МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
"САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЕВА
(национальный исследовательский университет)"

С.Ю. Гоголева

АЛГЕБРА И ГЕОМЕТРИЯ. КОНСПЕКТ ЛЕКЦИЙ

Электронное учебное пособие

САМАРА 2011

УДК 512.8

Автор:Гоголева Софья Юрьевна

Гоголева, С. Ю. Алгебра и геометрия. Конспект лекций

[Электронный ресурс] : электрон. учеб. пособие / С. Ю. Гоголева; М-во образования и науки РФ, Самар. гос. аэрокосм. ун-т им. С. П. Королева (нац. исслед. ун-т). - Электрон. и граф. дан. (8,25 Мбайт). - Самара, 2011. -1 эл. опт. диск (CD-ROM).

Рассмотрены теоретические сведения по курсу "Алгебра и геометрия"и "Алгебра и геометрия (вариативная часть)". Учебное пособие рекомендуется в качестве руководства при проведении практических занятий и для самостоятельной работы.

Учебное пособие предназначено для студентов 6 факультета для бакалавров направления: 010400.62 "Прикладная математика и информатика", изучающих дисциплины "Алгебра и геометрия"в 1 семестре и "Алгебра и геометрия (вариативная часть)"во 2 и 3 семестрах.

Разработано на кафедре прикладной математики.

Комплексные числа. Определение комплексного числа.

То, что действительные числа геометрически изображаются точками координатной прямой, приводит к мысли построить систему чисел, геометрически изображаемых всеми точками координатной плоскости. Поэтому в качестве исходных элементов для построения новой числовой системы берутся всевозможные упорядоченные пары $z = (x, y)$, где $x, y \in \mathbb{R}$. Эти пары будем называть *комплексными числами*, если для неё понятие равенства и операции сложения и умножениye определены следующим образом:

1) две пары $z_1 = (x_1, y_1), z_2 = (x_2, y_2)$ считаются равными $z_1 = z_2$ тогда и только тогда, когда $x_1 = x_2, y_1 = y_2$

2) суммой двух пар $z_1 = (x_1, y_1), z_2 = (x_2, y_2)$ считается пара$(x_1, y_1)+(x_2, y_2) = (x_1+x_2, y_1+y_2)$;

3) произведением двух пар $z_1 = (x_1, y_1), z_2 = (x_2, y_2)$ считается пара$(x_1, y_1)(x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$.

Обозначение. $\mathbb{C}$ - множество комплексных чисел.

Операции сложение и умножение обладают следующими свойствами:

1. Коммутативность.
$$z_1z_2 = z_2z_1;$$
$$z_1 + z_2 = z_2 + z_1;$$
2. Ассоциативность.
$$(z_1 + z_2) + z_3 = z_1 = (z_2 + z_3);$$
$$(z_1z_2)z_3 = z_1(z_2z_3);$$
3. Дистрибутивность(Распределительный закон).
$$(z_1 + z_2)z_3 = z_1z_3 + z_2z_3.$$

Из равенства комплексных чисел для пар вида $(x, 0)$ операции выполняются также, как и для действительных чисел:
$$z_1 = (x_1, 0), \quad z_2 = (x_2, 0).$$
$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0);$$
$$(x_1, 0)(x_2, 0) = (x_1x_2, 0).$$

Определение. *Мнимой единицей* будем называть $i = (0, 1)$.
$$ii = i^2;$$
$$(0, 1)(0, 1) = (-1, 0) = -1;$$
$$i^2 = -1.$$

Определение. *Алгебраической формой* комплексного числа называется представление комплексного числа в виде $z = x + iy$, где $x, y \in \mathbb{R}$.
$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1)(y, 0).$$
x-действительная часть комплексного числа.
$$x = Rez.$$
y-мнимая часть комплексного числа.
$$y = Imz.$$

Определение. Комплексное число $x-iy$ называется *сопряженным* комплексному числу $x+iy$ и обозначается $\overline{z}$.

Свойства операции сопряжения:

1. $\overline{\overline{z}} = z$;
2. $\overline{z} = z \Leftrightarrow z \in \mathbb{R}$;
3. $z + \overline{z} = 2x, \quad \forall z = x + iy$;
4. $\overline{z} = x^2 + y^2, \quad \forall z = x + iy$;
5. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}; \quad \overline{z_1 - z2} = \overline{z_1} - \overline{z_2}$;
6. $\overline{z_1 z_2} = \overline{z_1 z_2}$;
7. $\overline{z_1/z_2} = \frac{\overline{z_1}}{\overline{z_2}}$, при $z_2 \neq 0$.

$$z_1 = x_1 + iy_1, \quad z_2 = x_2 + iy_2.$$
$$z_1 \pm z_2 = (x_1 \pm x_2) + i(y_1 \pm y_2);$$
$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = x_1x_2 - y_1y_2 + i(x_2y_1 + x_1y_2);$$
$$\frac{z_1}{z_2} = \frac{(x_1 + iy_1)}{(x_2 + iy_2)} \frac{(x_2 - iy_2)}{(x_2 - iy_2)} = \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + i\frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2};$$

Пример.
Даны два комплексных числа:

$$z_1 = 1 + 3i, \quad z_2 + 1 - 2i.$$

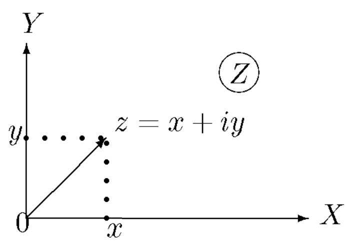
Выполнить операции сложения, вычитания, произведения и деления этих чисел.

Решение:

1. $z_1 + z_2 = 2 - i$;
2. $z_1 - z_2 = 5i$;
3. $z_1 z_2 = 7 + i$;
4. $\frac{z_1}{z_2} = \frac{1+3i}{1-2i}\frac{1+2i}{1+2i} = \frac{1-6}{1+4} + i\frac{3+2}{1+4} = -\frac{5}{5} + i\frac{5}{5} = -1 + i$;

Геометрическое представление комплексных чисел.

Рассмотрим координатную плоскость OXY, каждой точке плоскости поставим в соответствие комплексное число. Плоскость будем называть *комплексной*.



OX-действительная ось, OY-мнимая ось.

Определение.
Радиус-вектором точки z, называется вектор, начало которого находится в точке 0, а конец в данной точке z с координатами (x, y).

Определение.
Модулем комплексного числа называется длина радиус-вектора

$$|z| = \sqrt{x^2 + y^2}.$$

$\forall z \in \mathbb{C}$;

1. $|\overline{z}| = |z|$;

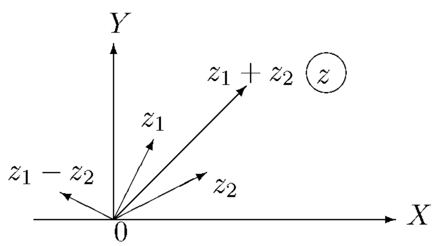
2. $z\overline{z} = |z|^2$;

3. $\forall z_1, z_2 \in \mathbb{C}, \quad |z_1 z_2|^2 = |z_1|^2 |z_2|^2$;

$$|z_1 z_2|^2 = (z_1 z_2)(\overline{z_1 z_2}) = z_1 \overline{z_1} z_2 \overline{z_2} = |z_1|^2 |z_2|^2;$$

4. $\forall z_1 \in \mathbb{C}, \ z_2 \neq 0, \quad \frac{|z_1|}{|z_2|} = \left|\frac{z_1}{z_2}\right|$;

$$|z_1| = \left|z_1 \frac{z_2}{z_2}\right| = \left|\frac{z_1}{z_2} z_2\right| = \left|\frac{z_1}{z_2}\right| |z_2|.$$



Расстояние между двумя точками равно модулю радиус-вектора разности этих векторов:

$$d(z_1, z_2) = |z_1 - z_2|.$$

Неравенство треугольника.

