ПРИЛОЖЕНИЯ

Приложение 1

П1. Основные матричные операции

В этом подразделе приведены основные понятия из теории матричного исчисления. При его формировании использованы материалы книг [3, 9].

Матрицей называется таблица, содержащая в общем случае n строк и m столбцов, элементами которой могут быть числа, символы и выражения различного вида. Матрицу A можно представить как

$$A = \begin{bmatrix} a_{11} & . & . & . & a_{1m} \\ . & . & . & . & . \\ a_{i1}. & . & a_{ij} & . & a_{im} \\ . & . & . & . & . \\ a_{n1} & . & . & . & a_{nm} \end{bmatrix}. \qquad (П1.1)$$

Для матрицы могут быть использованы также следующие обозначения:

$$A = \{a_{ij}\};\ A = \{A(i, j)\};\ A = \{A[i, j]\},\ i = \overline{1.n},\ j = \overline{1.m}.$$

Вектор

$$a = \begin{bmatrix} a_1 \\ . \\ . \\ a_n \end{bmatrix} \qquad (П1.2)$$

и **строка**

$$a = [a_1, ...a_n] \qquad (П1.3)$$

представляют собой частные случаи матриц размерности $n \times 1$ и $1 \times n$ соответственно.

Квадратная матрица. Квадратной называется матрица, у которой число строк и столбцов одинаково, т.е. матрица размерности $n \times n$**.**

Диагональная матрица. Квадратная матрица A размерности n называется диагональной, если все ее элементы, кроме тех, которые стоят на диагонали, равны нулю.

$$D = \left[\begin{array}{c:c:c:c:c} d_{11} & 0 & . & 0 & 0 \\ \hdashline 0 & d_{22} & . & 0 & 0 \\ \hdashline . & . & . & . & . \\ \hdashline 0 & 0 & . & d_{n-1n-1} & 0 \\ \hdashline 0 & 0 & . & 0 & d_{nn} \end{array}\right]. \tag{П1.4}$$

Для диагональной матрицы нередко используется следующее обозначение: $D = diag\{d_{ii}\}, \quad i = \overline{1.n}$.

Единичная матрица. Единичной матрицей называется диагональная матрица с единицами на диагонали. Обычно для этой матрицы используется обозначение E, т.е.

$$E = \left[\begin{array}{c:c:c:c:c} 1 & 0 & . & 0 & 0 \\ \hdashline 0 & 1 & . & 0 & 0 \\ \hdashline . & . & . & . & . \\ \hdashline 0 & 0 & . & 1 & 0 \\ \hdashline 0 & 0 & . & 0 & 1 \end{array}\right]. \tag{П1.5}$$

Транспонированная матрица. Пусть задана матрица $A = \{a_{ij}\}, i = \overline{1.n}, j = \overline{1.m}$ размерности $n \times m$. Транспонированной матрицей $A^{\text{т}}$ называется такая матрица размерности $m \times n$, которая определяется как

$$A^{\text{т}} = \{a_{ji}\}, \; j = \overline{1.m}, i = \overline{1.n}. \tag{П1.6}$$

Очевидно, что строки такой матрицы являются столбцами матрицы A, а строки A – столбцами $A^{\text{т}}$.

Верхней треугольной матрицей называется матрица, у которой отличны от нуля элементы, стоящие на диагонали и выше:

$$A = \left[\begin{array}{c:c:c:c:c} * & * & . & * & * \\ \hdashline 0 & * & . & * & * \\ \hdashline . & . & . & . & . \\ \hdashline 0 & 0 & . & * & * \\ \hdashline 0 & 0 & . & 0 & * \end{array}\right]. \tag{П1.7}$$

У **нижней треугольной матрицы** от нуля отличны элементы, стоящие на диагонали и ниже.

Симметричная (симметрическая) матрица. Квадратная матрица размерности n называется симметричной, если

$$A = A^{\text{т}}. \qquad (\text{П}1.8)$$

Блочная матрица. Матрица, имеющая вид

$$A = \left[\begin{array}{c:c:c:c:c} A_{11} & . & . & . & A_{1m} \\ \hdashline . & . & . & . & . \\ \hdashline A_{i1} & . & A_{ij} & . & A_{im} \\ \hdashline . & . & . & . & . \\ \hdashline A_{n1} & . & . & . & A_{nm} \end{array}\right], \qquad (\text{П}1.9)$$

где A_{ij}, $i = \overline{1.n}, j = \overline{1.m}$ представляют собой матрицы, т.е. матрица, составленная из блоков, называется блочной матрицей.

Блочно-диагональная матрица – такая, у которой отличны от нуля только блоки, стоящие на диагонали.

След матрицы. Следом квадратной матрицы A называется величина, равная сумме диагональных элементов, т.е.

$$TrA = \sum_{i=1}^{n} a_{ii}. \qquad (\text{П}1.10)$$

Такое обозначение происходит от английского слова **trace**, одно из значений которого – след.

Иногда эту величину обозначают как

$$SpA = \sum_{i=1}^{n} a_{ii}. \qquad (\text{П}1.11)$$

Это обозначение соответствует немецкому варианту этого слова, которое произносится как «**шпур**».

Сложение матриц. Пусть заданы три матрицы A, B, C размерности $n \times m$. Элементы матрицы C, равной сумме двух матриц A и B

$$C = A + B, \qquad (\text{П}1.12)$$

определяются с помощью следующего выражения

$$c_{ij} = a_{ij} + b_{ij}, i = \overline{1.n}, \ j = \overline{1.m}. \qquad (\text{П}1.13)$$

Умножение матрицы на число. При умножении матрицы на число на это число умножается каждый элемент матрицы, т.е.

$$\alpha A = \{\alpha a_{ij}\}. \qquad (\text{П}1.14)$$

Умножение матриц. Пусть заданы матрицы A и B размерности $n \times m$ и $m \times l$ соответственно. Элементы $n \times l$ матрицы C, равной произведению двух матриц A и B,

$$C = AB, \qquad (\text{П}1.15)$$

определяются как $c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}, i = \overline{1.n}, \ j = \overline{1.l}$.

Из приведенного определения следует, что операция умножения корректна только в случае, когда число столбцов матрицы, стоящей в произведении на первом месте, равно числу строк матрицы, стоящей на втором месте, т.е. когда

$$C = AB, \tag{П1.16}$$

и имеет следующую структуру:

$$\underbrace{n\left\{\overbrace{\begin{bmatrix} c_{11} & \cdot & \cdot & c_{1m} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ c_{n1} & \cdot & \cdot & c_{nm} \end{bmatrix}}^{C}\right.}_{m} = \underbrace{n\left\{\overbrace{\begin{bmatrix} a_{11} & \cdot & \cdot & \cdot & a_{l1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \cdot & \cdot\cdot & \cdot & a_{nl} \end{bmatrix}}^{A}\right.}_{l} \underbrace{l\left\{\overbrace{\begin{bmatrix} b_{11} & \cdot & \cdot & \cdot & b_{1m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{l1} & \cdot & \cdot & \cdot & b_{lm} \end{bmatrix}}^{B}\right.}_{m}.$$

Будем говорить, что в этом случае **размерности матриц согласованы.** В результате умножения $1\times n$ строки $a^{\mathrm{т}}$ на $n\times 1$ столбец b получаем скаляр, в то время как умножение столбца b на строку a дает квадратную матрицу, т.е.

$$ba^{\mathrm{т}} = \begin{bmatrix} b_1a_1 & \cdot & \cdot & b_1a_n \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b_na_1 & \cdot & \cdot & b_na_n \end{bmatrix}. \tag{П1.17}$$

Величину

$$\|a\| = \sqrt{\sum_{i=1}^{n} a_i^2} = \sqrt{a^{\mathrm{т}}a} \tag{П1.18}$$

будем называть **нормой вектора** a или его **модулем**. С другими определениями нормы вектора, а также нормы матрицы можно ознакомиться в работе [3].

В общем случае операция умножения матриц некоммутативна, т.е.

$$AB \neq BA.$$

Если операция перестановки справедлива, то матрицы называют **коммутативными**.

Из сказанного с очевидностью следует, что $a^{\mathrm{т}}a = Sp(aa^{\mathrm{т}})$.

При выполнении операций **умножения блочных матриц** можно использовать правила для обычных матриц, т.е.

$$C=\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}=\begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \end{bmatrix}.$$

Здесь важно иметь в виду, что при выполнении этих операций размерности перемножаемых блоков должны быть согласованы.

