, даже такой эмоциональный фактор как встречу тоже можно «подсчитать». Например, как ноль или единица, в зависимости от того была она или нет, или как 100, 10 или 0 в зависимости отеё «ценности». Ограничимся в данном случае теми случайными событиями, которые можно представить в виде чисел. Такие события обладают замечательной особенностью – они обычно изначально несовместны. Действительно, если разным событиям сопоставлены разные числа, то два различных числа произойти не могут. Впрочем оговорка об обычной несовместности сделана не спрста. Напримр, совместны события A и B, если A – выигрыш до ста рублей, а событие B – выигрыш до тысячи рублей. Это события составные – оставим их для отдельного исследования, ограничиваясь элементарными событиями, для которых несовместность не вызываес сомнений.

Таким образом рассмотрим те случайные события, которым сопоставляется число и при том только одно. Математически это означает определение числовой функции на алгебре событий. Обозначим эту функцию как $X(\omega)$ (функция X от события ω) или просто X. Функция X называется *случайной величиной.*

Например, при подбрасывании игрального кубика выпадают случайные величины – число точек от 1 до 6 на верхней грани кубика.

Рассмотрим случай когда случайная величина принимает конечное число значений. В этом случае мы говорим о *дискретной случайной величине – ДСВ.* Сопоставим дискретным случайным величинам их вероятность как случайных событий и запишем результат в таблицу.

X	x_1	x_2	$\dots$	x_n
P	p_1	p_2	$\dots$	p_n

Данная таблица называется *законом рапределения дискретной случайной величины.*

Например, при подбрасывании кубика в соответствии с классическим определением вероятности всем событиям сопоставим число $1/6$, поэтому закон распределения дискретной случайной величины – «число на верхней грани кубика» имеет вид:

X	1	2	3	4	5	6
P	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

Важным обязательным свойством закона распределения ДСВ является то, что сумма всех вероятностей таблицы распределения равна единице.

Имеются и другие случайные величины. Например, температура на улице – это *непрерывная случайная величина – НСВ*, так как её значениями может быть любое число из некоторого интервала. В этом случае вместо таблицы значений случайной величины используют график неотрицательной функции, которую называют *функцией плотности распределения.* Важным свойством функции плотности является равенство единице площади под графиком функции плотности.

Наиболее часто используют функцию плотности *равномерного распределения непрерывной случайной величины* и *нормального распределения непрерывной случайной величины.*

Всегда предполагается, что равномерно распределённая случайная величина распределена на конечном промежутке $[a, b]$, на котором принимает постоянное значение, равное $1/(b-a)$ и нулю – вне отрезка $[a,b]$. Её функция плотности распределения указана на рис. 45.

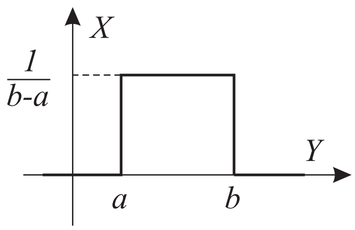


рис. 45.

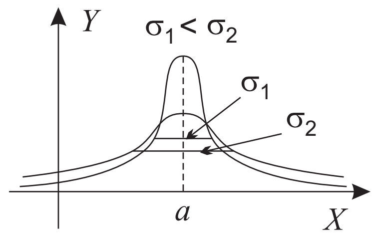


рис. 46.

Равномерное распределение характеризуется двумя параметрами. В качестве параметров, как видно из рис. 45, выступают концы отрезка, на котором равномерно распределённая случайная величина отлична от нуля.

$$f(x) = \begin{cases} \dfrac{1}{b-a}, x \in [a,b], a < b \\ 0, \quad x \notin [a,b], a < b \end{cases} \tag{7.13}$$

Нормально распределённая случайная величина определена на всей числовой оси и её график изображён на рис. 46. Нормально распределённая случайная величина также определяется двумя параметрами. Роль параметров распределения играют числа a – центр симметрии графика функции и параметр σ, характеризующий ширину распределения.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x-a)^2}{2\sigma^2}} \tag{7.14}$$

Смысл функции плотности распределения таков, что площадь под графиком на отрезке $[x_1, x_2]$ равна вероятности того, что событие примет значение на этом отрезке,

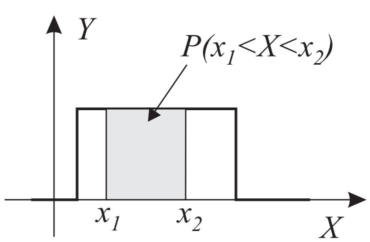


рис. 47.

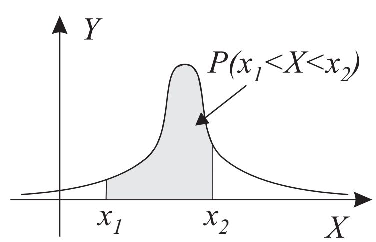


рис. 48.

поэтому её называют также *функцией плотности вероятности*.

Из графиков 45 – 46 заключаем, что внутри отрезка $[a,b]$ для одинаковых по длине промежутков $(x_2-x_1=x_2'-x_1')$ вероятности попадания в них одинаковы, что согласуется с названием рапределения – равномерное. В случае нормального распределения эта вероятность уменьшается с удалением от оси симметрии распределения.

7.8 Совместное распределение случайных величин

Рассмотрим случайную величину Z, возмжные значения которой определяются несколькими случайными значениями. Например, работа фирмы может определяться производительностью труда X и коньюктурой рынка Y, то есть двумя ***компонентами случайной величины*** $Z=(X,Y)$. Такую случайную величину называют двумерной. Случайная величина может быть и трёхмерной, и четырёхмерной и т. д. Ограничимся пока двумерным распределением.

Законом распределения дискретной двумерной величины называют перечень всевозможных значений пар чисел $(x_i,y_j), i=1,2,\ldots,n, j=1,2,\ldots,m$ и их вероятностей $p(x_i,y_j)=p_{ij}$, которую представляют в виде таблицы

Y	X			
	x_1	x_2	$\ldots$	x_n
y_1	p_{11}	p_{12}	$\ldots$	p_{1n}
y_2	p_{21}	p_{22}	$\ldots$	p_{2n}
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
y_m	p_{m1}	p_{m2}	$\ldots$	p_{mn}

Пример. Построить одномерные законы распределения случайных величин X и Y, зная их двумерное распределение

Y	X		
	4	6	9
2	0.04	0.1	0.3
7	0.16	0.2	0.2

Решение. Если случайнайная величина Y принимает значение $y_1=2$ и мы не интересуемся какие при этом значения примает случайнайная величина X, то это означает, что мы интересуемся вероятностью того, что случайная величина $Y = y_1$, при всевозможных значениях случайной величины X. Но, если $X = x_1 = 4$, то X не может принять значение $X = x_2 = 6$, так как она уже приняла значение, поэтому $X = x_1 = 4$ и $X = x_2 = 6$ – несовместны. Аналогичные рассуждения применимы и к $X = x_3 = 9$. В соответствии с аксиомой теории вероятностей вероятность несовместных событий складывается, таким образом