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

Тригонометрическая и показательная формы комплексного числа.

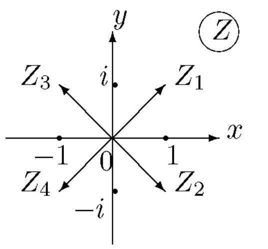
Определение. Главным аргументом комплексного числа $z \neq 0$ будем называть угол между положительным направлением действительной оси и радиус-вектором точки Z:

$$\varphi = arg\ z, \quad \varphi \in (-\pi; \pi] \,.$$

Определение. Произвольный аргумент: $Arg\ z = arg\ z + 2\pi k, \quad k \in \mathbb{Z}$.

Примеры.
Найти главный аргумент всех четырех чисел: $z_1 = 1 + i, z_2 = 1 - i, z_3 = -1 + i, z_4 = -1 - i$.

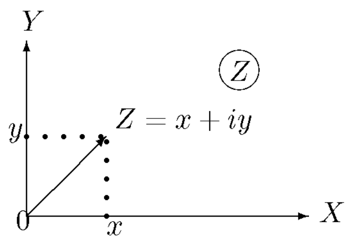
Решение: Нарисуем на комплексной плоскости радиус-векторы всех четырех чисел:



По определению главного аргумента: $\varphi_1 = \frac{\pi}{4},\ \varphi_2 = -\frac{\pi}{4},\ \varphi_3 = \frac{3\pi}{4},\ \varphi_4 = -\frac{3\pi}{4}$.

При $z = 0$ аргумент неопределен.

Рассмотрим комплексное число $z = x + iy,\ r = |z|\,,\ \varphi = arg\ z$.
Нарисуем радиус-вектор этого числа:



Очевидно, что $x = r\cos\varphi,\ y = \sin\varphi$.

Тогда $z = r\cos\varphi + ir\sin\varphi = r(\cos\varphi + i\sin\varphi)$, запись такого вида называется *тригонометрической формой* представления комплексного числа.

Умножение и деление комплексных чисел в тригонометрической форме.

$$z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1); \quad z_1 \neq 0;$$

$$z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2); \quad z_2 \neq 0;$$

$$z_1 z_2 = r_1 r_2(\cos\varphi_1 + i\sin\varphi_1)(\cos\varphi_2 + i\sin\varphi_2) =$$

$$r_1 r_2((\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i(\sin\varphi_1\cos\varphi_2 + \cos\varphi_1\sin\varphi_2)) = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2));$$

$$z_1 z_2 = r_1 r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)); \tag{4}$$

$$\frac{z_1}{z_2} = \frac{r_1(\cos\varphi_1 + \sin\varphi_1)}{r_2(\cos\varphi_2 + i\sin\varphi_2)} \frac{(\cos\varphi_2 - i\sin\varphi_2)}{(\cos\varphi_2 - i\sin\varphi_2)} =$$

$$\frac{r_1}{r_2}\left(\frac{\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2}{1} + i\frac{\sin\varphi_1\cos\varphi_2 - \cos\varphi_1\sin\varphi_2}{1}\right) = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2));$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)). \tag{5}$$

$$z = r(\cos\varphi + i\sin\varphi); \tag{6}$$

$$|z| = r = 1; \quad \varphi = arg\ z; \quad z \neq 0.$$

Формула Эйлера.

$$e^{i\varphi} = \cos\varphi + i\sin\varphi. \tag{7}$$

Заменой φ на $-\varphi$ получаем:

$$e^{-i\varphi} = \cos\varphi - i\sin\varphi.$$

$$\cos\varphi = \frac{e^{i\varphi} + e^{-i\varphi}}{2}; \quad \sin\varphi = \frac{e^{i\varphi} - e^{-i\varphi}}{2i}.$$

$$e^{i\varphi_1}e^{i\varphi_2} = e^{i(\varphi_1+\varphi_2)}; \tag{8}$$

$$\frac{e^{i\varphi_1}}{e^{i\varphi_2}} = e^{i(\varphi_1-\varphi_2)}; \tag{9}$$

$$(e^{i\varphi})^n = e^{i\varphi n}. \tag{10}$$

Формула Муавра.

$$(\cos\varphi + i\sin\varphi)^n = (\cos(n\varphi) + i\sin(n\varphi)). \tag{11}$$

Из (6) и (7) мы можем записать представление комплексного числа в *показательной форме.*

$$z = re^{i\varphi}.$$

Используя (7), (8), (9) мы можем, также как и для тригонометрического представления, записать правила нахождения частного, произведения и возведения в степень комплексных чисел в показательной форме:

$$z_1 = r_1e^{i\varphi_1}, \quad z_2 = r_2e^{i\varphi_2}.$$

1. $z_1z_2 = r_1r_2e^{i(\varphi_1+\varphi_2)}$;
2. $\frac{z_1}{z_2} = \frac{r_1}{r_2}e^{i(\varphi_1-\varphi_2)}$;
3. $(re^{i\varphi})^n = r^ne^{i\varphi n}$;
4. $z^n = r^n(\cos(n\varphi) + i\sin(n\varphi))$.

Пример.
Вычислить, используя тригонометрическую и показательную формы комплексного числа:

$$(1 - i\sqrt{3})^3(1 + i)^2.$$

Решение:

1. Вычисления в тригонометрической форме:

$$z_1 = 1 - i\sqrt{3} = 2(\cos(-\frac{\pi}{3}) + i\sin(-\frac{\pi}{3}));$$

$$z_2 = 1 + i = \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4});$$

$$z_1^3 = 8(\cos(-\pi) + i\sin(-\pi));$$

$$z_2^2 = 2(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2});$$

$$z_1^3 z_2^2 = 2^4(\cos(-\frac{\pi}{2} + i\sin(-\frac{\pi}{2}))) = -16i.$$

2. Вычисления в показательной форме:

$$z_1 = 2e^{i\frac{\pi}{3}};$$

$$z_2 = \sqrt{2}ei\frac{\pi}{4};$$

$$z_1^3 = 8e^{-i\pi};$$

$$z_2^2 = 2e^{i\frac{\pi}{2}};$$

$$z_1^3 z_2^2 = 8e^{-i\pi}2e^{i\frac{\pi}{2}} = 16e^{-\pi+\frac{\pi}{2}i} = 16e^{-1\frac{\pi}{2}} = 16(\cos(\frac{\pi}{2}) + i\sin(\frac{\pi}{2})) = -16i.$$

Извлечение корней из комплексных чисел.

$$\omega^n = Z; \quad n \in \mathbb{Z}; \quad Z \neq 0;$$
$$\omega^n = \rho^n e^{in\theta}; \quad Z = re^{i\varphi};$$
$$\rho^n e^{in\theta} = re^{i\varphi} \Leftrightarrow \rho^n = r;\ n\theta = \varphi + 2\pi k; \quad k \in \mathbb{Z};$$
$$\rho = \sqrt[n]{r}; \quad \theta = \frac{\varphi + 2\pi k}{n};$$
$$\omega_k = \sqrt[n]{r} e^{i\left(\frac{\varphi+2\pi k}{n}\right)}, \quad k \in \mathbb{Z}. \tag{11}$$