Определитель или **детерминант** квадратной матрицы $A=\{a_{ij}\}$ размерности *n* может быть вычислен с использованием определителей для матриц размерности $n-1$ с помощью следующего выражения:

$$\det(A)=\sum_{j=1}^{n} a_{1j}A_{1j}, \qquad (П1.19)$$

в котором A_{1j} – **алгебраические дополнения**, т.е. определители для матриц порядка $n-1$, получаемых вычеркиванием 1-й строки и j-го столбца у матрицы *A*, умноженные на $(-1)^{1+j}$.

Аналогичное представление может быть получено для произвольной *i*-й строки

$$\det(A)=\sum_{j=1}^{n} a_{ij}A_{ij},$$

в котором A_{ij} – определители порядка $n-1$, получаемые вычеркиванием i-й строки и j-го столбца и умножением на $(-1)^{i+j}$.

К примеру, для двухмерной матрицы легко получить

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

Для определителя также используется обозначение

$$\det(A)=|A|.$$

Квадратная матрица называется **невырожденной** или **несингулярной**, если ее определитель отличен от нуля, т.е.

$$\det(A)=|A|\neq 0.$$

В противном случае матрица называется **вырожденной** или **сингулярной**.

Введем сокращенное обозначение для определителей, составленных из элементов прямоугольной *n*×*m* матрицы

$$A=\begin{pmatrix} i_1\ i_2 \ldots i_p \\ k_1\, k_2 \ldots k_p \end{pmatrix}=\begin{vmatrix} a_{i_1k_1} & a_{i_1k_2} & \ldots & a_{i_1k_p} \\ a_{i_2k_1} & a_{i_2k_2} & \ldots & a_{i_2k_p} \\ . & . & . & . \\ a_{i_pk_1} & a_{i_pk_2} & \ldots & a_{i_pk_p} \end{vmatrix}.$$

Такой определитель называется **минором** p-го порядка, если $1 \le i_1 < i_2 < ... < i_p \le n$ и $1 \le k_1 < k_2 < ... < k_p \le m$. Миноры, у которых $i_1 = k_1, i_2 = k_2, ..i_p = k_p$, называются **главными**.

Наибольший из порядков, отличных от нуля миноров, порождаемых матрицей, называется **рангом матрицы**. Если r – ранг прямоугольной $n \times m$ матрицы, то очевидно, что $r < n, m$.

Обратная матрица. Пусть задана невырожденная $n \times n$ матрица $A = \{a_{ij}\}$. Обратной матрицей A^{-1} называется такая, для которой справедливо следующее соотношение:

$$AA^{-1} = E. \tag{П1.20}$$

Обозначим обратную матрицу как $A^{-1} = B = \{b_{ij}\}$.

Для элементов обратной матрицы справедливо соотношение [3, с.345]

$$b_{ij} = \frac{A_{ji}}{\det(A)},\ \det(A) = a_{11}A_{11} + .. + a_{1n}A_{1n}, \tag{П1.21}$$

в котором A_{ji} – алгебраические дополнения.

К примеру, для двухмерной матрицы легко получить

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Заметим, что для обратной матрицы справедливо равенство $A^{-1}A = E$, т.е. прямая и обратная матрицы коммутативны.

Ортогональная матрица. Квадратная матрица A называется ортогональной при условии, что

$$AA^{\mathrm{т}} = E. \tag{П1.22}$$

Характеристическим многочленом квадратной матрицы A называется многочлен

$$p(\lambda) = \det(A - \lambda E). \tag{П1.23}$$

К примеру, для двухмерной матрицы этот многочлен имеет вид

$$p(\lambda) = \det(A - \lambda E) = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}.$$

Характеристическое уравнение определяется как

$$p(\lambda) = \det(A - \lambda E) = 0.$$

Собственные значения матрицы. Собственными значениями квадратной матрицы A размерности n называются n корней ее характеристического уравнения

$$p(\lambda) = \det(A - \lambda E) = 0. \tag{П1.24}$$

Совокупность всех n собственных чисел называется **спектром матрицы.**

Собственным вектором квадратной матрицы A называется такой вектор x, для которого выполняется следующее равенство:

$$Ax = \lambda x, \tag{П1.25}$$

т.е. $(A - \lambda E)x = 0$, где λ – собственное число.

Это равенство означает, что умножение матрицы на такой вектор не меняет его направления. Поскольку характеристическое уравнение матрицы размерности n имеет n корней, соответственно для каждого корня (собственного числа) будет свой собственный вектор. Таким образом, матрица имеет n собственных чисел и собственных векторов. Существенно, что некоторые из собственных чисел могут между собой совпадать, а одному и тому же собственному значению может соответствовать несколько различных собственных векторов. Задача их нахождения известна как **проблема собственных чисел.**

Преобразование подобия. Пусть заданы две квадратные $n\times n$ матрицы A и C, причем матрица C не вырождена. Определим квадратную матрицу B как

$$B = CAC^{-1}. \tag{П1.26}$$

Матрицы B и A называются **подобными матрицами**, а матрица C **преобразованием подобия**. Нетрудно убедиться в том, что подобные матрицы имеют одинаковые характеристические многочлены. Действительно,

$$\det(B - \lambda E) = \det(CAC^{-1} - \lambda CC^{-1}) = \det(A - \lambda E)\det(C)\det(C^{-1}) =$$

$$= \det(A - \lambda E)\det(CC^{-1}) = \det(A - \lambda E).$$

Отсюда вытекает, что подобные матрицы имеют одинаковые собственные числа.

Легко убедиться в том, что определитель и след квадратной матрицы A могут быть представлены с использованием собственных чисел, т.е.

$$TrA = \sum_{i=1}^{n} \lambda_i, \tag{П1.27}$$

$$\det(A) = \prod_{i=1}^{n} \lambda_i. \tag{П1.28}$$

Отсюда следует, что определители и следы подобных матриц одинаковы.

Диагонализация симметричных матриц. Симметричная матрица A всегда с помощью ортогонального преобразования подобия может быть приведена к диагональному виду, т.е. всегда можно найти такую ортогональную матрицу

$$T^{\mathrm{т}}T = E,\ T^{\mathrm{т}} = \begin{bmatrix} t_1 & \dots & t_j & \dots & t_n \end{bmatrix}, \tag{П1.29}$$

для которой

$$TAT^{\mathrm{т}} = \left\{ \lambda_j \delta_{ij} \right\}, \quad j = \overline{1.n}, \tag{П1.30}$$

где λ_j, t_j, $j=\overline{1.n}$ – собственные числа и собственные векторы матрицы *A*, т.е.

$$A\,t_j = \lambda_j t_j, \tag{П1.31}$$

причем $t_i^{\text{т}} t_j = \delta_{ij}$. Здесь величина δ_{ij} представляет собой **символ Кронекера**

$$\delta_{ij} = \begin{cases} 1, & i=j, \\ 0, & i \neq j. \end{cases} \tag{П1.32}$$

Последнее условие означает дополнительное требование, накладываемое на значение модулей собственных векторов, т.е. их модули должны быть равны единице.

Важно заметить, что возможность преобразовать матрицу к диагональному виду в общем случае, когда матрица не является симметричной, существует не всегда. Однако произвольную квадратную матрицу можно преобразовать к другим матрицам специального (говорят еще – нормального или канонического) вида, таким, например, как матрица Жордана, матрица Фробениуса и т.д. [9].

Квадратичная форма. Пусть заданы квадратная матрица A и вектор x размерности n. Квадратичной формой называется выражение

$$y = x^{\text{т}} A x. \tag{П1.33}$$

Квадратичная форма называется **положительно определенной** в случае, если при любых отличных от нуля значениях вектора x величина квадратичной формы положительна, т.е.

$$y = x^{\text{т}} A x > 0, \text{при } x \neq 0. \tag{П1.34}$$

Матрицу A в этом случае называют **положительно определенной.**

Если знак строгого неравенства заменить на $\geq$, то тогда квадратичная форма и соответствующая ей матрица будут называться **неотрицательно определенными**.

Если знак в неравенстве заменить на противоположный, то получим **отрицательно определенную** квадратичную форму и матрицу.