$P(Y = y_1 = 2) = p_{1.} = 0.04 + 0.1 + 0.3 = 0.44$
$P(Y = y_2 = 7) = p_{2.} = 0.16 + 0.2 + 0.2 = 0.56$
$P(X = x_1 = 4) = p_{.1} = 0.04 + 0.16 = 0.2$
$P(X = x_2 = 6) = p_{.2} = 0.1 + 0.2 = 0.3$
$P(Y = x_3 = 9) = p_{.3} = 0.3 + 0.2 = 0.5$

Строим одномерные распределения

Y	2	7
P	0.44	0.56

X	4	6	9
P	0.2	0.3	0.5

Как и должно быть, сумма вероятностей как для X так и для Y равна единице:

$$0.44 + 0.56 = 1 \qquad 0.2 + 0.3 + 0.5 = 1$$

Одномерные законы распределения записывают в таблицу совместного распределения, дописывая правый столбец и нижнюю строку

	X				
Y	x_1	x_2	$\ldots$	x_n	$P_{i.}$
y_1	p_{11}	p_{12}	$\ldots$	p_{1n}	$p_{1.}$
y_2	p_{21}	p_{22}	$\ldots$	p_{2n}	$p_{2.}$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
y_m	p_{m1}	p_{m2}	$\ldots$	p_{mn}	$p_{m.}$
$P_{.j}$	$p_{.1}$	$p_{.2}$	$\ldots$	$p_{.n}$	1

В таблице правый крайний столбец соответствует одномерной случайной величине Y

$$p_{1.} = p_{11} + p_{12} + \ldots + p_{1n} = P(Y = y_1)$$

и т. д., а самый нижней столбец – одномерной случайной величине X

$$p_{.1} = p_{11} + p_{21} + \ldots + p_{m1} = P(X = x_1)$$

и т. д. Сумма всех вероятностей как по столбцу, так и по строке равна единице.

Заметим, что зная закон совместного распределения случайных величин мы строим законы распределения одномерной случайной величины, но оказывается, что обратное в общем случае не верно.

7.9 Условное распределение случайных величин

Пусть случайная величина X принимает значения $(x_1, x_2, \dots, x_n)$, а случайная величина Y принимает значения $(y_1, y_2, \dots, y_m)$.

Обозначим условную вероятность того, что случайная величина Y примет значение y_j, при условии, что случайная величина X приняла значение x_i через $p(y_j|x_i)$.

Условным распределением дискретной случайной величины Y при $X = x_i$ называют совокупность условных вероятностей $p(y_1|x_i), p(y_2|x_i), \dots,$ вычисленных в предположении, что событие $X = x_i$ уже наступило. Тогда, зная закон распределения дискретной случайной величины $Z = (X, Y)$, можно, пользуясь формулой (7.8), вычислить условный закон распределения $P(Y|X)$:

$$p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}, \quad \text{где} \quad p(x_i) = \sum_i p(x_i, y_j), \quad i, j = 1, 2, \dots .$$

Понятно, что условная вероятность $p(y_j|x_i)$ обладает свойством неотрицательности и

$$\sum_j p(y_j|x_i) = 1.$$

7.10 Числовые характеристики случайных величин

1^0. **Математическое ожидание.**

Закон распределения случайной величины её полностью определяет. Однако при практических расчётах часто можно ограничиться более простыми характеристиками. Одной из таких важнейших характеристик является математическое ожидание случайной величины. Определение математического ожидания связано с обычным представлением о среднем значении.

Математическим ожиданием MX случайной величины X, имеющей закон распределения (1), называется число

$$M(X) = \sum_{\omega_i \in \Omega} \xi(\omega_i) p(\omega_i) = \sum_{i=1}^{n} x_i \cdot p_i \tag{7.15}$$

Если в случае ДСВ значения $x_i, i = 1, 2, \ldots, n$ случайной величины $\xi(\omega_i)$ считать компонентами вектора $x = (x_1, x_2, \ldots, x_n)$, то случайная величина представима n-мерным вектором x в пространстве R^n.

Пример. Пусть в некоторой игре вероятность выигрыша и проигрыша 100 руб. одинакова. Найдём математическое ожидание дискретной случайной величины – выигрыша.

$$M(\xi) = \sum_{i=1}^{2} \xi(\omega_i) p(\omega_i) = -100 \cdot \frac{1}{2} + 100 \cdot \frac{1}{2} = 0$$

Этот результат полностью согласуется с представлением о математическом ожидании как о среднем выигрыше.

2^0**. Дисперсия.**

Дисперсией $D(\xi)$ случайной величины ξ называется число

$$D(\xi) = M \overset{\circ}{\xi}{}^2 = M((\xi - M(\xi))^2), \tag{7.16}$$

если математическое ожидание в правой части (7.16) существует. Дисперсия является мерой рассеяния значений случайной величины около её математического ожидания. Величину $\sigma = \sqrt{D(\xi)}$ называют средним квадратическим отклонением. На рис. 46 она достаточно хорошо согласуется с геометрической полушириной кривой распределения плотности вероятности вблизи математического ожидания.

В формуле дисперсии введена новая случайная величина $\overset{\circ}{\xi} = \xi - M(\xi)$, которая называется *центрированной*.

Если воспользоваться свойствами математического ожидания, то

$$D(\xi) = M(\xi^2) - (M(\xi))^2.$$

Для дискретной случайной величины ξ, применяя (7.16) получим

$$D(\xi) = \sum_{i=1}^{\infty} (x_i - M(\xi))^2 P(\xi = x_i).$$

В случае векторного представления x ДСВ центрированный вектор $\overset{\circ}{x}$ строится по исходному вектору по правилу $\overset{\circ}{x}= (x_1 - MX, x_2 - MX, \ldots, x_n - MX)$. В этом случае дисперсия случайной величины ξ оказывается равной длине вектора x, то есть $D\xi = || \overset{\circ}{\xi} ||$.

Пример. Найдём дисперсию случайной величины.

3^0. **Ковариация.**

Ковариацией $\operatorname{cov}(\xi_1, \xi_2)$ двух случайных величин ξ_1 и ξ_2 называется число

$$\operatorname{cov}(\xi_1, \xi_2) = M(\overset{\circ}{\xi}_1, \overset{\circ}{\xi}_2) = M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)]. \tag{7.17}$$

Отсюда, используя свойства математического ожидания, легко получить следующую формулу:

$$\operatorname{cov}(\xi_1, \xi_2) = M\xi_1\xi_2 - M\xi_1 M\xi_2.$$

Из определения ковариации следует, что

$$\operatorname{cov}(\xi, \xi) = D\xi.$$

Ковариация показывает имеется или нет зависимость между случайными величинами.

4^0. **Коэффициент корреляции.**

Нетрудно доказать, что

$$\left|\frac{\operatorname{cov}(\xi_1, \xi_2)}{\sqrt{D\xi_1 D\xi_2}}\right| \leq 1. \tag{7.18}$$

Величина, стоящая в левой части неравенста (7.18) называется *коэффициентом корреляции* случайных величин ξ_1, ξ_2 и обозначается как $\rho(\xi_1, \xi_2)$.

Важным свойством коэффициента корреляции является то, что в случае, если связь между случайными величинами ξ_1 и ξ_2 линейна, то коэффициент корреляции равен ± 1.

Таким образом, если известно, что между двумя случайными величинами ξ_1, ξ_2 имеется линейная связь, и при возрастании значения случайной величины ξ_1 возрастает значение случайной величины ξ_2 – коэффициент корреляции равен «$+1$». Если же при возрастании значения случайной величины ξ_1 значение случайной величины ξ_2 убывает – коэффициент корреляции равен «-1».