Покажем, что при нахождении корней по формуле (11) мы получаем ровно n различных корней:

$$\omega_0 = \sqrt[n]{r} e^{i\left(\frac{\varphi}{n}\right)};$$
$$\omega_1 = \sqrt[n]{r} e^{i\left(\frac{\varphi+2\pi}{n}\right)};$$
$$\vdots$$
$$\omega_{n-1} = \sqrt[n]{r} e^{i\left(\frac{\varphi+2\pi(n-1)}{n}\right)};$$
$$\omega_n = \sqrt[n]{r} e^{i\left(\frac{\varphi}{n}+2\pi\right)}.$$

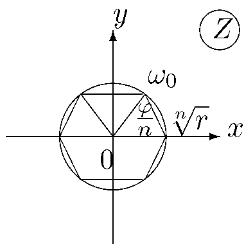
Очевидно, что первый и последний члены совпадают.

Формула нахождения корней:

$$\omega_n = \sqrt[n]{r}(\cos\left(\frac{\varphi + 2\pi k}{n}\right) + i \sin\left(\frac{\varphi + 2\pi k}{n}\right));$$
$$k = 0, 1, ..., n-1.$$

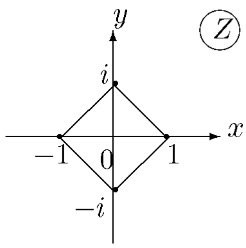
Геометрическая интерпретация.

На комплексной плоскости точки ω_k, $\quad k = 0, 1, ..., n-1$ являются вершинами правильного n-угольника, вписанного в окружность радиуса $R = \sqrt[n]{r}$ с центром в точке 0.



Пример.
Найти все значения корня $\sqrt[4]{1}$.
Решение:

$$1 = 1(\cos 0 + i \sin 0).$$
$$k = 0 \quad \omega_0 = \cos 0 + i \sin 0 = 1;$$
$$k = 1 \quad \omega_1 = \cos\frac{\pi}{2} + i \sin\frac{\pi}{2} = i;$$
$$k = 2 \quad \omega_2 = \cos\pi + i \sin\pi = -1;$$
$$k = 3 \quad \omega_3 = \cos\frac{3\pi}{2} + i \sin\frac{3\pi}{2} = -i;$$



y
Z
i
x
-1
0
1
$-i$

Понятие матрицы. Операции над матрицами

Пусть $m, n \in \mathbb{N}$. *Матрицей размера* $m \times n$ называется совокупность mn чисел, записанных в виде прямоугольной таблицы из m строк и n столбцов. При этом сами числа называются *элементами* матрицы.

Матрицу обозначают прописными латинскими буквами, при этом саму таблицу заключают в скобки (либо круглые, либо квадратные, либо двойные вертикальные):

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}; \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}; \quad \begin{Vmatrix} 1 \\ 2 \end{Vmatrix}.$$

Элементы матрицы обозначают строчными латинскими буквами, снабженными двумя индексами: a_{ij} – элемент матрицы, расположенный в i-й строке j-м столбце. В этих обозначениях матрица размера $m \times n$ в общем виде может быть записана следующим образом:

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

Используются обозначения :

$A = (a_{ij})$ – матрица A с элементами a_{ij};

$\mathbb{R}^{m \times n}$ – множество всех вещественных матриц размера $m \times n$.

Матриц размера $n \times n$ называется *квадратной матрицей* n*-го порядка.*

Квадратная матрица называется *диагональной*, если все ее внедиагональные элементы $a_{ij},\ i \neq j$ равны нулю.

Обозначение: $diag(a_{11}, \dots, a_{nn})$.

Диагональная матрица, у которой все диагональные элементы равны между собой, называется *скалярной.*

Скалярная матрица, у которой все диагональные элементы равны 1, называется *единичной.* Отметим, что для каждого порядка n существует своя единичная матрица.

Обозначение: E или I.

Матрица O, все элементы которой равны нулю, называется *нулевой.*

Квадратная матрица $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ называется *верхней (правой) треугольной*, если $a_{ij} = 0$ при $i > j$, и *нижней (левой) треугольной*, если $a_{ij} = 0$ при $i < j$.

Матрица $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ называется *верхней (правой) ступенчатой*, если она обладает следующими свойствами:

1) если i-я строка нулевая, то $(i+1)$-я строка также нулевая;

2) если первые ненулевые элементы i-й и $(i+1)$-й строк расположены в столбцах с номерами k_i и k_{i+1}, то $k_i < k_{i+1}$.

Эти свойства означают, что все нулевые строки являются последними и что все элементы, расположенные слева и под первым ненулевым элементом каждой строки, равны нулю.

Если в определении верхней ступенчатой матрицы поменять ролями строки и столбцы, то получим определение *нижней (левой)ступенчатой* матрицы.

Ступенчатая матрица, у которой $k_i = i$, называется *трапециевидной.*

Операции над матрицами. Две матрицы $A = (a_{ij})$ и $B = (b_{ij})$ одинакового размера $m \times n$ называются *равными*, если

$$a_{ij} = b_{ij}, \qquad i = 1, 2, \ldots, m, \qquad j = 1, 2, \ldots, n$$

.

Обозначение: $A = B$.

Суммой матриц $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ и $B = (b_{ij}) \in \mathbb{R}^{m\times n}$ называется матрица $C = (c_{ij}) \in \mathbb{R}^{m\times n}$, элементы которой определены равенством:

$$c_{ij} = a_{ij} + b_{ij}, \qquad i = 1, 2, \ldots, m, \qquad j = 1, 2, \ldots, n. \tag{1}$$

Обозначение: $C = A + B$.

Матрица $-A = (-a_{ij}) \in \mathbb{R}^{m\times n}$ называется *противоположной* к матрице $A = (a_{ij}) \in \mathbb{R}^{m\times n}$.

Свойства операции сложения:

$\forall\ A, B, C \in \mathbb{R}^{m\times n}$ и $O \in \mathbb{R}^{m\times n}$

1. $A + B = B + A$;
2. $(A + B) + C = A + (B + C)$;
3. $A + O = O + A = A$;
4. $A + (-A) = -A + A = O$.

Разностью матриц $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ и $B = (b_{ij}) \in \mathbb{R}^{m\times n}$ называется матрица $X = (x_{ij}) \in \mathbb{R}^{m\times n}$ такая, что $A = B + X$.

Обозначение: $X = A - B$.

Очевидно, что для $\forall\ A, B \in \mathbb{R}^{m\times n}$ существует единственная разность $A - B$, при этом

$$A - B = A + (-B) = (a_{ij} - b_{ij}).$$

Произведением матрицы $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ ***на число*** $\alpha \in \mathbb{R}$ называется матрица $C = (c_{ij}) \in \mathbb{R}^{m\times n}$, элементы которой определены равенством:

$$c_{ij} = \alpha a_{ij}, \qquad i = 1, 2, \ldots, m, \qquad j = 1, 2, \ldots, n. \tag{2}$$

Обозначение: $C = \alpha A$.