Матричные неравенства. Говорят, что квадратная матрица A больше другой квадратной матрицы B при условии, если неравенства

$$x^{\text{т}} A x > x^{\text{т}} B x > 0 \tag{П1.35}$$

или $x^{\text{т}}(A-B)x > 0$ справедливы при отличных от нуля значениях вектора x. Аналогично можно ввести и другие типы неравенств.

Можно показать, что в случае, когда выполняется какое-то неравенство для матриц, то такое же неравенство будет справедливо и для их собственных чисел. Так, если все собственные числа квадратной матрицы A положительны или неотрицательны (отрицательны или не положительны), т.е. $\lambda_j > 0$ или $\lambda_j \geq 0$ ($\lambda_j < 0$ или $\lambda_j \leq 0$), $j=\overline{1.n}$, то соответствующие неравенства будут справедливы и для самой матрицы, т.е. $A>0$ или $A\geq 0$ ($A<0$ или $A\leq 0$).

Функции от матриц. Некоторые функции от квадратных матриц могут быть введены с помощью степенных рядов. Матрица A^m степени m определяется как

$$A^m = \underbrace{A\cdot A...\cdot A}_{m}. \tag{П1.36}$$

Используя это определение, можно ввести, например, **матричную экспоненту**, задаваемую в виде

$$e^A = \exp(A) = \sum_{i=0}^{\infty}\frac{A^i}{i!}. \tag{П1.37}$$

Функции $f(A)$ от симметричных матриц могут быть введены и иначе. Для этого сначала матрицу A с помощью ортогонального преобразования сводят к диагональному виду, т.е. получают представление

$$TAT^{\mathrm{T}} = \Lambda = diag\left\{\lambda_j\right\}, \quad j=\overline{1.n},$$

из которого следует, что $A=T^{\mathrm{T}}\Lambda T$.

Далее $f(A)$ определяют как

$$f(A)=T^{\mathrm{T}}f(\Lambda)T, \tag{П1.38}$$

где $f(\Lambda)=diag\left\{f(\lambda_j)\right\}, \quad j=\overline{1.n}$.

Можно показать, что введенное таким образом определение будет совпадать с определением с помощью разложения в ряд [9].

Производная и интеграл матрицы, элементы которой зависят, например, от времени, т.е. $A(t)=\left\{a_{ij}(t)\right\}, i=\overline{1.n}, j=\overline{1.m}$, представляют собой матрицу $\dot{A}(t)=\left\{\dot{a}_{ij}(t)\right\}, i=\overline{1.n}, j=\overline{1.m}$ либо $\int_0^t A(\tau)d\tau=\left\{\int_0^t a_{ij}(\tau)d\tau\right\}, i=\overline{1.n}, j=\overline{1.m}$ $\int_0^t A(\tau)d\tau=\left\{\int_0^t a_{ij}(\tau)d\tau\right\}, i=\overline{1.n}, j=\overline{1.m}$ с элементами в виде производной или интеграла от элементов исходной матрицы.

Можно также ввести **производную скалярной функции** $s(x)$ **по векторному аргументу**

$$\frac{ds(x)}{dx^{\mathrm{T}}}=\left[\frac{\partial s(x)}{\partial x_1}...\frac{\partial s(x)}{\partial x_n}\right]. \tag{П1.39}$$

Запись $\dfrac{ds(x)}{dx}$ будет означать

$$\frac{ds(x)}{dx}=\begin{bmatrix}\frac{\partial s(x)}{\partial x_1}\\ \dots\dots\\ \frac{\partial s(x)}{\partial x_n}\end{bmatrix}, \tag{П1.40}$$

таким образом,

$$\frac{ds(x)}{dx}=\left[\frac{ds(x)}{dx^{\text{т}}}\right]^{\text{т}}. \tag{П1.41}$$

Отсюда следует, что

$$\frac{d^2s(x)}{dxdx^{\text{т}}}=\frac{d}{dx}\left[\frac{ds(x)}{dx^{\text{т}}}\right]=\begin{bmatrix}\frac{\partial^2 s(x)}{\partial x_1\partial x_1}\dots\frac{\partial^2 s(x)}{\partial x_1\partial x_n}\\ \dots\dots\dots\dots\dots\dots\\ \frac{\partial^2 s(x)}{\partial x_n\partial x_1}\dots\frac{\partial^2 s(x)}{\partial x_n\partial x_n}\end{bmatrix},\ \frac{d^2s(x)}{dxdx^{\text{T}}}=\left[\frac{d^2s(x)}{dx^{\text{T}}dx}\right]^{\text{т}}. \tag{П1.42}$$

Вектор-функцию векторного аргумента и ее производную по векторному аргументу $x=(x_{1,}\dots x_n)^{\text{т}}$ определим как

$$s(x)=\begin{bmatrix}s_1(x_1,\dots x_n)\\ .\\ .\\ s_m(x_1,\dots x_n)\end{bmatrix}. \tag{П1.43}$$

и

$$\frac{ds(x)}{dx^{\text{т}}}=\begin{bmatrix}\frac{\partial s_1(x)}{\partial x_1}\dots\frac{\partial s_1(x)}{\partial x_n}\\ \dots\dots\dots\dots\dots\dots\\ \frac{\partial s_m(x)}{\partial x_1}\dots\frac{\partial s_m(x)}{\partial x_n}\end{bmatrix},\text{ а }\frac{ds^{\text{т}}(x)}{dx}=\begin{bmatrix}\frac{\partial s_1(x)}{\partial x_1}\dots\frac{\partial s_m(x)}{\partial x_1}\\ \dots\dots\dots\dots\dots\dots\\ \frac{\partial s_1(x)}{\partial x_n}\dots\frac{\partial s_m(x)}{\partial x_n}\end{bmatrix}, \tag{П1.44}$$

т.е. $\frac{ds^{\text{т}}(x)}{dx}=\left[\frac{ds(x)}{dx^{\text{т}}}\right]^{\text{т}}$.

В общем случае, если A – квадратная симметричная матрица, то $\frac{d}{dx}[g^{\text{т}}(x)Ag(x)]=2\frac{dg^{\text{т}}(x)}{dx}Ag(x)$.

Производную скалярной функции $s(A)$ **по матрице (матричный градиент)** введем как

$$\frac{ds(A)}{dA} = \left\{ \frac{ds}{da_{ij}} \right\}.$$

К примеру, для квадратной матрицы А справедливо [32, с. 23]:

$$\frac{ds(Sp(A))}{dA} = E\,,\ \frac{ds(Sp(BAC))}{dA} = B^{\mathrm{т}}C^{\mathrm{т}}\,;$$

$$\frac{ds(Sp(ABA^{\mathrm{т}}))}{dA} = A(B + B^{\mathrm{т}})\,. \qquad (\Pi1.45)$$

Формула обращения блочных матриц [18, с. 107]. Пусть

$$P = \begin{bmatrix} P^x & P^{xy} \\ (P^{xy})^{\mathrm{т}} & P^y \end{bmatrix}, \qquad (\Pi1.46)$$

где P^x, P^y, P^{xy} – $n\times n$, $m\times m$ и $n\times m$ матрицы, причем для P^x, P^y существуют обратные матрицы. В этом случае обратная матрица P^{-1} определяется как

$$P^{-1} = \begin{bmatrix} A & B \\ B^{\mathrm{т}} & C \end{bmatrix}, \qquad (\Pi1.47)$$

где

$$A = \left[P^x - P^{xy}\left(P^y\right)^{-1} P^{yx} \right]^{-1},$$

$$B = -\left[P^x - P^{xy}\left(P^y\right)^{-1} P^{yx} \right]^{-1} P^{xy}\left(P^y\right)^{-1},$$

$$C = \left[P^y - P^{yx}\left(P^x\right)^{-1} P^{xy} \right]^{-1}.$$

Справедливы также следующие соотношения:

$$A = \left(P^x\right)^{-1} + \left(P^x\right)^{-1} P^{xy} C P^{yx} \left(P^x\right)^{-1},$$

$$B = -AP^{xy}\left(P^y\right)^{-1} = -\left(P^x\right)^{-1} P^{xy} C\,,$$

$$C = \left(P^y\right)^{-1} + \left(P^y\right)^{-1} P^{yx} A P^{xy} \left(P^y\right)^{-1}.$$

В табл. П1.1 представлены некоторые полезные матричные соотношения, в частности, так называемая **лемма об обращении матриц** [18, с. 216]

$$\left[P^{-1} + H^{\mathrm{т}} R^{-1} H \right]^{-1} = P - PH^{\mathrm{т}}\left(HPH^{\mathrm{т}} + R \right)^{-1} HP\,, \qquad (\Pi1.48)$$

где P, R – квадратные невырожденные матрицы размерности n и m; H – $m\times n$ матрица.