Случайные величины, между которыми коэффициент корреляции равен нулю, называются *некоррелированными*.

Необходимо отметить, что равенство нулю коэффициента корреляции не является достаточным условием независимости случайных величин. Таким образом некоррелированность случайных величин не означает их независимость.

5^0. **Моменты высших порядков.**

Наряду с рассмотренными выше числовыми характеристиками случайных величин часто используются моменты более высоких порядков. *Моментом порядка* k случайной величины ξ называется число $M\xi^k$. Число $M\overset{\circ}{\xi}{}^k$ называется *центральным моментом порядка* k.

Пусть задан случайный вектор $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_n)$. Величины

$$M(\xi_1^{k_1}\xi_2^{k_2}\ldots\xi_n^{k_n}), \qquad M(\overset{\circ}{\xi}{}_1^{k_1}\overset{\circ}{\xi}{}_2^{k_2}\ldots\overset{\circ}{\xi}{}_n^{k_n})$$

называются соответственно *смешанным моментом порядка* $k = k_1 + \ldots + k_n$ и *смешанным центральным моментом порядка* k соответственно.

7.11 Задачи на законы распределения

Пример 3. Из урны, содержащей 5 белых шаров и 3 чёрных, достают по одному три шара. Составить закон распределения дискретной случайной величины X – число белых шаров из выбранных наудачу трёх шаров. Найти математическое ожидание, дисперсию и с. к. о.

Решение. Случайная величина X – число белых шаров среди трёх взятых может принимать значения от 0 до 3-х. Найдём соответствующие вероятности.

Обозначим взятый белый шар как событие A и соответственно взятый чёрный как $\bar{A}$, тогда $X = 0$ означает, что все шары чёрные, откуда

$$P(X = 0) = P(\bar{A}_1) \cdot P(\bar{A}_2|\bar{A}_1) \cdot P(\bar{A}_3|\bar{A}_1\bar{A}_2)$$

Здесь $P(\bar{A}_1) = \frac{3}{8}$ – вероятность того, что первый взятый шар – чёрный;
$P(\bar{A}_2|\bar{A}_1) = \frac{2}{7}$ – вероятность того, что второй взятый шар – чёрный, при условии, что первый взятый шар – также чёрный;
$P(\bar{A}_3|\bar{A}_1\bar{A}_2) = \frac{1}{6}$ – вероятность того, что третий взятый шар – чёрный, при условии, что первый и второй взятые шары – также чёрные. Откуда

$$P(X=0)=\frac{3}{8}\cdot\frac{2}{7}\cdot\frac{1}{6}=\frac{1}{56}$$

Один белый шар может быть в трёх случаях – если он взят первым, вторым или третьим. Так как это соответствует случаю «или» для несовместных событий, то необходимо найти все три вероятности и их сложить, то есть

$$P(X=1)=P(A_1)\cdot P(\bar{A}_2|A_1)\cdot P(\bar{A}_3|\bar{A}_2A_1)+$$

$$+P(\bar{A}_1)\cdot P(A_2|\bar{A}_1)\cdot P(\bar{A}_3|\bar{A}_1A_2)+P(\bar{A}_1)\cdot P(\bar{A}_2|\bar{A}_1)\cdot P(A_3|\bar{A}_1\bar{A}_2)=$$

$$=\frac{5}{8}\cdot\frac{3}{7}\cdot\frac{2}{6}+\frac{3}{8}\cdot\frac{5}{7}\cdot\frac{2}{6}+\frac{3}{8}\cdot\frac{2}{7}\cdot\frac{5}{6}=3\cdot\frac{5}{56}=\frac{15}{56}$$

Два белых шара может быть в трёх случаях – если первым, вторым или третьим взят чёрный, а два других шара в каждом из случае – белые. Так как эти случаи опять же соответствует «или» для несовместных событий, то необходимо найти все три вероятности и их сложить, то есть

$$P(X=2)=P(\bar{A}_1)\cdot P(A_2|\bar{A}_1)\cdot P(A_3|A_2\bar{A}_1)+$$

$$+P(A_1)\cdot P(\bar{A}_2|A_1)\cdot P(A_3|A_1\bar{A}_2)+P(A_1)\cdot P(A_2|A_1)\cdot P(\bar{A}_3|A_1A_2)=$$

$$=\frac{3}{8}\cdot\frac{5}{7}\cdot\frac{4}{6}+\frac{5}{8}\cdot\frac{3}{7}\cdot\frac{4}{6}+\frac{5}{8}\cdot\frac{4}{7}\cdot\frac{3}{6}=3\cdot\frac{5}{28}=\frac{15}{28}$$

Наконец, $X=3$ – взято три белых шара. Аналогично предыдущим случаям, получим

$$P(X=3)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1A_2)=\frac{5}{8}\cdot\frac{4}{7}\cdot\frac{3}{6}=\frac{5}{28}$$

Запишем результаты в таблицу – закон распределения.

X	0	1	2	3
P	$\frac{1}{56}$	$\frac{15}{56}$	$\frac{15}{28}$	$\frac{5}{28}$

Проверим, что закон распределения построен правильно:

$$\frac{1}{56}+\frac{15}{56}+\frac{15}{28}+\frac{5}{28}=1-\text{верно}$$

Найдём математическое ожидание.

$$MX=0\cdot\frac{1}{56}+1\cdot\frac{15}{56}+2\cdot\frac{15}{28}+3\cdot\frac{5}{28}=\frac{15}{8}=1.875$$

Найдём дисперсию. Для нахождения дисперсии воспользуемся формулой $DX = M(X^2) - (MX)^2$, поэтому сначала вячислим $M(X^2)$.

$$M(X^2) = 0^2 \cdot \frac{1}{56} + 1^2 \cdot \frac{15}{56} + 2^2 \cdot \frac{15}{28} + 3^2 \cdot \frac{5}{28} = \frac{225}{56}$$

Откуда

$$DX = \frac{225}{56} - \frac{225}{64} = \frac{225}{448} \sim 0.5$$

Наконец, находим среднее квадратическое отклонение

$$\sigma = \sqrt{DX} = \frac{15}{8\sqrt{7}} \sim 0.7$$

Полученные результаты показывают, что ожидаемое число белых шаров в одном опыте близко к двум ($MX = 1.875$). Большое с. к. о. ($\sigma = 0.7$) следует понимать как погрешность прогноза и говорит, что наиболее вероятно число белых шаров от одного до двух.

Пример 4. Случайная величина распределена равномерно на отрезке $[4; 12]$. Найти вероятность того, что она принимет значение на отрезке $[10; 16]$, а также M и DX.

Решение. В соответствии с определением нахождения вероятности непрерывного распределения как площади под графиком на соответствующем промежутке условие задачи может быть представлено как заштрихованная площадка на рис. 49. Так как высота прямоугольника по формуле (7.13) равна $\frac{1}{12-4} = \frac{1}{8}$, а ширина ненулевой части равна $12 - 10 = 2$, то

$$P(10 < X < 16) = 2 \cdot \frac{1}{8} = \frac{1}{4}$$

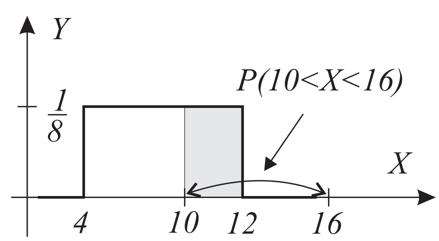


рис. 49.