Свойства операции умножения матрицы на число:

$\forall\ A, B \in \mathbb{R}^{m\times n}$ и $\alpha, \beta \in \mathbb{R}$
1. $(\alpha\beta)A = \alpha(\beta A)$;
2. $\alpha(A+B) = \alpha A + \alpha B$;
3. $(\alpha+\beta)A = \alpha A + \beta A$;
4. $1 \cdot A = A$;
5. $-A = (-1)A$.

Произведением матриц $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ и $B = (b_{ij}) \in \mathbb{R}^{n\times k}$ называется матрица $C = (c_{ij}) \in \mathbb{R}^{m\times k}$, элементы которой определены равенством:

$$c_{ij} = \sum_{s=1}^{n} a_{is}b_{sj}, \qquad i = 1, 2, \ldots, m, \qquad j = 1, 2, \ldots, k. \tag{3}$$

Обозначение: $C = AB$.

!*Произведение AB определено лишь в том случае, когда число столбцов матрицы A равно числу строк матрицы B.*

Свойства операции умножения матриц:
1. $(AB)C = A(BC)$;
2. $\alpha(AB) = (\alpha A)B = A(\alpha B), \quad \forall \alpha \in \mathbb{R}$;
3. $A(B+C) = AB + AC, \quad (A+B)C = AC + BC$,

выполненные для любых матриц A, B, C, для которых левые части равенств имееют смысл.

Целой положительной степенью A^k $(k > 1)$ *квадратной матрицы* называется произведение k матриц, каждая из которых равна A.

Нулевой степенью квадратной матрицы A называется единичная матрица E того же порядка, что и A, т. е. $A^0 = E$.

Пусть $A = (a_{ij}) \in \mathbb{R}^{m\times n}$. Матрица $A^T = (a^t_{ij}) \in \mathbb{R}^{n\times m}$называется ***транспонированной*** к матрице A, если

$$a^t_{ij} = a_{ji}, \qquad i = 1, 2, \ldots, n, \qquad j = 1, 2, \ldots, m.$$

Переход от матрицы A к A^T называется *транспонированием матрицы A*. При транспонировании матрицы A ее строки становятся столбцами A^Tс теми же номерами, а столбцы – строками.

Свойства операции транспонирования матриц:
1. $(A+B)^T = A^T + B^T$;
2. $(\alpha A)^T = \alpha A^T, \quad \forall \alpha \in \mathbb{R}$;
3. $(AB)^T = B^T A^T$;
4. $(A^T)^T = A$,

выполненные для любых матриц A, B, для которых левые части равенств имееют смысл.

П р и м е р 1. *Найти произведение матриц* $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \end{pmatrix}$ *и* $B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & 3 \end{pmatrix}$.

<u>Решение.</u> $A \in \mathbb{R}^{2\times3}$, $B \in \mathbb{R}^{3\times2}$ $\Rightarrow$ *произведение* AB *определено (число столбцов матрицы* A *равно числу строк матрицы* B *и равно трем) и* $AB = C \in \mathbb{R}^{2\times2}$.

По формуле (3) находим

$$AB = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 1 & 3 \end{pmatrix} = C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} =$$

$$= \begin{pmatrix} 1\cdot1+2\cdot2+3\cdot1 & 1\cdot1+2\cdot0+3\cdot3 \\ 4\cdot1+5\cdot2+0\cdot1 & 4\cdot1+5\cdot0+0\cdot3 \end{pmatrix} = \begin{pmatrix} 8 & 10 \\ 14 & 4 \end{pmatrix},$$

т.е. c_{ij} – элементы матрицы C, которые получаются перемножением i-й строки матрицы A на j-й столбец матрицы B.

П р и м е р 2. *Найти значение многочлена* $f(C)$, *если* $f(x) = x^2 - 2x + 5$; $C = AB$; $A = \begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ и $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 2 & -1 \end{pmatrix}$.

<u>Решение.</u> По формуле (3) $C = AB = \begin{pmatrix} 4 & -2 \\ 4 & -1 \end{pmatrix}$; $C^2 = CC = \begin{pmatrix} 8 & -6 \\ 12 & -7 \end{pmatrix}$;

Используя формулы (1) и (3) вычисляем

$$f(C) = C^2 - 2C + 5E = \begin{pmatrix} 8 & -6 \\ 12 & -7 \end{pmatrix} - 2\begin{pmatrix} 4 & -2 \\ 4 & -1 \end{pmatrix} + 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & -2 \\ 4 & 0 \end{pmatrix}.$$

Определители. Основные методы вычисления определителей.

Упорядоченная совокупность чисел $\alpha_1, \alpha_2, \ldots, \alpha_n$, в которой

1) $\alpha_i \in \{1, 2, \ldots, n\}, \quad i = 1, 2, \ldots, n$;

2) $\alpha_i \neq \alpha_j$ при $i \neq j$

называется *перестановкой* из чисел $i = 1, 2, \ldots, n$.

Говорят, что два числа α_i и α_j в перестановке $\alpha_1, \alpha_2, \ldots, \alpha_n$ образуют *инверсию (беспорядок)*, если $\alpha_i > \alpha_j$ при $i < j$ и порядок – в противном случае. Общее число инверсий в перестановке $\alpha_1, \alpha_2, \ldots, \alpha_n$ обозначается символами $N(\alpha_1, \alpha_2, \ldots, \alpha_n)$ или $N(\alpha)$.

Определителем n***-го порядка*** квадратной матрицы $A = (a_{ij})$ называется сумма всевозможных произведений $a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n}$ элементов матрицы, взятых по одному из каждой строки и каждого столбца, причем, если сомножители в этом произведении упорядочены в порядке возрастания номеров строк, то оно берется со знаком $(-1)^{N(\alpha)}$. Для обозначения определителя приняты символы $\Delta, |A|, \det A$. Итак,

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \sum_{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)} (-1)^{N(\alpha)} a_{1\alpha_1} a_{2\alpha_2} \ldots a_{n\alpha_n}, \qquad (4)$$

где суммирование ведется по всевозможным перестановкам $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ из чисел $1, 2, \ldots, n$.

Каждое произведение в сумме (4) называется *членом определителя*, а число $(-1)^{N(\alpha)}$ – его *знаком.*

Из свойств перестановки следует, что число всевозможных членов определителя n-го порядка равно $n!$ и что при $n \geq 2$ число положительных членов равно числу отрицательных и равно $n!/2$.

Определение (4) для $n = 2$ и $n = 3$ приобретает вид

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \qquad (5)$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{aligned} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ &- a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}. \end{aligned} \qquad (6)$$

Свойства определителя.

1. Определитель треугольной матрицы равен произведению диагональных элементов.

2. Определитель квадратной матрицы не изменяется при ее транспонировании: $|A| = |A^T|$.

Следствие. В определении (4) определителя можно поменять ролями строки и столбцы:

$$|A| = \sum_{\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)} (-1)^{N(\alpha)} a_{\alpha_1 1} a_{\alpha_2 2} \ldots a_{\alpha_n n},$$

т.к. эта сумма равна $|A^T|$.

3. Если одна из строк (столбцов) матрицы целиком состоит из нулей, то ее определитель равен нулю.

4. При умножении строки (столбца) матрицы на число ее определитель умножается на это число.