Из этой леммы следует, что

$$\left[P^{-1}+R^{-1}\right]^{-1}=P-P\left(P+R\right)^{-1}P, \qquad (П1.49)$$

а также

$$\left[r^2E+q^2I\right]^{-1}=\frac{1}{r^2}\left(E-\frac{q^2}{nq^2+r^2}I\right), \qquad (П1.50)$$

Последнее соотношение легко получается, если принять

$$P^{-1}=r^2E,\ H=\left[1,1...1\right],\ R^{-1}=q^2.$$

Таблица П1.1

Некоторые полезные матричные соотношения

Соотношение	Примечание
$(AB)^{\text{т}}=B^{\text{т}}A^{\text{т}}$ (П1.51)	A, B – произвольные матрицы согласованной размерности
$Tr(ABC)=Tr(BCA)=Tr(CAB)$ (П1.52)	A, B, C – произвольные матрицы согласованной размерности
$Tr(A)=\sum_{i=1}^{n}\lambda_i$ (П1.53)	A – квадратная матрица, λ_i – собственные числа A
$\det(AB)=\det(A)\det(B)$ (П1.54)	A, B – квадратные матрицы одинаковой размерности
$\det(A)=\prod_{i=1}^{n}\lambda_i$ (П1.55)	A – квадратная матрица, λ_i – собственные числа A
$(AB)^{-1}=B^{-1}A^{-1}$ (П1.56)	A, B – квадратные невырожденные матрицы одинаковой размерности
$\left[P^{-1}+H^TR^{-1}H\right]^{-1}=$ $=P-PH^T\left(HPH^T+R\right)^{-1}HP$ (П1.57)	P, R – квадратные невырожденные матрицы размерности n и m, H – $m\times n$ матрица
$\left[P^{-1}+R^{-1}\right]^{-1}=P-P\left(P+R\right)^{-1}P$ (П1.58)	P, R – квадратные невырожденные матрицы одинаковой размерности
$\left[r^2E+q^2I\right]^{-1}=\frac{1}{r^2}\left(E-\frac{q^2}{nq^2+r^2}I\right)$ (П1.59)	I – n-мерная квадратная матрица, составленная из единиц; r^2, q^2 – положительные величины
$\frac{\partial}{\partial x^{\text{т}}}(Ax)=A,\ \frac{\partial}{\partial x}(x^{\text{т}}A^{\text{т}})=A^{\text{т}}$ (П1.60)	A – квадратная симметричная матрица,

$\frac{\partial}{\partial x}(x^{\text{т}}Ax)=2Ax$, $\frac{\partial}{\partial x^{\text{т}}}(x^{\text{т}}Ax)=2x^{\text{т}}A$ (П1.61)	$x=(x_1,...x_n)^{\text{т}}$ – n-мерный вектор
$\frac{\partial^2}{\partial x\partial x^{\text{т}}}(x^{\text{т}}Ax)=2A$ (П1.62)	
$\frac{\partial}{\partial x}[g^{\text{т}}(x)Ag(x)]=2\frac{\partial g^{\text{т}}(x)}{\partial x}Ag(x)$, в частности, $\frac{\partial}{\partial x}[(y-Hx)^{\text{т}}A(y-Hx)]=$ $=2H^{\text{т}}A(y-Hx)]$ (П1.63)	A – квадратная симметричная матрица, $g(x)=(g_1(x_1,...x_n),...g_m(x_1,...x_n))^{\text{т}}$ – m-мерная вектор-функция, x, y – n- и m-мерный векторы, H – $m\times n$-матрица
$x^{\text{т}}Ax-2x^{\text{т}}z=$ $=(x-A^{-1}z)A(x-A^{-1}z)-z^{\text{т}}A^{-1}z$ (П1.64)	A – квадратная невырожденная симметричная матрица, x,z – n-мерные векторы
Задана $P=\begin{bmatrix} P^x & P^{xy} \\ (P^{xy})^{\text{т}} & P^y \end{bmatrix}$ Найти $P^{-1}=\begin{bmatrix} A & B \\ B^{\text{т}} & C \end{bmatrix}$ (П1.65)	$C=\left[P^y-P^{yx}\left(P^x\right)^{-1}P^{xy}\right]^{-1}$ $B=-\left[P^x-P^{xy}\left(P^y\right)^{-1}P^{yx}\right]^{-1}P^{xy}\left(P^y\right)^{-1}$ $A=\left[P^x-P^{xy}\left(P^y\right)^{-1}P^{yx}\right]^{-1}$

Приложение 2

П2. Случайные величины и векторы

П2.1. Случайная величина и ее описание

Случайной называется такая величина, значение которой заранее неизвестно, и можно лишь указать некую числовую меру (вероятность) того, что она будет принадлежать той или иной заранее определенной области значений [13, 14, 16, 25, 32].

Случайную величину будем считать заданной, если определена функция, позволяющая вычислять вероятность того, что случайная величина будет принадлежать тому или иному интервалу или их произвольному набору. Такой функцией является **функция распределения вероятностей**, представляющая собой скалярную функцию $F_{\mathbf{x}}(x)$ и определяющую вероятность того, что случайная величина $\mathbf{x}$ принадлежит открытому интервалу $(-\infty,x)$, т.е. вероятность того, что $\mathbf{x}<x$. Таким образом,

$$F_{\mathbf{x}}(x)=\mathbf{Pr}(\mathbf{x}:\mathbf{x}<x). \qquad (П2.1)$$

Функция (П2.1) является неотрицательной, неубывающей, непрерывной слева функцией, удовлетворяющей условиям:

$$F_{\mathbf{x}}(-\infty) = \mathbf{Pr}(\mathbf{x} : \mathbf{x} < -\infty) = 0; \tag{П2.2}$$

$$F_{\mathbf{x}}(\infty) = \mathbf{Pr}(\mathbf{x} : \mathbf{x} < \infty) = 1. \tag{П2.3}$$

Помимо функции распределения вероятности, для описания свойств случайных величин используют также **функцию плотности распределения вероятности (ф.п.р.в.)**, определяемую как

$$f_{\mathbf{x}}(x) = \frac{dF_{\mathbf{x}}(x)}{dx}. \tag{П2.4}$$

Ф.п.р.в. вероятностей облает следующими свойствами: эта функция является неотрицательной, $f_{\mathbf{x}}(x) \ge 0$);

$$F_{\mathbf{x}}(x) = \int_{-\infty}^{x} f_{\mathbf{x}}(u)du, \tag{П2.5}$$

$$\int_{-\infty}^{\infty} f_{\mathbf{x}}(u)du = 1, \tag{П2.6}$$

$$\mathbf{Pr}(x_1 \le \mathbf{x} < x_2) = \mathbf{Pr}(\mathbf{x} : \mathbf{x} < x_2) - \mathbf{Pr}(\mathbf{x} : \mathbf{x} < x_1) = F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1) \tag{П2.7}$$

или

$$\mathbf{Pr}(x_1 \le \mathbf{x} < x_2) = F_{\mathbf{x}}(x_2) - F_{\mathbf{x}}(x_1) = \int_{x_1}^{x_2} f_{\mathbf{x}}(u)du. \tag{П2.8}$$

Принимая во внимание (П2.4), можно записать

$$f_{\mathbf{x}}(x) = \lim_{dx \to 0} \frac{F_{\mathbf{x}}(x+dx) - F_{\mathbf{x}}(x)}{dx} = \lim_{dx \to 0} \frac{\mathbf{Pr}(x \le \mathbf{x} < x+dx)}{dx};$$

из этого соотношения следует, что при малых dx

$$\mathbf{Pr}(x \le \mathbf{x} < x+dx) = F_{\mathbf{x}}(x+dx) - F_{\mathbf{x}}(x) \approx f_{\mathbf{x}}(x)dx. \tag{П2.9}$$

Помимо ф.р.в. и ф.п.р.в., для описания статистических свойств с.в. используется набор ее числовых характеристик. Основными из них являются **математическое ожидание**, **моменты**, **дисперсия**, **среднеквадратическое отклонение** (**СКО**); для СКО иногда используют термин **стандартное отклонение или стандарт**. Если речь идет об ошибке, то используется термин **среднеквадратическая ошибка,** для которой будем использовать аббревиатуру **с.к.о.** В англоязычной литературе для среднеквадратической ошибки используют термин **Root Mean Square Error (RMSE).** Связь перечисленных характеристик с ф.п.р.в. определяется соотношениями, представленными в табл. П2.1.