Для равномерного распределения (7.13) известна формулы математического ожидания, дисперсии и с. к. о.:

$$MX = \frac{a+b}{2}, \qquad DX = \frac{(b-a)^2}{12}, \qquad \sigma = \frac{|b-a|}{2\sqrt{3}}$$

откуда

$$MX = \frac{12+4}{2} = 8, \qquad DX = \frac{(12-4)^2}{12} = \frac{16}{3}, \qquad \sigma = \frac{12-4}{2\sqrt{3}} = \frac{4}{\sqrt{3}}$$

Пример 5. Случайная величина X – ожидаемый процент избираетелей, проголосовавших за депутата распределена по норамальному закону с $MX = 61$ и $DX = 25$. Найти вероятность того, что за кандидита проголосует не менее 50% избирателей.

Решение. В соответствии с определением нахождения вероятности непрерывного распределения как площади под графиком на промежутке $[50; +\infty)$ вероятность – это заштрихованная площадка на рис. 50.

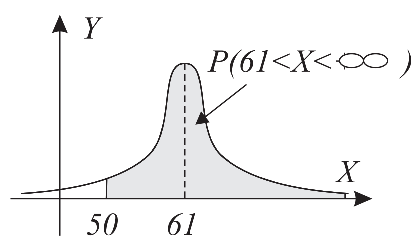


рис. 50.

Функция плотности $f_n(x)$ определена формулой (7.14). Известно математическое ожидание $MX = a = 61$ и с. к. о. $\sigma = \sqrt{DX} = \sqrt{25} = 5$. Следовательно,

$$f_X(x) = f_n(x) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{(x-61)^2}{10}}$$

и

$$P(50 < X < \infty) = \int_{50}^{+\infty} f_X(x)dx$$

Однако нет необходимости вычислять непосредственно этот интеграл. Если функция плотности $f(x)$ является функцией нормального распределения $f_n(x)$, то имеется формула:

$$P(\alpha < X < \beta) = \Phi(\frac{\beta - a}{\sigma}) - \Phi(\frac{\alpha - a}{\sigma}) \qquad (7.19)$$

Здесь

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt \qquad (7.20)$$

и соответствует «половинке» «колокола» функции плотности нормального распределения. Для функции $\Phi(x)$ на рис. 51 выполняется свойство

$$\Phi(-x) = -\Phi(x)$$

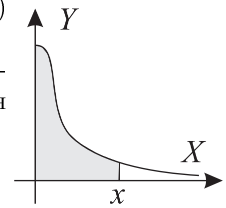


рис. 51.

Применим выписанные формулы к задаче.

$$P(50 < X < +\infty) = \Phi(\frac{\infty - 61}{5}) - \Phi(\frac{50 - 61}{5}) =$$

$$= \Phi(\infty) - \Phi(-2.2) = \Phi(\infty) + \Phi(2.2)$$

Значения функции Лапласа можно найти по таблицам, которые обычно прилагаются в конце учебников и задачников по теории вероятностей или с помощью интернет-вопроса «таблица функции лапласа».

В нашем случае $\Phi(2.2) = 0.4861$, а табличное значение $\Phi(+\infty)$ указывется не всегда. Обычно принимается, что для всех значений $x > 3$ $\Phi(x) = 0.5$. Таким образом $\Phi(+\infty) = 0.5$. В результате получим

$$P(50 < X < +\infty) = 0.4861 + 0.5 = 0.9861$$

Ответ: с вероятностью 0.9861 за кандидата проголосует не менее 50% избирателей.

Пример 6. Случайная величина X – ожидаемая эффективность (в млн руб.) рекламной кампании распределена по норамальному закону с $MX = 11$ и $DX = 16$. Найти интервал, симметричный относительно математического ожидания, возможной эффективности рекламной кампании, которая ожидается с вероятностью 0.8.

Решение. Из условия следует, что для случайной величины X – ожидаемая эффективность необходимо найти интервал вида $a - \delta < X < a + \delta)$, который может быть записан как $|X - a| < \delta$. В этом случае формула (7.19) упрощается и принимает вид

$$P(|X - a| < \delta) = 2\Phi(\frac{\delta}{\sigma}) \qquad (7.21)$$

В соответствии с решением предыдущей задачи и формулой (7.21), в которой $\sigma = \sqrt{DX} = \sqrt{16} = 4$ получим

$$P(|X - 11| < \delta) = 2\Phi(\frac{\delta}{4}) = 0.8, \qquad \Phi(\frac{\delta}{4}) = 0.4$$

По таблице функции Лапласа находим, что $\Phi(1.28) = 0.4$. Следовательно, $1.28 = \frac{\delta}{4}$ и $\delta = 1.28 \cdot 4 = 5.12$. Находим $a - \delta = 11 - 5.12 = 5.88$ и $a + \delta = 11 + 5.12 = 16.12$.

Ответ: с вероятностью 0.8 ожидаемая эффективность рекламной кампании будет не ниже 5.88 млн руб. и не выше 16.12 млн руб.

7.12 Закон больших чисел

В результате отдельного испытания нельзя заранее знать, какое значение примет случайная величина. Однако, оказывается, что при выполнении сравнительно широких условий, при большом числе испытаний суммарное поведение большого числа случайных величин почти утрачивает случайный характер.

Рассмотрим дискретную случайную величину X, заданную законом распределения:

X	x_1	x_2	$\ldots$	x_n
P	p_1	p_2	$\ldots$	p_n

Найдём математическое ожидание MX.

Расположим числовые значения $x_1, x_2, \ldots, x_n$ и MX случайной величины X на числовой оси.

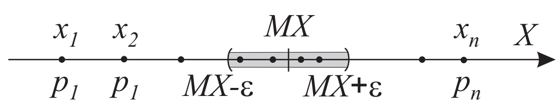


рис. 52.

Вокруг математического ожидания M отметим симметричную полоску шириной 2ε. Пусть в эту полоску попали значения $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$.

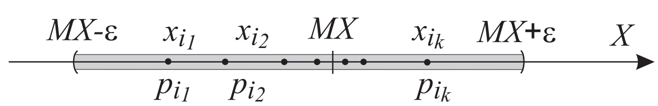


рис. 53.

Тогда вероятность того, что случайная величина оказалась внутри этой полоски равна

$$P(|X - MX| < \varepsilon) = p_{i_1} + p_{i_2} + \ldots + p_{i_k}$$

что видно из рис. 53, где отдельно изображена симметричная полоска 2ε вокруг математического ожидания MX.

Оказывается, что эта вероятность определяется шириной интервала 2ε и дисперсией DX и определяется следующей теоремой

Теорема 24. *(Неравенство Чебышева.) Вероятность того, что отклонение случайной величины X от её математического ожидания по абсолютной величине меньше положительного числа ε, не меньше, чем $1 - DX/\varepsilon^2$, то есть*

$$P(|X - MX| < \varepsilon) \geq 1 - \frac{DX}{\varepsilon^2} \tag{7.22}$$

Если в качестве ε взять утроенное с. к. о., то есть $\varepsilon = \sigma = \sqrt{DX}$, тогда

$$P(|X - MX| < \varepsilon) \geq \frac{8}{9} \tag{7.23}$$

что называется «правилом 3-х сигм» в общем случае. Фактически это правило утверждает, что с вероятностью большей, чем $8/9{=}0.88\%$ в результате испытания случайная величина примет значение, которое отклоняется от математического ожидания не более, чем на 3σ.