5. Если каждый элемент некоторой i-й строки матрицы представлен в виде суммы:

$$a_{ik} = a'_{ik} + a''_{ik}, \quad k = 1, 2, \ldots, n,$$

то определитель матрицы можно представить в виде суммы двух определителей: $\Delta = \Delta' + \Delta''$, где

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a'_{i1} + a''_{i1} & a'_{i2} + a''_{i2} & \ldots & a'_{in} + a''_{in} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

$$\Delta' = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a'_{i1} & a'_{i2} & \ldots & a'_{in} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}, \quad \Delta'' = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a''_{i1} & a''_{i2} & \ldots & a''_{in} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

6. При перестановке местами двух строк (столбцов) матрицы ее определитель меняет знак.

7. Определитель матрицы, имеющий две одинаковые строки (столбца), равен нулю.

8. Если одна строка (столбец) матрицы является линейной комбинацией других ее строк (столбцов), то определитель матрицы равен нулю.

9. Если к какой-либо строке (столбцу) матрицы прибавить линейную комбинацию других ее строк (столбцов), то ее определитель не изменится.

Определитель k-го порядка, составленный из элементов матрицы $A \in \mathbb{R}^{m\times n}$, стоящих на пересечении строк и столбцов с номерами $i_1 < i_2 < \ldots < i_k$ и $j_1 < j_2 < \ldots < j_k$ соответственно, называется ***минором*** k-го порядка матрицы A и обозначается

$$M^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k} = \begin{vmatrix} a_{i_1 j_1} & \ldots & a_{i_1 j_k} \\ \ldots & \ldots & \ldots \\ a_{i_k j_1} & \ldots & a_{i_k j_k} \end{vmatrix}.$$

Минор порядка $n-k$, оставшийся после вычеркивания в квадратной матрице $A \in \mathbb{R}^{n\times n}$ строк и столбцов с номерами $i_1 < i_2 < \ldots < i_k$ и $j_1 < j_2 < \ldots < j_k$ соответственно, называется ***дополнительным минором к минору*** $M^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k}$ и обозначается $\overline{M}^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k}$.

Число

$$A^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k} = (-1)^{\sum_{p=1}^{k}(i_p+j_p)} \overline{M}^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k}$$

называется ***алгебраическим дополнением*** к минору $M^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k}$.

П р и м е р 3. *Дана матрица*

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

Найти алгебраическое дополнение к минору M^{13}_{34}.

<u>Решение.</u> Вычеркнем из данной матрицы 1-ю и 3-ю строки, 3-й и 4-й столбцы. Минор

$$\overline{M}^{13}_{34} = \begin{vmatrix} -3 & 2 \\ 1 & -2 \end{vmatrix}$$

является дополнительным к минору M^{13}_{34}. Алгебраическим дополнением к минору M^{13}_{34} будет

$$(-1)^{1+3+3+4}\overline{M}^{13}_{34} = -\begin{vmatrix} -3 & 2 \\ 1 & -2 \end{vmatrix} = -4.$$

Теорема Лапласа. *Пусть $A \in \mathbb{R}^{m\times n}$ и $k \in \mathbb{N}$, $1 \le k \le n-1$. Пусть в матрице A выбраны произвольные k строк (или столбцов). Тогда определитель матрицы A равен сумме всевозможных произведений миноров k-го порядка, расположенных в выбранных строках (соответственно столбцах), на их алгебраические дополнения, т.е.*

$$\det A = \sum_{(j_1, j_2, \ldots, j_k)} M^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k} A^{i_1 i_2 \ldots i_k}_{j_1 j_2 \ldots j_k}.$$

Если в теореме Лапласа выбрать $k = 1$ и строку (столбец) с номером i, то минорами первого порядка, расположенными в i-й строке (столбце), будут сами элементы $a_{ij}(a_{ji})$. Обозначив через A_{ij} алгебраическое дополнение к элементу a_{ij}, получим из теоремы Лапласа, что

$$\det A = \sum_{j=1}^{n} a_{ij}A_{ij} \quad \det A = \sum_{j=1}^{n} a_{ji}A_{ji}. \tag{7}$$

Представление определителя (7) называется ***разложением определителя по i-й строке (столбцу)***.

П р и м е р 4. *Дана матрица*

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

Найти алгебраические дополнения элементов 2-го столбца.

Решение.

$$A_{12} = (-1)^{1+2} \begin{vmatrix} 1 & 5 \\ 4 & 7 \end{vmatrix} = 13;$$

$$A_{22} = (-1)^{2+2} \begin{vmatrix} -2 & 0 \\ 4 & 7 \end{vmatrix} = -14;$$

$$A_{32} = (-1)^{3+2} \begin{vmatrix} -2 & 0 \\ 1 & 5 \end{vmatrix} = 10.$$

Теорема. *Определитель произведения квадратных матриц равен произведению определителей матриц-сомножителей*:

$$\det AB = \det A \det B.$$

Основные методы вычисления определителей.

1. Приведение к треугольному виду. Этот метод заключается в преобразовании матрицы определителя к такому виду, когда элементы, стоящие по одну сторону от главной (побочной) диагонали, равны нулю. Полученный определитель по свойству 1 равен произведению элементов главной диагонали (побочной диагонали, умноженной на $(-1)^{n(n-1)/2}$).

Для вычисления определителя таким способом используют ***метод Гаусса***, который приводит определитель n-го порядка матрицы $A = (a_{ij})$ к верхнему треугольному виду:

1. Если $a_{11} = 0$, то переставляем строки (столбцы) матрицы определителя так, чтобы элемент $a_{11} \neq 0$.

2. Умножаем 1-ю строку матрицы определителя последовательно на числа $-a_{21}/a_{11}, -a_{31}/a_{11}, \ldots, -a_{n1}/a_{11}$ и складываем со 2-ой, 3-й, $\ldots$ n-ой строками соответственно, получая нули в 1-ом столбце ниже элемента a_{11}.

3. Повторяем процедуру п.1-2, применяя ее к измененной подматрице $(n-1)$-го порядка, у которой в верхнем левом углу стоит элемент $\tilde{a}_{22}$ и так далее.

Замечание. Преобразование определителя легче производить с целыми числами, поэтому диагональный элемент, если возможно, выбирают равным единице, меняя строки (столбцы) местами или вынося общий множитель строки (столбца) за знак определителя.

П р и м е р 5. *Вычислить определитель*

$$\Delta = \begin{vmatrix} 3 & -9 & 3 & 12 \\ -1 & 5 & 2 & -3 \\ 2 & -2 & 6 & 3 \\ 0 & 2 & -1 & 2 \end{vmatrix}$$

приведением к треугольному виду.

Решение. Вынесем за знак определителя общий множитель 1-ой строки:

$$\Delta = 3 \cdot \begin{vmatrix} 1 & -3 & 1 & 4 \\ -1 & 5 & 2 & -3 \\ 2 & -2 & 6 & 3 \\ 0 & 2 & -1 & 2 \end{vmatrix}.$$

В полученном определителе ко 2-ой строке прибавим 1-ю и к 3-й строке – 1-ю, умноженную на (-2), получим:

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

Прибавив к 3-й строке 2-ю, умноженную на (-2), и к 4-й строке – 2-ю, умноженную на (-1), получим:

$$\Delta = 3 \cdot \begin{vmatrix} 1 & -3 & 1 & 4 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & -2 & -7 \\ 0 & 0 & -4 & 1 \end{vmatrix}.$$

В полученном определителе к 4-ой строке прибавим 3-ю, умноженную на (-2), получим:

$$\Delta = 3 \cdot \begin{vmatrix} 1 & -3 & 1 & 4 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & -2 & -7 \\ 0 & 0 & 0 & 15 \end{vmatrix}.$$

Таким образом, данный определитель приведен к треугольному виду, и, следовательно,

$$\Delta = 3 \cdot 1 \cdot 2 \cdot (-2) \cdot 15 = -180.$$

2. Метод понижения порядка основан на использовании формул (7). Формула разложения определителя по строке (столбцу) принимает особенно простой вид, когда в этой строке (столбце) все элементы равны нулю, кроме одного.