Случайная величина, имеющая нулевое математическое ожидание, называется **центрированной случайной величиной.**

Предположим, что с помощью некоторой в общем случае нелинейной функции $g(\bullet)$ в результате преобразования случайной величины $\mathbf{x}$ с известной ф.п.р.в. $f_{\mathbf{x}}(x)$ сформирована новая с.в. $\mathbf{y} = g(\mathbf{x})$. Справедливы следующие соотношения:

$$\overline{y} = M_{\text{y}}\{\text{y}\} = M_{\text{x}}\{g(x)\} = \int g(x) f_{\text{x}}(x)dx;$$
$$M_{\text{y}}\left\{\text{y}^n\right\} = M_{\text{x}}\left\{g^n(x)\right\} = \int g^n(x) f_{\text{x}}(x)dx.$$

Таблица П2.1

Основные числовые характеристики случайной величины

Характеристика	Определение
Математическое ожидание	$M_{\text{x}}(x) = \overline{x} = \int x f_{\text{x}}(x)dx$ (П2.10)
Момент n-го порядка	$M_{\text{x}}(x^n) = \int x^n f_{\text{x}}(x)dx$ (П2.11)
Центральный момент n-го порядка	$M_{\text{x}}(x-\overline{x})^n = \int (x-\overline{x})^n f_{\text{x}}(x)dx$ (П2.12)
Дисперсия	$D = \int (x-\overline{x})^2 f_{\text{x}}(x)dx$ (П2.13)
	$D = M_{\text{x}}x^2 - \overline{x}^2$ (П2.14)
Среднеквадратическое отклонение (СКО)	$\sigma = \left(\int (x-\overline{x})^2 f_{\text{x}}(x)dx\right)^{1/2}$ (П2.15)

Примечание. В приведенных соотношениях пределы интегрирования считаются бесконечными. В дальнейшем, когда пределы не указываются, они предполагаются бесконечными.

П2.2. Гауссовские случайные величины и их характеристики

Гауссовской или нормальной случайной величиной называется такая, для которой ф.р.в. и ф.п.р.в. имеют вид

$$F_{\text{x}}(x) = \frac{1}{(2\pi)^{1/2}\sigma} \int_{-\infty}^{x} \exp\left\{-\frac{(t-\overline{x})^2}{2\sigma^2}\right\} dt; \qquad (\text{П}2.16)$$
$$f_{\text{x}}(x) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left\{-\frac{(x-\overline{x})^2}{2\sigma^2}\right\}.$$

Эти функции называют соответственно **гауссовским (нормальным) распределением** вероятности и **гауссовской (нормальной) функцией плотности распределения** вероятностей.

Для гауссовской ф.п.р.в. используется следующее обозначение:

$$f_{\text{x}}(x) = \frac{1}{(2\pi)^{1/2}\sigma} \exp\left\{-\frac{(x-\overline{x})^2}{2\sigma^2}\right\} = N(x;\overline{x},\sigma^2). \qquad (\text{П}2.17)$$

Вид гауссовских ф.р.в. и ф.п.р.в. и их зависимость от математического ожидания и СКО показан на рис. П2.1.

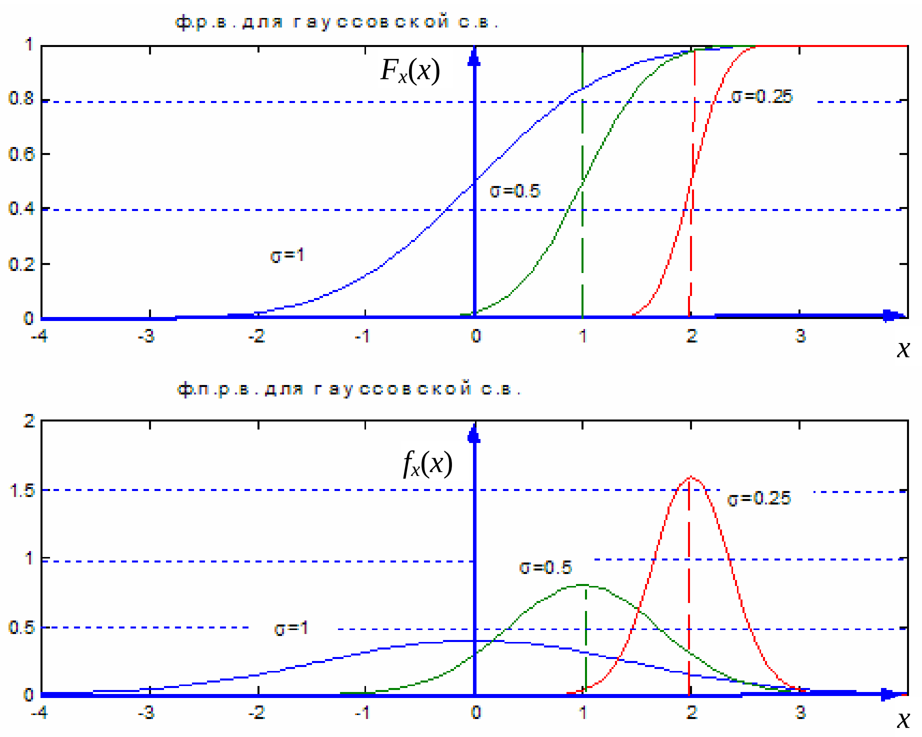

Рис. П2.1. Графики ф.р.в. и ф.п.р.в. гауссовской случайной величины при различных значениях математического ожидания ($\overline{x}=0, \overline{x}=1, \overline{x}=2$) и СКО $\sigma=1; \sigma=0,5; \sigma=0,25$

Нетрудно также заметить, что для гауссовской ф.п.р.в. медиана, математическое ожидание и мода между собой совпадают. Понятно, что такое распределение является унимодальным.

Нечетные центральные моменты гауссовской случайной величины равны нулю, т.е. [14, с.576]

$$\int (x-\overline{x})^{2k-1} f_{\mathbf{x}}(x)dx = 0,$$

а для четных моментов справедливо выражение

$$\int (x-\overline{x})^{2k} f_{\mathbf{x}}(x)dx = 1\times 3\times ..(2k-1)\sigma^{2k}, k=1,2\dots .$$

В табл. П2.2 представлены значения вероятности $\mathbf{Pr}\left[\overline{x}-k\sigma \le \mathbf{x} < \overline{x}+k\sigma\right] = \mathbf{Pr}\left[\left|\mathbf{x}-\overline{x}\right| \le k\sigma\right]$ для гауссовской с.в. при различных значениях k.

Таблица П2.2

Значения вероятности $\mathbf{Pr}\left[\left|\mathbf{x}-\overline{x}\right| \le k\sigma\right]$ для гауссовского распределения при различных k

k	1	2	3	4
$\mathbf{Pr}\left[\left|\mathbf{x}-\overline{x}\right| \le k\sigma\right]$	0,6827	0,9545	0,9973	0,9999

Обычно величину, равную 3σ, называют **предельным значением** или **предельной ошибкой,** если с.в. описывает погрешности тех или иных измерений**.** Тот факт, что для гауссовской случайной величины $\mathbf{Pr}\left[\left|\mathbf{x}-\bar{x}\right|\le 3\sigma\right]$=0,997, называют **правилом трех сигм.**

П2.3. Случайные векторы и их описание

Случайным называется вектор, каждая компонента которого является случайной величиной. Для случайного вектора $\mathbf{x}=(\mathbf{x}_1,...\mathbf{x}_n)^{\text{т}}$ его свойства в полном объеме задаются **совместной ф.р.в.** или **совместной ф.п.р.в.**, определяемыми в следующем виде:

$$F_{\mathbf{x}}(x)=\mathbf{Pr}(\mathbf{x}_1<x_1,...,\mathbf{x}_n<x_n); \qquad (П2.18)$$

$$f_{\mathbf{x}}(x)=\frac{\partial^n F_{\mathbf{x}}(x)}{\partial x_1...\partial x_n}; \qquad (П2.19)$$

$$F_{\mathbf{x}}(x)=\int\limits_{-\infty}^{x_n}...\int\limits_{-\infty}^{x_1} f_{\mathbf{x}}(u)du_1...du_n\,. \qquad (П2.20)$$

Выражение (П2.18) задает вероятность события, при котором для каждой компоненты выполняется неравенство $\mathbf{x}_j<x_j$, $j=\overline{1.n}$.