От рассмотрения одной случайной величины X перейдём к нескольким случайным величинам: пусть имеется n случайных величин $X_1, X_2, \ldots$, каждая из которых имеет математическое своё ожидание MX_i.

Из этих случайных величин составим новую случайную величину

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

математическое ожидание которой равно

$$M\bar{X} = \frac{MX_1 + MX_2 + \ldots + MX_n}{n}$$

Договоримся называть совокупность числовых значений $x_\alpha, \alpha \in M$ (M – произвольное множество, например, это могут быть натуральные числа $1, 2, \ldots, n$) ***равномерно ограниченной***, если существует число C, что $x_\alpha \leq C$, для всех $\alpha \in M$.

Теорема 25. *(Чебышева). Если случайные величины $X_1, X_2 \ldots, X_n$ – попарно независимы, имеют одно и то же математическое ожидание, а их дисперсии равномерно ограничены, то при любом $\varepsilon > 0$:*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n} - \frac{MX_1+MX_2+\ldots+MX_n}{n}\right| < \varepsilon\right) = 1. \tag{7.24}$$

Другими словами, теорема Чебышева утверждает, что если рассматривается достаточно большое число независимых случайных величин, имеющих ограниченные дисперсии, то почти достоверным можно считать событие, состоящее в том, что отклонение среднего арифметического случайных величин от среднего арифметического их математических ожиданий будет по абсолютной величине сколь угодно малым.

Из этой теоремы следует, что хотя отдельные независимые случайные величины могут иметь значительный разброс, тем не менее их среднее арифметическое рассеянно мало. Таким образом, нельзя уверенно предсказать, какое возможное значение примет каждая из случайных величин, но можно предвидеть, какое значение примет их среднее арифметическое.

В частном случае, когда случайные величины X_i имеют одинаковое математическое ожидание равное a, получим

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n} - a\right| < \varepsilon\right) = 1. \tag{7.25}$$

Итак, среднее арифметическое достаточно большого числа независимых случайных величин (дисперсии которых равномерно ограничены) утрачивает характер случайной величины. Объясняется это тем, что отклонения каждой из величин от своих математических ожиданий могут быть как положительными, так и отрицательными, а в среднем арифметическом они взаимно погашаются.

На теореме Чебышева основан широко применяемый в статистике выборочный метод, суть которого состоит в том, что по сравнительно небольшой случайной выборке судят о всей совокупности исследуемых объектов. Например, о качестве кипы хлопка заключают по небольшому пучку, состоящему из волокон, наудачу отобранных из разных мест кипы. Хотя число волокон в пучке значительно меньше, чем в кипе, сам пучок содержит достаточно большое количество волокон.

Пусть производится n независимых испытаний, в, каждом из которых вероятность появления некоторого события A равна p. Можно ли предвидеть, какова примерно будет относительная частота появлений события? Положительный ответ на этот вопрос дает теорема, доказанная Якобом Бернулли (опубликована в 1713 г.), которая получила название «закона больших чисел» и положила начало теории вероятностей как науки. Доказательство Бернулли было сложным; простое доказательство дано П. Л. Чебышевым в 1846 г.

Теорема 26. *(Бернулли.) Пусть μ_n – число успехов в n испытаниях Бернулли и p – вероятность успеха в отдельном испытании. Тогда при любом $\varepsilon > 0$*

$$\lim_{n\to\infty} P\left(\left|\frac{\mu_n}{n} - p\right| < \varepsilon\right) = 1.$$

Итак, при росте числа испытаний относительная частота стремится к p. Этим теорема Бернулли объясняет, почему относительная частота при достаточно большом числе испытаний оправдывает статистическое определение вероятности.

Убедимся в справедливости теоремы Бернулли.

В случае одного испытания $\mu_1 = X_1$ с законом распределения:

X_1	0	1
P	$q = 1 - p$	p

Математическое ожидание и дисперсия равны

$$MX_1 = 0 \cdot q + 1 \cdot p = p, \quad M(X_1^2) = 0^2 \cdot q + 1^2 \cdot p = p$$

$$DX_1 = p - p^2 = p(1 - q) = pq \leq 1$$

Рассмотрим $\mu_n = X_1 + X_2 + \ldots + X_n$. Например, при подбрасывании игрального кубика с цифрами от единицы до шести на гранях, вероятность, скажем, числа «три» равна $p = 1/6$. Тогда $q = 1 - 1/6 = 5/6$. При, например, пяти подбрасываниях $n = 5$, допустим число «три» появилось два раза, то есть $k = 2$. Это может быть, например, в случае $\mu_5 = 0 + 0 + 1 + 0 + 1$.

Тогда число $\mu_5/5 = 2/5$ – это относительная частота появления числа «три». Таким образом μ_n/n является относительной частотой. Так как дисперсия суммы независимых случайных величин равна сумме дисперсий

$$D(X_1 + X_2 + \ldots + X_n) = D(X_1) + D(X_2) + \ldots + D(X_n)$$

а коэффициент для дисперсии выносится в квадрате

$$D(kX) = k^2 DX$$

то в нашем случае

$$D(\mu_n) = D(X_1 + X_2 + \ldots + X_n) = npq$$

и

$$D(\frac{\mu_n}{n}) = \frac{D(\mu_n)}{n^2} = \frac{pq}{n} \to_{n\to\infty} 0$$

Откуда из неравенства Чебышева следует, что при бесконечно большом числе испытаний $n \to \infty$

$$P(|\frac{\mu_n}{n} - p| < \varepsilon) \geq_{n\to\infty} 1$$

или

$$P(|\frac{\mu_n}{n} - p| < \varepsilon)_{n\to\infty} = 1$$

для любого положительного ε.

Таким образом теорема Бернулли является следствием неравенства Чебышева.

В заключении отметим, что приведённые теоремы приведены на примере дискретной случайной величины. Но все они, что можно показать, справедливы и для непрерывной случайной величины.

7.13 Центральная предельная теорема

Пусть независимые случайные величины $X_1, X_2, \ldots, X_n$ распределены равномерно. Возникает вопрос: будет ли иметь равномерное распределение их среднеарифметическое? Оказывается, что нет. Чтобы понять причину этого, казалось бы невероятного факта, рассмотрим дискретные случайные величины, принимающие значения $1, 2, 3$ с одинаковой вероятностью $p = 1/3$.

Случайная величина $\bar{X} = (X_1 + X_2 + \ldots + X_n)/n$ будет иметь уже следующие законы распределения:
при $n = 1$

ξ	1	2	3
P	1/3	1/3	1/3

при $n = 2$

ξ	1	1.5	2	2.5	3
P	1/9	2/9	3/9	2/9	1/9

при $n = 3$

ξ	1	4/3	5/3	2	7/3	8/3	3
P	1/27	3/27	6/27	7/27	6/27	3/27	1/27

Теперь по найденным законам распределения для $n = 1, 2, 3$ построим полигоны частот.