П р и м е р 6. *Вычислить определитель*

$$\Delta = \begin{vmatrix} 3 & 1 & 2 & 4 \\ 0 & 0 & -1 & 6 \\ 2 & 1 & 3 & 1 \\ 2 & -2 & 3 & 1 \end{vmatrix}$$

методом понижения порядка.

<u>*Решение.*</u> *Вычтем из 3-й строки 4-ю и получим*

$$\Delta = \begin{vmatrix} 3 & 1 & 2 & 4 \\ 0 & 0 & -1 & 6 \\ 0 & 3 & 0 & 0 \\ 2 & -2 & 3 & 1 \end{vmatrix}.$$

Полученный определитель разложим по 3-й строке:

$$\Delta = 3 \cdot (-1)^{3+2} \cdot \begin{vmatrix} 3 & 2 & 4 \\ 0 & -1 & 6 \\ 2 & 3 & 1 \end{vmatrix}.$$

Прибавив к 3-му столбцу 2-ой, умноженный на 6, получим:

$$\Delta = -3 \cdot \begin{vmatrix} 3 & 2 & 16 \\ 0 & -1 & 0 \\ 2 & 3 & 19 \end{vmatrix}.$$

Полученный определитель разложим по 2-й строке:

$$\Delta = -3 \cdot (-1) \cdot (-1)^{2+2} \cdot \begin{vmatrix} 3 & 16 \\ 2 & 19 \end{vmatrix} = 3 \cdot 25 = 75.$$

Обратная матрица.

Матрица A^{-1} называется ***обратной к матрице*** A, если

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

Матрица A, для которой существует обратная матрица, называется *обратимой.*

Из определения следует, что обратимой может быть лишь квадратная матрица, так как равенство $AA^{-1} = A^{-1}A$ возможно лишь для квадратных матриц A и A^{-1} одинакового порядка.

Кваратная матрица A называется *вырожденной (особенной)*, если $|A| = 0$, и *невырожденной (неособенной)*, если $|A| \neq 0$.

Пусть $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. Матрица

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

составленная из алгебраических дополнений A_{ij} к элементам a_{ij} матрицы A называется *присоединенной* к матрице A.

Теорема (критерий обратимости). *Матрица обратима тогда и только тогда, когда она не вырождена.*

Обратная матрица вычисляется по формуле

$$A^{-1} = \frac{1}{|A|}\tilde{A}. \tag{8}$$

Свойства обратной матрицы.

1. $E^{-1} = E$, так как $E \cdot E = E$
2. $|A^{-1}| = 1/|A|$, так как $|A| \cdot |A^{-1}| = 1$.
3. $(A^{-1})^{-1} = A$, так как $AA^{-1} = A^{-1}A = E$.
4. $(A^T)^{-1} = (A^{-1})^T$, так как $(A^{-1})^T A^T = (AA^{-1})^T = E^T = E$.
5. $(AB)^{-1} = B^{-1}A^{-1}$, так как $(AB)(B^{-1}A^{-1}) = E$.

Вычисление обратной матрицы.

Соотношение (8) дает явный вид обратной матрицы. Оно полезно в теоретических исследованиях и совершенно неэффективно для практического вычисления (разве что для матриц второго порядка) вследствие большого объема требуемых вычислений. Для получения обратной матрицы к матрице n-го порядка согласно (8) требуется вычислить n^2 определителей $(n-1)$-го порядка и один определитель n-го порядка. В вычислительной математике используются различные дополнительные приемы вычисления обратной матрицы, которые по объему вычислений равносильны вычислению всего лишь двух определителей n-го порядка. Рассмотрим один из таких методов, в основе которого лежит *метод Гаусса*:

1) формируем расширенную матрицу $(A|E)$ приписыванием к матрице A справа матрицы E того же порядка;

2) с помощью метода Гаусса, производя элементарные преобразования **только над строками**, приводим сформированную расширенную матрицу к виду $(E|B)$, что всегда возможно, если A не вырождена.

Элементарными преобразованиями матрицы называются преобразования следующих типов:

а) перестановка двух строк (столбцов) матрицы;

б) умножение строки (столбца) матрицы на число, отличное от нуля;

в) прибавление к одной строке (столбцу) матрицы другой ее строки (соответственно столбца), умноженной на любое число.

Тогда $A^{-1} = B$.

П р и м е р 7. *Выяснить, существует ли матрица, обратная матрице*

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

и если существует, то найти ее.

<u>Решение.</u> Так как $\det A = -6 \neq 0$, *то матрица A невырожденная и A^{-1} существует.*

Способ 1. Найдем матрицу A по формуле (8). Алгебраические дополнения соответствующих элементов матрицы A:

$$A_{11} = \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} = -6; \quad A_{12} = -\begin{vmatrix} 0 & 2 \\ -1 & 1 \end{vmatrix} = -2; \quad A_{13} = \begin{vmatrix} 0 & 0 \\ -1 & 3 \end{vmatrix} = 0;$$

$$A_{21} = -\begin{vmatrix} 0 & 1 \\ 3 & 1 \end{vmatrix} = 3; \quad A_{22} = \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2; \quad A_{23} = -\begin{vmatrix} 1 & 0 \\ -1 & 3 \end{vmatrix} = -3;$$

$$A_{31} = \begin{vmatrix} 0 & 1 \\ 0 & 2 \end{vmatrix} = 0; \quad A_{32} = -\begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = -2; \quad A_{33} = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} = 0.$$

Следовательно,

$$A^{-1} = -\frac{1}{6}\begin{pmatrix} -6 & 3 & 0 \\ -2 & 2 & -2 \\ 0 & -3 & 0 \end{pmatrix}.$$

Способ 2. Найдем A^{-1} *с помощью расширенной матрицы и метода Гаусса. Составим расширенную матрицу*

$$C = \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 \\ -1 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \sim$$

Прибавим к 3-й строке 1-ю, получим

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

Поменяем местами 2-ю и 3-ю строки, тогда

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

Прибавив ко 2-й строке 3-ю, умноженную на (-1), получим

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

Умножив 2-ю строку на $1/3$*, а 3-ю-на* $1/2$*, имеем*

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

Вычтем из 1-й строки 3-ю, тогда

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

и

$$A^{-1} = \begin{pmatrix} 1 & -1/2 & 0 \\ 1/3 & -1/3 & 1/3 \\ 0 & 1/2 & 0 \end{pmatrix} = -\frac{1}{6}\begin{pmatrix} -6 & 3 & 0 \\ -2 & 2 & -2 \\ 0 & -3 & 0 \end{pmatrix}.$$

4. Ранг матрицы.

4.1. Теоретические сведения

Терминология и обозначения. *Рангом* ненулевой матрицы называется максимальный порядок ненулевых миноров этой матрицы. Ранг нулевой матрицы по определению считается равным нулю.

Обозначение: rg A, rang A и др.