Совместная ф.п.р.в., так же, как и в одномерном случае, является неотрицательной функцией и удовлетворяет **условию нормировки**

$$\int\limits_{-\infty}^{\infty}...\int\limits_{-\infty}^{\infty} f_{\mathbf{x}}(x)dx_1...dx_n=1. \qquad (П2.21)$$

Кроме того, совместная ф.п.р.в. удовлетворяет **условиям согласованности**, которые при $m<n$ записываются как

$$f_{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_m}(x_1,x_2,...,x_m)=\int...\int f_{\mathbf{x}}(x_1,x_2,...,x_m,x_{m+1}...x_n)dx_{m+1}...dx_n\,, \quad (П2.22)$$

и является симметричной функцией своих аргументов. Последнее означает, что ф.п.р.в. для вектора $\mathbf{x}=(\mathbf{x}_1,...\mathbf{x}_n)^{\text{т}}$ не зависит от того, в какой последовательности расположены его компоненты, в частности $f_{\mathbf{x}_i,\mathbf{x}_j}(x_i,x_j)=f_{\mathbf{x}_j,\mathbf{x}_i}(x_j,x_i)$.

Вероятность попадания случайного вектора в область Ω, его математическое ожидание $\bar{x}$ и **матрица ковариаций** P, которая является обобщением понятия дисперсии на многомерный случай, имеют вид

$$\mathbf{Pr}(\mathbf{x}\in\Omega)=\int\limits_{\Omega} f_{\mathbf{x}}(x)dx; \qquad (П2.23)$$

$$\bar{x}=\int x f_{\mathbf{x}}(x)dx=M_{\mathbf{x}}(x); \qquad (П2.24)$$

$$P=\int\;(x-\bar{x})(x-\bar{x})^{\text{т}} f_{\mathbf{x}}(x)dx=M_{\mathbf{x}}(xx^{\text{т}})-\bar{x}\bar{x}^{\text{т}}\,. \qquad (П2.25)$$

Здесь и в дальнейшем интегралы понимаются как многократные, причем если область интегрирования не указана, как уже отмечалось выше, то пределы по каждой компоненте предполагаются от $-\infty$ до $+\infty$. Диагональные элементы матрицы ковариаций определяют дисперсии соответствующих компонент случайного вектора.

Математическое ожидание $M_{\mathbf{x}_i,\mathbf{x}_j}\left\{(x_i-\bar{x}_i)(x_j-\bar{x}_j)\right\}$ для двух случайных величин $\mathbf{x}_i$ и $\mathbf{x}_j$ называется **коэффициентом корреляции.** Таким образом, недиагональные элементы

$$P_{ij}=M_{\mathbf{x}_i,\mathbf{x}_j}\left\{(x_i-\bar{x}_i)(x_j-\bar{x}_j)\right\}=\int\int(x_i-\bar{x}_i)(x_j-\bar{x}_j)\,f_{\mathbf{x}_i,\mathbf{x}_j}(x_i,x_j)dx_i dx_j\,,\ i\neq j,\ i,j=\overline{1.n}$$

определяют коэффициенты корреляции между различными компонентами.

В силу свойства симметричности ф.п.р.в. справедливо равенство $P_{ij}=P_{ji}$, означающее симметричность матрицы ковариаций, т.е. $P=P^{\text{т}}$. Важным свойством матрицы ковариаций является тот факт, что она является неотрицательно определенной матрицей, т.е. такой, для которой $x^{\text{т}}Px\geq 0$ при любом $x\neq 0$.

Если математическое ожидание случайного вектора – нулевое, то, как и в скалярном случае, такой вектор называется **центрированным**.

Если $M_{\mathbf{x}_i,\mathbf{x}_j}\left\{(x_i-\bar{x}_i)(x_j-\bar{x}_j)\right\}=0$, то случайные величины называются **некоррелированными** или **ортогональными**. Отсюда следует, что для случайного вектора, у которого компоненты не коррелированы между собой, матрица ковариаций имеет диагональный вид. Если заданы два случайных вектора, то можно ввести **матрицу взаимной корреляции,** определяемую как

$$B=\int\int(x-\bar{x})(y-\bar{y})^{\text{т}}f_{\mathbf{x},\mathbf{y}}(x,y)dxdy\,,$$

где $f_{\mathbf{x},\mathbf{y}}(x,y)$ – совместная ф.п.р.в. Если эта матрица равна нулю, то говорят о том, что **случайные векторы некоррелированы или ортогональны**.

Важным также является понятие **независимости** с.в. Случайные величины $\mathbf{x}_i$, $i=\overline{1.n}$ называются взаимно **независимыми**, если совместная ф.п.р.в. равна произведению ф.п.р.в. для каждой из этих с.в., т.е.

$$f_{\mathbf{x}}(x_1...x_n)=\prod_{i=1}^{n}f_{\mathbf{x}_i}(x_i)\,.$$

Аналогично вводится определение и для **независимых случайных векторов**.

Независимые случайные величины являются некоррелированными, поскольку

$$M_{\mathbf{x}_i,\mathbf{x}_j}\left\{(x_i-\bar{x}_i)(x_j-\bar{x}_j)\right\}=\int\int(x_i-\bar{x}_i)(x_j-\bar{x}_j)\,f_{\mathbf{x}_i,\mathbf{x}_j}(x_i,x_j)dx_i dx_j=$$
$$=\int(x_i-\bar{x}_i)\,f_{\mathbf{x}_i}(x_i)dx_i\int(x_j-\bar{x}_j)f_{\mathbf{x}_j}(x_j)dx_j=0\,.$$

Обратное утверждение в общем случае несправедливо.

Приведенные выше понятия, конкретизированные для двухмерного вектора $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)^{\mathrm{т}}$, представлены в табл. П2.3.

Таблица П2.3

Основные определения и соотношения для двухмерного случайного вектора

Соотношение	Определение
Функция распределения вероятностей	$F_{\mathbf{x}}(x) = \mathbf{Pr}(\mathbf{x}_1 < x_1, \mathbf{x}_2 < x_2);$
Связь ф.п.р.в. и ф.р.в.	$f_{\mathbf{x}}(x) = \dfrac{\partial^2 F(x)}{\partial x_1 \partial x_2}$, $F_{\mathbf{x}}(x) = \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{x_2} f_{\mathbf{x}}(u) du_2 du_1$
Условие нормировки	$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f_{\mathbf{x}}(x) dx_1 dx_2 = 1$
Симметричность ф.п.р.в.	$f_{\mathbf{x}}(x_1, x_2) = f_{\mathbf{x}}(x_2, x_1)$
Условия согласованности	$f_{\mathbf{x}_1}(x_1) = \int f_{\mathbf{x}}(x_1, x_2) dx_2$ $f_{\mathbf{x}_2}(x_2) = \int f_{\mathbf{x}}(x_1, x_2) dx_1$
Независимость	$f_{\mathbf{x}}(x_1, x_2) = f_{\mathbf{x}_1}(x_1) f_{\mathbf{x}_2}(x_2)$
Некоррелированность	$M_{\mathbf{x}}(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) = 0$
Вероятность попадания случайного вектора в область Ω	$\Pr(\mathbf{x} \in \Omega) = \int\limits_{\Omega} f_{\mathbf{x}}(x) dx_1 dx_2$
Математическое ожидание	$\bar{x}_i = \int \int x_i f_{\mathbf{x}}(x_1, x_2) dx_1 dx_2 = \int x_i f_{\mathbf{x}_i}(x_i) dx_i$, $i = 1, 2$
Коэффициенты корреляции	$P_{12} = P_{21} = \iint (x_1 - \bar{x}_1)(x_2 - \bar{x}_2) f_{\mathbf{x}}(x_1, x_2) dx_1 dx_2$
Дисперсии компонент	$\sigma_i^2 = \iint (x_i - \bar{x}_i)^2 f_{\mathbf{x}_i}(x_i) dx_i$, $i = 1, 2$
Матрица ковариаций	$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$, $P_{ii} = \sigma_i^2$, $i = 1, 2$, $P = P^{\mathrm{т}} \geq 0$

П.2.4. Гауссовские случайные векторы и их характеристики

Гауссовским случайным вектором называется такой, для которого ф.п.р.в. определяется в виде

$$f_{\mathbf{x}}(x) = N(x; \bar{x}, P) = \frac{1}{(2\pi)^{n/2} (\det P)^{1/2}} \exp\left\{-0.5(x - \bar{x})^{\mathrm{т}} P^{-1} (x - \bar{x})\right\}. \quad (\text{П2.26})$$

В этом выражении $\bar{x}$ и P представляют собой математическое ожидание и матрицу ковариаций, которые, как и в одномерном случае, полностью определяют гауссовскую ф.п.р.в.