Сначала построим для $n = 1$. По точкам полигона частот проведена огибающая линия. Видим, что она совпадает с плотностью функции равномерного распределения.

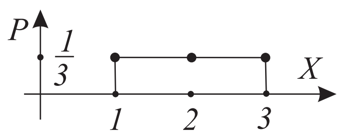


рис. 54.

Построим соответствующие полигоны частот для $n = 2$ и $n = 3$.

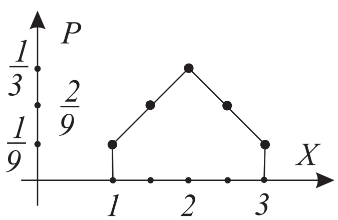


рис. 55.

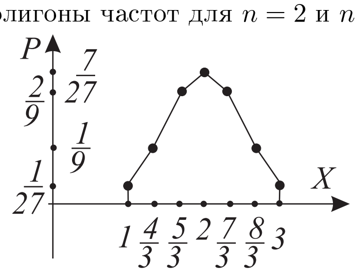


рис. 56.

Таким образом закон равномерного распределения нарушается. Очевидно это связано с тем, среднее арифметическое случайных величин вблизи математического ожидания может принимать значения большим числом комбинаций отдельных случайных величин, поэтому вероятность среднего арифметического становится доминирующей. Вероятность же того, что случайная величина примет значение на концах промежутка в силу независимости случайных величин равна произведению вероятностей этого же события для каждой случайной величины, поэтому результирующая вероятность оказывается близкой к нулю.

То, как изменяется распределение среднего арифметического n случайных величин с ростом n наглядно видно, если таблицу распределения изобразить графически:

Возникает вопрос: можно ли указать функцию плотности распределения, которая наиболее близко проходила бы по точкам рисунка. Ответ на этот вопрос был дан выдющимся русским математиком А. М. Ляпуновым в доказанной им центральной предельной теореме), которую можно сформулировать следующим образом: *если случайная величина X представляет собой сумму очень большого числа взаимно независимых случайных величин, влияние каждой из которых на всю сумму ничтожно мало, то X имеет распределение, близкое к нормальному.*

При построении функции распределения случайной величины $\bar{X} = (X_1 + X_2 + \ldots + X_n)/n$ мы воспользовались равномерным закон распределения каждой случайной величины X_i.

Подобные рассуждения применимы и к другим распределениям.

Из вида графика функции $f_X(x)$ следует, что изменение параметра a – математического ожидания – не изменяет формы нормальной кривой, а приводит лишь к ее сдвигу вдоль оси OX. Если параметр σ – среднее квадратическое отклонение – возрастает, то максимальная ордината нормальной кривой убывает, а сама кривая становится более пологой, т. е. сжимается к оси OX; если σ уменьшается, то нормальная кривая становится более «остро-вершинной».

Ранее, в неравенстве Чебышева была найдена вероятность попадания в промежуток, отклоняющийся от математического ожидание не более, чем три с. к. о. Эта вероятность для произвольного закона распределения была равна 0.89. В случае нормального распределения вероятность такого события равна практически единице, то есть данное событие можно считать достоверным. В самом деле:

$$P(|\xi - a| < 3\sigma) = 0.9973.$$

Другими словами, вероятность того, что абсолютная величина отклонения превысит утроенное среднее квадратическое отклонение, очень мала и составляет 0,27% случаев. Такие события обычно считают практически невозможными. Правило, согласно которому вероятность отклонения нормально распределённой случайной величины от её математического ожидания не превосходит утроенного среднего квадратического отклонения назывется *правилом трёх сигм.* Если это правило не выполняется, то случайная величина не распределена нормально.

Глава 8

Математическая статистика

8.1 Выборка и её представление

Результатом наблюдения за случайным событием является набор данных, который называется *выборочной совокупностью* или просто *выборкой*. Каждое новое наблюдение даже за тем же случайным событием даст новый набор данных. Все мыслимые выборки набюдения за одним и тем же случайным событием называется *генеральной совокупностью*. Задачей математической статистики является указать способ получения выборки из генеральной совокупности и предложить её метод исследования.

Математически случайное событие описывается как набор признаков. Для того, чтобы выборка правильно представляла генеральную совокупность и можно было уверенно судить об интересующем признаке необходимо, чтобы она правильно её представляла. Другими словами выборка должна правильно отражать пропорции генеральной совокупности. Это требование называется *репрезентативностью выборочной совокупности*.

Пусть в результате наблюдения получено n чисел. Число n называется *объёмом выборки* и на основании закона больших чисел должно быть достаточно большим, чтобы правильно представять генеральную совукупность. Полученную совокупность n чисел необходимо подвергнуть первичному анализу, а именно, – их группировки на повторяемость. Например, выборку 20 чисел $\{2, 2, 3, 6, 4, 3, 3, 2, 3, 4, 6, 4, 4, 4, 3, 2, 2, 2, 3, 3\}$, которые представим в виде таблицы № 1:

Таблица №1

x_i	2	3	4	6
n_i	6	7	5	2

или

Таблица № 2

x_i	2	3	4	6
p_i	0.3	0.35	0.25	0.1

Наблюдаемые значения x_i называются ***выриантами***, последовательность вариант, записанных в возрастающем порядке – ***вариационным рядом***, числа n_i – ***частотами***, а числа $p_i = n_i/n$ – ***относительными частотами***, которые совпадают со статистическим определением вероятности, поэтому вторую таблицу можно назвать статистическим законом распределние случайной величины $X = \{x_1, x_2, x_3, x_4\}$.

Первичный графический анализ выборки небольшого объёма проводят с помощью *полигона* относительных частот (см. рис. 57, построенный по закону рапределения таблицы № 2.). Если же число наблюдений достаточно велико, то целесообразно нарисовать *гистограмму относительных – частот* ступенчатую функцию, состоящую из прямоугольников, основанием которой служат частичные интервалы длиной h, а высоты равны $n_i/(n \cdot h)$.

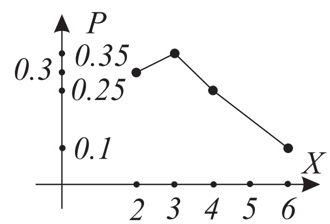


рис. 57.

Площадь i-го частичного прямоугольника равна относительной частоте p_i, где в качестве приведённой варианты выступает середина интервала прямоугольника. Полная площадь под графиком равна единице. Разность между максимальным и минимальным значениями вариант называется ***размахом выборки***, по которой находится интервал h.

Пример. Пострить гистограмму по вариационному ряду

x_i	10	15	21	25	32	35	45	48	51	60
n_i	5	7	10	5	9	12	10	5	5	7

1. Найдём объём выборки: $n = 75$. Число различных вариант равно 10 при объёме выборки 75 позволяет нам строить гистограмму относительных частот, группируя варианты на пять частей. Деление на более крупные части ($n < 5$) считается нецелесообразным. Деление на более

мелкие части в данном случае возможно, но часто приводит к малому числу вариант в отдельном интервале h, что не допустимо.

2. Найдём размах выборки $H = x_{max} - x_{min} = 60 - 10 = 50$.

3. Найдём интервал $h = H : 5 = 50 : 5 = 10$.