Из определения вытекают следующие факты:

1) ранг матрицы не превосходит ее размеров: если $A \in \mathbb{R}^{m\times n}$, то $\mathrm{rg}A \le \min(m, n)$;

2) равенство $\mathrm{rg}A = r > 0$ равносильно выполнению двух условий:

   а) в матрице A существует ненулевой минор r-го порядка,

   б) любой минор более высокого порядка равен нулю.

Пусть $\mathrm{rg}A = r > 0$. Любой ненулевой минор r-го порядка этой матрицы называется *базисным минором*, а строки и столбцы, в которым расположен базисный минор, – базисными строками и столбцами.

Разумеется, у матрицы может быть не один базисный минор, но все они имеют один и тот же порядок, равный рангу этой матрицы.

Метод Гаусса вычисление ранга матрицы.

Теоретическую основу этого метода для решения данной задачи составляют следующие факты:

– ранг верхней (нижней) трапециевидной матрицы равен количеству ненулевых строк (соответственно столбцов);

– элементарные преобразования не изменяют ее ранга;

– любая матрица элементарными преобразованиями строк и столбцов приводится к трапециевидной форме.

Метод Гаусса вычисления ранга матрицы состоит в приведении этой матрицы элементарными преобразованиями к верхне (нижней) трапециевидной форме и подсчете ее ненулевых строк (столбцов).

П р и м е р 8. *Найти ранг матрицы*

$$A = \begin{pmatrix} 2 & -1 & 1 & 3 & -4 \\ 1 & 3 & 4 & -1 & 1 \\ 5 & 1 & 6 & 5 & -7 \\ 7 & 0 & 7 & 7 & -11 \end{pmatrix}.$$

Решение.

$$A = \begin{pmatrix} 2 & -1 & 1 & 3 & -4 \\ 1 & 3 & 4 & -1 & 1 \\ 5 & 1 & 6 & 5 & -7 \\ 7 & 0 & 7 & 7 & -11 \end{pmatrix} \sim$$

Поменяем местами 1-ю и 2-ю строки, чтобы на месте a_{11} *оказалась единица:*

$$\sim \begin{pmatrix} 1 & 3 & 4 & -1 & 1 \\ 2 & -1 & 1 & 3 & -4 \\ 5 & 1 & 6 & 5 & -7 \\ 7 & 0 & 7 & 7 & -11 \end{pmatrix} \sim$$

Умножим 1-ю строку матрицы на -2, -5, -7 и прибавим соответственно ко 2-й, 3-й, 4-й строкам, получим

$$\sim \begin{pmatrix} 1 & 3 & 4 & -1 & 1 \\ 0 & -7 & -7 & 5 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix} \sim$$

Поменяем местами 3-ю и 4-ю строки и 3-й и 4-й столбцы, получим

$$\sim \begin{pmatrix} 1 & 3 & -1 & 4 & 1 \\ 0 & -7 & 5 & -7 & -6 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$rgA = 3$, *так как трапециевидная матрица имеет три ненулевых строки.*

Линейные пространства
Понятие линейного пространства

Рассмотрим множества объектов любой природы, для элементов которых каким-либо способом (причем, безразлично каким) определены операция сложения двух элементов и операция умножения на число элемента этого множества. Такие множества, называемые линейными пространствами, обладают целом рядом общих свойств, которые и будут установлены ниже.

Множество $\mathcal{L}$ элементов любой природы будем называть *линейным пространством*, если выполнены следующие требования.

I. Имеется правило, посредством которого $\forall x, y \in \mathcal{L}$ ставится в соответствие элемент $z \in \mathcal{L}$, называемый *суммой* и обозначаемый $z = x + y$.

II. Имеется правило, посредством которого $\forall x \in \mathcal{L}$ и $\forall \lambda \in \mathbb{R}(\mathbb{C})$ ставится в соответствие элемент $u \in \mathcal{L}$, называемый *произведением элемента* x *на число* λ и обозначаемый $u = \lambda x$.

III. Указанные два правила подчинены следующим восьми аксиомам:
$\forall x, y, z \in \mathcal{L}, \quad \forall \lambda, \mu \in \mathbb{R}(\mathbb{C})$

1) $x + y = y + x$;
2) $(x + y) + z = x + (y + z)$;
3) $\exists \theta \in \mathcal{L} : x + \theta = x$;
4) $\forall x \ \exists x' \in \mathcal{L}$(противоположный элемент): $x + x' = \theta$;
5) $1 \cdot x = x$;
6) $\lambda(\mu x) = (\lambda\mu)x$;
7) $\lambda(x + y) = \lambda x + \lambda y$;
8) $(\lambda + \mu)x = \lambda x + \mu x$.

Если число $\lambda \in \mathbb{R}$, то множество $\mathcal{L}$ называется вещественным линейным пространством. Если число $\lambda \in \mathbb{C}$, то множество $\mathcal{L}$ называется комплексным линейным пространством.

Примеры линейных пространств.

1. Множество функций $C_{[a,b]}$, определенных и непрерывных на $[a, b]$. Операции сложения таких функций и умножения их на вещественные числа определены обычными правилами математического анализа. Элементарно проверяется справедливость восьми аксиом, в частности, нулевым элементом является функция, тождественно равная нулю на отрезке $[a, b]$. Это позволяет заключить, что $C_{[a,b]}$ является линейным пространством.

2. Множество $P_n(x)$ алгебраических многочленов степени не выше $n \in \mathbb{N}$, с операциями, опеределенными так же, как в предыдущем примере. Заметим, что множество $P_n(x)$, если его рассматривать на отрезке $[a, b]$, является подмножеством линейного пространства $C_{[a,b]}$, рассмотренного в предыдущем примере.

П р и м е р 1. *Определить, является ли множество матриц $M^{2\times 2}$ вида $\begin{pmatrix} a & b \\ c & 1 \end{pmatrix}$, где $a, b, c \in \mathbb{R}$, линейным пространством?*

Решение.

Пусть $A = \begin{pmatrix} a_1 & b_1 \\ c_1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} a_2 & b_2 \\ c_2 & 1 \end{pmatrix}$, *тогда*

$C = A + B = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & 2 \end{pmatrix}$.

$C \notin M^{2\times 2}$, *следовательно, множество не является ни вещественным, ни комплексным линейным пространством.*

П р и м е р 2. *Определить, является ли множество матриц* $S^{2\times 2} = \begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$, где $a, b, c \in \mathbb{R}$, *вещественным линейным пространством*?

Решение.

Пусть $A = \begin{pmatrix} a_1 & b_1 \\ c_1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} a_2 & b_2 \\ c_2 & 0 \end{pmatrix}$, *тогда*

$C = A + B = \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c_1 + c_2 & 0 \end{pmatrix} \in S^{2\times 2}$,

$\lambda A = \begin{pmatrix} \lambda a_1 & \lambda b_1 \\ \lambda c_1 & 0 \end{pmatrix} \in S^{2\times 2}$, *где* $\lambda \in \mathbb{R}$, *все 8 аксиом выполняются тоже, следовательно, множество матриц является вещественным линейным пространством.*

Свойства линейного пространства.

1. Любое линейное пространство имеет только один нулевой элемент.

2. Каждый элемент линейного пространства имеет только один противоположный элемент.

3. Если элемент $(-x)$ противоположен элементу x, то элемент x является противоположным для $(-x)$.