Векторы называются **совместно гауссовскими**, если их совместная плотность распределения – гауссовская. Заметим, что возможны ситуации, при которых каждый вектор или с.в. по отдельности – гауссовские, а их совместная плотность таковой не является.

Если совместно гауссовские векторы **x** и **y** не коррелированы, т.е. $\mathrm{M}_{\mathbf{x},\mathbf{y}}\{(x-\bar{x})(y-\bar{y})^{\mathrm{т}}\}=0$, то они и независимы между собой и, таким образом, $f_{\mathbf{x},\mathbf{y}}(x,y)=f_{\mathbf{y}}(y)\,f_{\mathbf{x}}(x)$.

В частности, в этом легко убедиться на примере, когда **x** и **y** – скаляры, поскольку

$$f_{\mathbf{x},\mathbf{y}}(x,y)=N\left(\begin{matrix}x\\y\end{matrix};\begin{bmatrix}\bar{x}\\\bar{y}\end{bmatrix},\begin{bmatrix}P_{11} & 0\\0 & P_{22}\end{bmatrix}\right)=\frac{1}{(2\pi)\ P_{11}^{1/2}P_{22}^{1/2}}\exp\left\{-\frac{(x-\bar{x})^2}{2P_{11}}-\frac{(y-\bar{y})^2}{2P_{22}}\right\}=$$

$$=\frac{1}{(2\pi)^{1/2}P_{11}^{1/2}}\exp\left\{-\frac{(x-\bar{x})^2}{2P_{11}}\right\}\frac{1}{(2\pi)^{1/2}P_{22}^{1/2}}\exp\left\{-\frac{(y-\bar{y})^2}{2P_{22}}\right\}.$$

Как отмечалось ранее, в общем случае такое утверждение несправедливо.

Проанализируем вид двухмерной гауссовской ф.п.р.в. Предположим, что матрица ковариаций – недиагональная, т.е.

$$P=\begin{bmatrix}\sigma_1^2 & r*\\ r* & \sigma_2^2\end{bmatrix}. \qquad (П2.27)$$

Вводя **нормированный коэффициент корреляции** в виде

$$r=\frac{r*}{\sigma_1\sigma_2}, \qquad (П2.28)$$

нетрудно убедиться в том, что

$$P^{-1}=\frac{1}{(1-r^2)}\begin{bmatrix}\dfrac{1}{\sigma_1^2} & -\dfrac{r}{\sigma_1\sigma_2}\\ -\dfrac{r}{\sigma_1\sigma_2} & \dfrac{1}{\sigma_2^2}\end{bmatrix}. \qquad (П2.29)$$

Таким образом, ф.п.р.в. для двухмерного гауссовского вектора может быть записана как

$$N(x;0,P)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}\exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{x_1^2}{\sigma_1^2}-\frac{2rx_1x_2}{\sigma_1\sigma_2}+\frac{x_2^2}{\sigma_2^2}\right)\right\}. \qquad (П2.30)$$

Уравнение

$$g(x_1,x_2)=\frac{x_1^2}{\sigma_1^2}-\frac{2rx_1x_2}{\sigma_1\sigma_2}+\frac{x_2^2}{\sigma_2^2}=c^2 \qquad (П2.31)$$

при разных значениях c задает эллипсы. Но при этом их оси развернуты относительно вертикальной оси на некоторый угол. В качестве примера на рис. П2.3 изображены изолинии при $r = 0,75$.

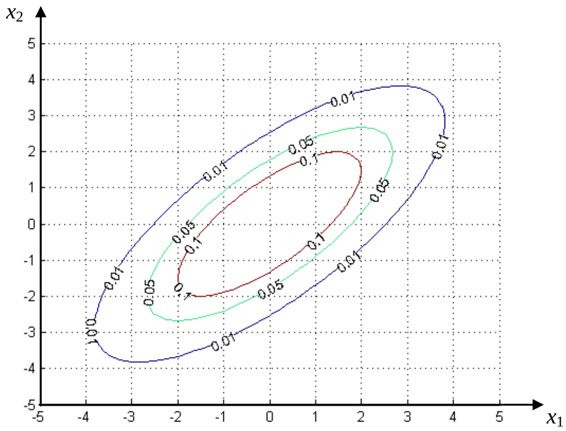

Рис. П2.3. Изолинии ф.п.р.в. для центрированного гауссовского вектора при $\sigma_1 = \sigma_2 = 1$

При решении навигационных задач на плоскости нередко полагают, что координаты объекта представляют собой гауссовский случайный вектор с математическим ожиданием в точке его предполагаемого местонахождения. Для описания неопределенности расположения точки на плоскости используют введенные выше **эллипсы равных вероятностей**, в частности эллипс, соответствующий уравнению (П2.31) при $c = 1$.

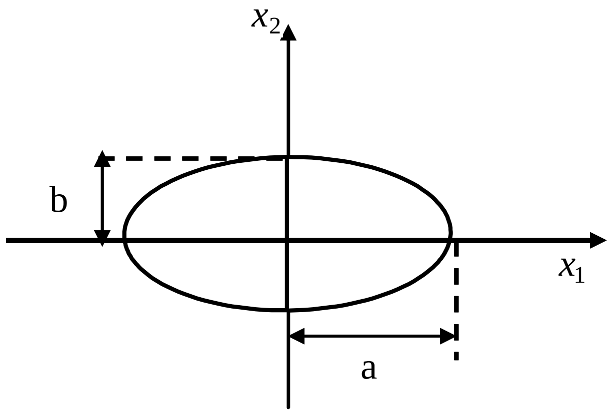

Рис. П2.4. Эллипс ошибок для двухмерного гауссовского вектора с независимыми компонентами

Поскольку этот эллипс пересекает оси в точках, совпадающих со значениями соответствующих СКО, т.е. при $x_2 = 0, x_1 = \sigma_1$, а при $x_1 = 0, x_2 = \sigma_2$ $x_1 = 0, x_2 = \sigma_2$, он получил наименование **среднеквадратического эллипса ошибок, или стандартного эллипса** [10]. В навигационных приложениях для его описания используют **параметры эллипса**: **большую** a и **малую** b **полуоси** и **дирекционный угол** τ, задающий ориентацию большой полуоси относительно оси x_2. Эти три параметра полностью определяют матрицу ковариаций двухмерной гауссовской плотности [25]. На рис. П2.4 изображен частный случай, когда $\sigma_2 = b$, $\sigma_1 = a$, $\tau = 90^\circ$, и таким образом

$$P = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix}, \qquad (П2.32)$$

т.е. размеры полуосей эллипса определяют значения СКО по каждой координате.

При оценивании точности местоположения подвижных объектов весьма важным представляется умение охарактеризовать неопределенность местоположения одним числом. Для этих целей обычно используют значения вероятности попадания точки на плоскости в ту или иную заданную область Ω. Для двухмерного центрированного гауссовского вектора с ф.п.р.в. (П2.30) эта вероятность определяется как

$$\mathbf{Pr}(\mathbf{x} \in \Omega) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \iint\limits_{\Omega} \exp\left\{-\frac{1}{2(1-r^2)} g(x_1, x_2)\right\} dx_1 dx_2, \qquad (П2.33)$$

где $g(x_1, x_2)$ – эллипс равных вероятностей (П2.31).