4. Строим *интервальный ряд частот*, группируя варианты по интервалам.

x_i	(10,20)	(20,30)	(30,40)	(40,50)	(50,60)
p_i	12	15	21	15	12

5. Строим *интервальный ряд относительных частот.*

x_i	(10,20)	(20,30)	(30,40)	(40,50)	(50,60)
$\frac{n_i}{n}$	$\frac{12}{75}$	$\frac{15}{75}$	$\frac{21}{75}$	$\frac{15}{75}$	$\frac{12}{75}$

6. Строим *интервальный ряд плотностей относительных частот.*

x_i	(10,20)	(20,30)	(30,40)	(40,50)	(50,60)
$\frac{n_i}{nh}$	$\frac{12}{750}$	$\frac{15}{750}$	$\frac{21}{750}$	$\frac{15}{750}$	$\frac{12}{750}$

7. Строим гистограмму относительных частот.

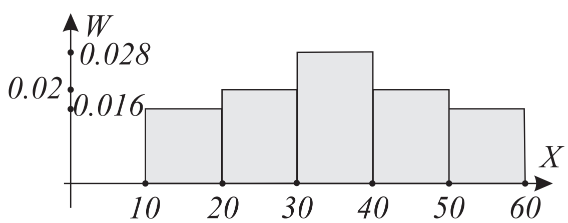


рис. 58.

Площадь каждого прямоугольника равна относительной частоте $p_i = W_i h$. Полная площадь под графиком на рис. 58 равна единице.

8.2 Оценка параметров выборки

Пусть требуется изучить количественный признак генеральной совокупности по выборке $x_1, x_2, \ldots, x_n$, которую можно представить в виде *вектора выборки* $x = (x_1, x_2, \ldots, x_n)$ в n-мерном пространстве. Например, если заранее известно, что генеральная совокупность имеет нормальное распределение, то необходимо оценить математическое ожидание и среднеквадратическое отклонение, так как эти два параметра полностью определяют нормальное распределение. На самом деле в качестве поставленной задачи может быть выяснение вида распределения, тем самым оцениваемым параметром будет некая особая величина, свойственная распределению.

Пусть Θ – параметр теоретического распределения. Например, это может быть математическим ожиданием, то есть $MX = \Theta$ или дисперсией $DX = \Theta$.

Пусть Θ^* – статистическая оценка параметра Θ теоретического распределения. Например, это может быть оценка математического ожидания или дисперсии.

Повторяя опыт мы будем получать каждый раз различные значения оценки $\Theta_1^*, \Theta_2^*, \ldots, \Theta_k^*$ параметра Θ. Если каждый раз будем получать, например, чуть большее значение, нежели правильное Θ, то математическое ожидание всех оценок Θ_i^* не будет равно Θ (что записывается как $M(\Theta^*) \neq \Theta$). По этой причине естественно потребовать, чтобы формула оцениваемого параметра была, как говорят, *несмещённой* или

$$M(\Theta^*) = \Theta$$

в противном случае оценку называют *смещённой*.

Однако ошибочно считать, что всякая несмещённая оценка всегда даёт хорошее приближение параметра распределения, так возможные значения Θ могут быть сильно рассеяны вокруг оцениваемого параметра, и соответственно большой дисперсия. Необходимо чтобы используемая оценка имела минимальную дисперсию из возможных, такая оценка называется *эффективной*.

Наконец, оценка называется *состоятельной*, если при увеличении числа наблюдений она приближается по вероятности к оцениваемому параметру. Например, если дисперсия несмещённой оценки при $n \to \infty$ стремится к нулю, то такая оценка называется и *состоятельной*.

Несмещённой оценкой генеральной средней (математического ожидания генеральной совокупности) по выборке называется число

$$\bar{x}_v = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Если же значения x_i имеют частоты n_i, причём $n_1 + \ldots + n_k = n$ то

$$\bar{x}_v = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_k x_k}{n}$$

или

$$\bar{x}_v = \frac{1}{n} \sum_{i=1}^{k} n_i x_i$$

Несмещённость, эффективность и состоятельность такой оценки математического ожидания строго доказывается методами математической статистики, тем самым являясь одним из её замечательных достижений.

Для дальнейших расчётов вместо вектора выборки переходят к *центрированному вектору выборки* $\overset{\circ}{x}_v = x_v - \bar{x}_v = (x_1 - \bar{x}_v, x_2 - \bar{x}_v, \ldots, x_n - \bar{x}_v)$.

Для оценки дисперсии генеральной совокупности используют следующую формулу *выборочной дисперсии*

$$D_v = \frac{1}{n}((x_1 - \bar{x}_v)^2 + (x_2 - \bar{x}_v)^2 + \ldots + (x_n - \bar{x}_v)^2) =$$

$$= \frac{1}{n} \sum_{i=1}^{n} (x_i - x_v)^2 = \frac{1}{n} \sum_{i=1}^{n} \overset{\circ}{x}_i^{\,2}$$

и *выборочным средним квадратичным отклонением* как оценкой среднеквадратичного отклонения генеральной совокупности называют величину

$$\sigma_v = \sqrt{D_v}$$

Для вариант с частотами выборочная средняя находится по формуле

$$D_v = \frac{n_1(x_1 - \bar{x}_v)^2 + \ldots + n_k(x_k - \bar{x}_v)^2}{n} = \frac{1}{n} \sum_{i=1}^{k} n_i (x_i - x_v)^2$$

Оказывается, что выборочная дисперсия приводит к систематическим ошибкам, поэтому используют исправленную дисперсию, обозначаемую как S или s^2

$$S = s^2 = \frac{1}{n-1} \sum_{i=1}^{k} n_i (x_i - x_v)^2$$

и исправленное среднеквадратическое отклонение

$$s = \sqrt{S}$$

Если уже найдена выборочная дисперсия D_v, то исправленная дисперсия S находится по формуле

$$S = \frac{n}{n-1} D_v$$

Из этой формулы видим, что чем больше n тем ближе исправленная к выборочной дисперсии, поэтому в практических расчётах часто используют именно выборочную дисперсию и выборочное среднеквадратическое отклонение.

8.3 Доверительный интервал и надёжность

Точечной называют оценку, которая определяется одним числом. Все оценки, которые определялись выше были точечными. *Интервальной* оценкой называют оценку, которая определяется двумя числами – концами интервала.

Пусть Θ^* – оценка параметра Θ является числом. Ясно, что чем меньше модуль отклонения $|\Theta - \Theta^*|$, тем лучше найдена оценка. Таким образом всякое положительное число $\delta > 0$ характеризует точность измерения и является *точностью оценки* в том смысле, что неравенство $|\Theta - \Theta^*| < \delta$ выполняется лучше, чем меньше δ. Но статистические методы не в состоянии категорически утверждать, что для заданного $\delta > 0$ будет выполняться неравенство $|\Theta - \Theta^*| < \delta$, зато можно говорить о вероятности, с которой это неравенство выполняется. Эту вероятность называют *надёжностью* оценки или её *доверительной вероятностью* и обозначают как γ и записывают как

$$P(|\Theta - \Theta^*| < \delta) = \gamma$$

Заменив неравенство $|\Theta - \Theta^*| < \delta$ на $|\Theta^* - \Theta| < \delta$ и равносильное ему двойное $-\delta < \Theta^* - \Theta < \delta$ получим

$$P(\Theta^* - \delta < \Theta < \Theta^* + \delta) = \gamma$$

Это выражение понимаем следующим образом: вероятность того, что интервал $(\Theta^* - \delta, \Theta^* + \delta)$ заключает в себе (покрывает) неизвестный параметр Θ равна γ.