Если в качестве Ω выступает область, ограниченная $g(x_1, x_2)$,то, переходя к полярным координатам, можно показать, что [16, с. 68]

$$\mathbf{Pr}(\mathbf{x} : g(x_1, x_2) \le c^2) = 1 - \exp\left\{-\frac{c^2}{2(1-r^2)}\right\}. \qquad (П2.34)$$

Для случая независимых с.в. при $\sigma_1 = \sigma_2 = \sigma$ эллипс превращается в окружность радиусом $R = c\sigma$ и, таким образом, из (П2.34) получаем, что вероятность нахождения случайного вектора в круге с таким радиусом определяется распределением Рэлея

$$\mathbf{Pr}\left(\mathbf{x} : \frac{\sqrt{x_1^2 + x_2^2}}{\sigma^2} \le \frac{R^2}{\sigma^2}\right) = \mathbf{Pr}\left(\mathbf{x} : \sqrt{x_1^2 + x_2^2} \le R\right) = F(R) = 1 - \exp\left\{-\frac{R^2}{2\sigma^2}\right\},\ R>0. \qquad (П2.35)$$

Величина R, соответствующая 50-процентному попаданию гауссовского случайного вектора в круг заданного радиуса, т.е. когда вероятность попадания равна 0,5, называется **круговой вероятной ошибкой (КВО)**, а круг, соответственно, **кругом равных вероятностей.** В случае когда эллипс представляет собой окружность, т.е. при независимых с.в. и равных СКО $\sigma_1 = \sigma_2 = \sigma$, 50-процентное попадание в круг (**Pr**=0,5) достигается при его радиусе, равном 1,177σ. Для $R = 3{,}4\sigma$ **Pr**=0,997. Если это не так, то радиус круга, при котором достигается вероятность попадания в него, равная 0,5, следует отыскивать с помощью соотношения (П2.33).

Иногда используют понятие **радиальной среднеквадратической ошибки** (**Distance Root Mean Square (DRMS)**)

$$DRMS = \sqrt{\sigma_1^2 + \sigma_2^2}. \qquad (П2.36)$$

В зависимости от значений матрицы ковариаций или параметров эллипса этой величине соответствует 65–68% попадания в круг такого радиуса. Нередко используют удвоенную радиальную среднеквадратическую ошибку (2DRMS). Ей соответствует вероятность попадания в круг радиуса, равного удвоенной радиальной ошибке. Точное значение вероятности зависит от конкретных соотношений дисперсий и коэффициента корреляции, а примерная ее величина определяется как **Pr**=0,95.

Приложение 3

П3. Обыкновенные дифференциальные уравнения

Обыкновенным дифференциальным уравнением в форме Коши называется уравнение вида [6, 14, 26, 32]

$$\dot{x}(t) = F(t, x(t), u(t)), \tag{П3.1}$$

где $x(t) = (x_1(t), \ldots x_n(t))^{\text{т}}$ – n-мерный вектор; $u(t) = (u_1(t) \ldots u_p(t))^{\text{т}}$ – p-мерный вектор входного воздействия, под которым может пониматься как управление, так и возмущение; $F(.) = (F_1(.), \ldots F_n(.))^{\text{т}}$ – n-мерная в общем случае нелинейная вектор-функция. Вектор $x(t) = (x_1(t), \ldots x_n(t))^{\text{т}}$ называют **фазовым вектором**.

В тех случаях, когда $F(t, x(t), u(t)) = F(x(t), u(t))$ не зависит явно от времени, уравнение называется **стационарным**. При наличии такой зависимости уравнение называется **нестационарным**.

При $u(t) \equiv 0$ уравнение называют **однородным**, а при $u(t) \neq 0$ – **неоднородным**.

Уравнение

$$\dot{x}(t) = F(t)x(t) + G(t)u(t), \tag{П3.2}$$

в котором $F(t), G(t)$ – матрицы соответствующей размерности, называется **линейным дифференциальным уравнением**.

Если матрицы $F(t), G(t)$ зависят от времени, то уравнение называется **нестационарным линейным дифференциальным уравнением**.

Если матрицы F, G не зависят от времени, т.е.

$$\dot{x}(t) = Fx(t) + Gu(t), \tag{П3.3}$$

уравнение называется **стационарным линейным дифференциальным уравнением**.

Решением дифференциального уравнения называется такая функция времени $x(t)$, $x(t_0) = x_0$, подстановка которой в исходное дифференциальное уравнение превращает его в тождество. Значение функции $x(t_0) = x_0$ в начальный момент называется **начальным условием**.

Фундаментальной, или переходной, матрицей системы уравнений (П3.2) называется такая матрица, которая удовлетворяет уравнению

$$\frac{d\Phi(t,t_0)}{dt} = F(t)\Phi(t,t_0) \qquad (П3.4)$$

при начальном условии вида $\Phi(t_0,t_0) = E$.

Общее решение дифференциального уравнения (П3.2), удовлетворяющее начальному условию $x(t_0) = x_0$, имеет вид [6, с. 507; 14, с. 42]

$$x(t) = \Phi(t,t_0)x(t_0) + \int_{t_0}^{t} \Phi(t,\tau)G(\tau)u(\tau)\mathrm{d}\tau. \qquad (П3.5)$$

Первое слагаемое представляет общее решение линейного однородного нестационарного дифференциального уравнения

$$\dot{x}(t) = F(t)x(t).$$

Второе слагаемое является **частным решением** дифференциального уравнения (П3.2) при $x(t_0) = 0$. Частным такое решение называется потому, что оно соответствует нулевым начальным условиям и зависит от конкретного (частного) вида входного воздействия $u(t)$.

Для того чтобы показать, что (П3.5) является решением (П3.3), необходимо воспользоваться следующим правилом [14, с. 114]:

$$\frac{d}{dt}\int_{0}^{t} g(t,\tau)d\tau = g(t,t) + \int_{0}^{t} \frac{d}{dt} g(t,\tau)d\tau. \qquad (П3.6)$$

Дифференцируя (П3.8) и учитывая, что

$$\dot{\Phi}(t,t_0) = F(t)\Phi(t,t_0),$$

получаем

$$\dot{x}(t) = F(t)\Phi(t,t_0)x(t_0) + \Phi(t,t)\mathrm{G}(t)u(t) + \int_{t_0}^{t} F(t)\Phi(t,\tau)\mathrm{G}(\tau)u(\tau)\mathrm{d}\tau =$$

$$= F(t)x(t) + \mathrm{G}(t)u(t).$$

Для стационарных уравнений (П3.3) фундаментальная матрица зависит от разности аргументов и представляет собой **матричную экспоненту** для F

$$\Phi(t-t_0) = e^{F(t-t_0)} = \sum_{i=0}^{\infty} \frac{1}{i!} F^i (t-t_0)^i. \qquad (П3.7)$$

Поскольку $Fe^{F(t-t_0)} = F\sum_{i=0}^{\infty}\frac{1}{i!}F^i(t-t_0)^i$, легко убедиться в том, что $e^{F(t-t_0)}$ удовлетворяет уравнению типа (П3.4).

Таким образом, общее решение стационарного линейного дифференциального уравнения (П3.3) может быть записано как

$$x(t) = e^{F(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{F(t-\tau)}\mathrm{G}u(\tau)\mathrm{d}\tau. \qquad (П3.8)$$

Рассмотрим линейное матричное дифференциальное уравнение

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + Q(t). \quad (П3.9)$$

Нетрудно убедиться, что общим решением этого уравнения является матрица

$$P(t) = \Phi(t,t_0)P(t_0)\Phi^T(t,t_0) + \int_{t_0}^{t} \Phi(t,\tau)Q(\tau)\Phi^T(t,\tau)d\tau, \quad (П3.10)$$

где $\Phi(t,t_0)$ – фундаментальная матрица для уравнения $\dot{x}(t) = F(t)x(t)$.

Действительно, осуществляя дифференцирование этого соотношения, с учетом (П3.6) получаем

$$\dot{P}(t) = F(t)\Phi(t,t_0)P(t_0)\Phi^T(t,t_0) + \Phi(t,t_0)P(t_0)\Phi^T(t,t_0)F^T(t) +$$

$$Q(t) + \int_{t_0}^{t} F(t)\Phi(t,\tau)Q(\tau)\Phi^T(t,\tau)d\tau + \int_{t_0}^{t} \Phi(t,\tau)Q(\tau)\Phi^T(t,\tau)F(t)^T d\tau.$$

Принимая во внимание (П3.10), получаем (П3.9). Для стационарного случая решение (П3.10) имеет вид

$$P(t) = e^{F(t-t_0)}P(t_0)e^{F^T(t-t_0)} + \int_{t_0}^{t} e^{F(t-\tau)}Qe^{F^T(t-\tau)}d\tau. \quad (П3.11)$$