Интервал $(\Theta^* - \delta, \Theta^* + \delta)$ называют *доверительным интервалом*, а число γ – надёжностью или доверительной вероятностью того, что интервал $(\Theta^* - \delta, \Theta^* + \delta)$ покрывает параметр Θ.

С помощью методов математической статистики предлагаются формулы нахождения интервальных оценок параметров распределения случайных величин. Эти формулы различны для различных законов распределения и зависят от объёма информации о законе распределения, если он известен.

1. Пусть количественный признак X генеральной совокупности распределён нормально, при этом **среднее квадратическое отклонение** σ **известно** из каких-то других исследований. Требуется **оценить неизвестное математическое ожидание** a по выборочной средней $\bar{x}_v$ с помощью доверительного интервала.

Будем рассматривать выборочную среднюю $\bar{x}_v$ как случайное значение случайной величины $\bar{X}$, которая изменяется от выборки к выборке и выборочные значения $x_1, x_2, \ldots, x_n$ как случайные значения одинаково распределённых независимых случайных величин $X_1, X_2, \ldots, X_n$.

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Математическое ожидание каждой из этих величин равно a, среднее квадратическое отклонение равно σ. Тогда

$$P(\bar{x}_v - t\sigma/\sqrt{n} < a < \bar{x}_v + t\sigma/\sqrt{n}) = \gamma$$

где с помощью интегральной функции Лапласа $\Phi(x)$ и доверительной вероятности γ находят число t по формуле

$$2\Phi(t) = \gamma$$

Пример. Случайная величина X имеет нормальное распределение с известным средним квадратическим отклонение $\sigma = 14$. Найти доверительный интервал для оценки математического ожидания a по выборочной средней $\bar{x}_v = 80$ с надёжностью $\gamma = 0.95$, если объём выборки равен $n = 49$.

Решение. Найдём t. Так как по условию $2\Phi(t) = 0.95$ и $\Phi(t) = 0.475$, то по таблице Приложения 1 находим $t = 1.96$. Найдём точность оценки

$$\delta = t\sigma/\sqrt{n} = 1.96 \cdot 14/\sqrt{49} = 3.92$$

Находим левую и правую границы доверительного интервала.

$$\bar{x}_v - t\sigma/\sqrt{n} = 80 - 3.92 = 76.08 \qquad \bar{x}_v + t\sigma/\sqrt{n} = 80 + 3.92 = 83.92$$

Ответ: с вероятностью 0.95 математическое ожидания случайной величины X находится в интервале $(76.08, 83.92)$.

2. Пусть количественный признак X генеральной совокупности распределён нормально, но **среднее квадратическое отклонение** σ **неизвестно**. Требуется найти интервальную оценку **неизвестного математического ожидания** a по выборке $x_1, x_2, \ldots, x_n$.

Тогда

$$P(\bar{x}_v - t_{\gamma,n}s/\sqrt{n} < a < \bar{x}_v + t_{\gamma,n}s/\sqrt{n}) = \gamma$$

где с помощью распределения Стьюдента $S(x,n)$ и доверительной вероятности γ находят число $t_{\gamma,n}$ по формуле

$$2S(t_{\gamma,n}) = \gamma$$

Пример. Случайная величина X имеет нормальное распределение. Найти доверительный интервал для оценки математического ожидания a с надёжностью $\gamma = 0.95$, по выборке $(2,4,5,3,4,2,6,4,3,5,4,3,4,2,4,5)$.

Решение. Из условия получаем, что объём выборки равен 16. Найдём выборочное среднее

$$\bar{x}_v = (2+4+5+3+4+2+6+4+3+5+4+3+4+2+4+5)/16 = 3.75$$

и исправленную дисперсию

$$S = \frac{1}{15}((2-3.75)^2 + (4-3.75)^2 + (5-3.75)^2 +$$

$$+(3-3.75)^2+(4-3.75)^2+(2-3.75)^2+(6-3.75)^2+(4-3.75)^2+(3-3.75)^2+$$

$$+(5-3.75)^2 + (4-3.75)^2 + (3-3.75)^2 + (4-3.75)^2 + (2-3.75)^2 +$$

$$+(4-3.75)^2 + (5-3.75)^2) = 1.4$$

Следовательно исправленное с. к. о. равно

$$s = \sqrt{1.4} = 1.183$$

Найдём $t_{\gamma=0.95,n}$ при $n = 16$. Так как по условию $2S(t_{\gamma,n}, n) = 0.95$ и $S(t_{\gamma,n}, n) = 0.475$, то по таблице распределения Стьюдента находим $t_{\gamma,n} = 2.13$. Найдём точность оценки

$$\delta = t_{\gamma,n} \cdot s/\sqrt{n} = 2.13 \cdot 14/\sqrt{16} = 0.63$$

Находим левую и правую границы доверительного интервала.

$$\bar{x}_v - t_{\gamma,n}s/\sqrt{n} = 3.75 - 0.63 = 3.12 \qquad \bar{x}_v + t_{\gamma,n}s/\sqrt{n} = 3.75 + 0.63 = 4.38$$

Ответ: с вероятностью 0.95 математическое ожидания случайной величины X находится в интервале $(3.12, 4.38)$.

Заметим, что во всех рассмотренных задачах стояло условие, что выборочная совокупности данных подчиняется закону нормального распределения. На самом деле это заранее никогода не известно и в общем случае такого рода утверждение является гипотезой. Но существуют методы, с помощью которых эту гипотезу можно проверить. В целом же предположение о нормальном распределении случайных величин хорошо согласуется с законом больших чисел, поэтому в случае большого числа испытаний предположение о нормальности распределения является разумным.

В этом параграфе указаны формулы для оценки математического ожидания случайных величин. Методами математической статистики получены формулы для интервальной оценки дисперсии и т. д. В виду их сложности здесь они не приводятся, но они могут быть найдены в соответствующищей литературе по математической статистике (например, в литературе в конце этого пособия).

Литература

[1] Дорофеев В. Ю. Линейная алгебра и аналитическая геометрия. СПб:Ютас, 2005.

[2] Беклемишев Д. В. Курс аналитической геометрии и линейной алгебры. Москва:Наука, 1987.

[3] Ведина О. И. и др. Математический анализ для экономистов. Москва:Филинъ, 2000.

[4] Кудрявцев Л. Д. Краткий курс математического анализа. т. 1, 1998.

[5] Банди Б. Основы линейного программирования. Москва:Радио и Связь, 1989.

[6] Феллер В. *Введение в теорию вероятностей и её приложения.* М:Мир, 1984.

[7] Гмурман В. Е. *Теории вероятностей и математическая статистика.* М:Наука, 2003.

[8] Капитоненко В.В. *Финансовая математика и её приложения.* М:ПРИОР, 1999.

[9] Калихман И. Л. Линейное алгебра и программирование. М:Высшая школа, 1967.

[10] Таха Х. Введение в исследование операций. М:Мир, 1985.

[11] Шикин Е. В. От игр к играм. Математическое введение. Москва:изд. УРСС, 2003.

[12] Дорофеев В. Ю., Савинов Г. В. «Mathematica» для линейных экономических моделей. СПб:Изд. СПбЭУ, 2018.