4. Для любых двух элементов a и b уравнение $a + x = b$ относительно x имеет решение, и притом единственное.
   Разностью двух элементов $b - a$ называется такой элемент x, который является решением уравнения $a + x = b$

5. Произведение произвольного элемента линейного пространства на число 0 равно нулевому элементу: $0 \cdot x = \theta$.

6. Элемент, противоположный данному элементу x, равен произведению x на число -1: $(-x) = (-1)x$.

7. Произведение нулевого элемента на любое число есть нулевой элемент: $\lambda\theta = \theta$.

Базис и размерность линейного пространства

<u>**Линейно независимые и линейно зависимые системы элементов**</u>.

Рассмотрим произвольное линейное пространство $\mathcal{L}$ с элементами $x, y, \ldots, z$.

Линейной комбинацией элементов $x, y, \ldots, z$ пространства $\mathcal{L}$ мы будем называть сумму произведений этих элементов на произвольные числа, т.е. выражения вида

$$\alpha x + \beta y + \ldots + \gamma z,$$

где $\alpha, \beta, \ldots, \gamma \in \mathbb{R}(\mathbb{C})$, в зависимости от того, вещественное или комплексное пространство $\mathcal{L}$.

Элементы $x, y, \ldots, z \in \mathcal{L}$ называются *линейно зависимыми*, если найдутся такие числа $\alpha, \beta, \ldots, \gamma$, из которых хотя бы одно отлично от нуля, что линейная комбинация элементов $x, y, \ldots, z$ с указанными числами является нулевым элементом пространства $\mathcal{L}$, т.е. имеет место равенство

$$\alpha x + \beta y + \ldots + \gamma z = \theta. \tag{1}$$

Элементы $x, y, \ldots, z$ линейного пространства $\mathcal{L}$ называются *линейно независимыми*, если равенство (1) выполняется только при условии $\alpha = \beta = \ldots = \gamma = 0$.

Теорема 1. *Для того чтобы элементы $x, y, \ldots, z \in \mathcal{L}$ были линейно зависимыми, необходимо и достаточно, чтобы один из этих элементов являлся линейной комбинацией остальных.*

П р и м е р 3. *Выяснить, является ли линейно независимой каждая из следующих систем элементов*:

1) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ *в пространстве* $\mathbb{R}^{2\times 2}$*вещественнозначных матриц* ;

2) $1,\ \sin^2 x,\ \cos 2x$ *в пространстве* $C_{(-\infty,+\infty)}$ *вещественнозначных функций, непрерывных на* $[a, b]$?

<u>*Решение.*</u>

1) Составим линейнейную комбинацию данных матриц

$$\alpha_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 & 2\alpha_2 \\ 0 & \alpha_1 \end{pmatrix}.$$

Она равна нулевой матрице $O = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ *только в том случае, когда* $\alpha_1 = \alpha_2 = 0$, *а это означает, что данная система матриц является линейно независимой.*

2) *Составим линейную комбинацию данных функций и приравняем ее к нулевому элементу (функции, тождественно равной нулю):*

$$\alpha_1 \cdot 1 + \alpha_2 \sin^2 x + \alpha_3 \cos 2x = \theta.$$

Это равенство справедливо, например, при $\alpha_1 = 1,\ \alpha_2 = -2,\ \alpha_3 = -1$, *следовательно, данная система элементов является линейно зависимой.*

Свойства систем элементов.

1. Если в системе элементов есть нулевой элемент, то эта система будет линейно зависимой.

2. Если подсистема элементов линейно зависимая, то вся система будет линейно зависимой.

3. Любая подсистема линейно независимой системы элементов линейно независимая.

4. Если система элементов $x, y, \ldots, z$ линейно независимая, а система элементов $x, y, \ldots, z, z'$ линейно зависимая, то z' представляется в виде линейной комбинации элементов $x, y, \ldots, z$.

Система линейно независимых элементов $e_1, e_2, \ldots, e_n \in \mathcal{L}$ называется *базисом пространства* $\mathcal{L}$, если для каждого элемента $x \in \mathcal{L}$ найдутся числа $x_1, x_2, \ldots, x_n$:

$$x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n, \tag{2}$$

При этом равенство (2) называется *разложением элемента x по базису* $e_1, e_2, \ldots, e_n$, а $x_1, x_2, \ldots, x_n$ – *координатами* элемента x (относительно базиса $e_1, e_2, \ldots, e_n$).

Любой элемент $x \in \mathcal{L}$ может быть разложен по базису $e_1, e_2, \ldots, e_n$ единственным образом, т.е. координаты любого элемента x относительно базиса $e_1, e_2, \ldots, e_n$ определяются однозначно.

Значение базиса заключается также и в том, что операции сложения элементов и умножения их на числа при задании базиса превращаются в соответствующие операции над числами – координатами этих элементов.

Теорема 2. *При сложении любых двух элементов линейного пространства $\mathcal{L}$ (относительно любого базиса пространства $\mathcal{L}$) их координаты складываются, а при умножении произвольного элемента на любое число λ соответствующие координаты этого элемента умножаются на число λ.*

Примеры базисов различных пространств.

1. Для линейного пространства матриц $\mathbb{R}^{2\times 2}$ базисом является, например, система элементов $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

2. Для линейного пространства многочленов степени не выше двух базисом является, например, система элементов $1, x, x^2$. Координатами элемента $5 - x - x^2$ данного пространства в этом базисе будут $(x_1, x_2, x_3)^T = (5, -1, 1)^T$.

П р и м е р 4. *Определить, какая из следующих систем элементов является базисом линейного пространства $\mathbb{R}^3$ векторов с вещественными коэффициентами:*

1. $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}$.

2. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

3. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$.

4. $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$

Решение.

Системы элементов 2, 4 являются базисом линейного пространства $\mathbb{R}^3$*, т.к. эти системы линено независимые и любой элемент этого пространства представляется в виде линейной комбинации элементов одной из этих систем.*

Системы элементов 1, 3 не являются базисом линейного пространства $\mathbb{R}^3$*, т.к. система 3 является линейно зависимой, а через систему 1 нельзя представить в виде линейной комбинации любой элемент этого пространства.*

Линейное пространство $\mathcal{L}$ называется *n-мерным*, если в нем существует линейно независимая система из n элементов, а любая система из $n+1$ элементов линейно зависимая. При этом число n называется *размерностью* пространства $\mathcal{L}$.

Обозначение: $\dim(\mathcal{L})$.

Линейное пространство L называется *бесконечномерным*, если в нем существует любое число линейно независимых элементов.

Обозначение: $\dim(L) = \infty$.

Примеры размерностей линейных пространств.

1. $\dim(\mathbb{R}^5) = 5$.

2. $\dim(\mathbb{R}^{3\times 3}) = 9$.

3. $\dim(P_n(x)) = n+1$.

Теорема 3. *Если $\mathcal{L}$ – линейное пространство размерности n, то в нём любая линейно независимая система из n элементов образует базис.*

Теорема 4. *Если в пространстве $\mathcal{L}$ существует базис из n элементов, то это пространство размерности n.*

1.3 Подпространства линейных пространств

Подмножество $\mathcal{L}'$ линейного пространства $\mathcal{L}$ называется *линейным подпространством*, если выполняются для этого подмножества следующие требования:

I. Если $x, y \in \mathcal{L}'$, то $x + y \in \mathcal{L}'$.

II. Если $x \in \mathcal{L}', \lambda \in \mathbb{R}(\mathbb{C})$, то $\lambda x \in \mathcal{L}'$.

Теорема 5. *Линейное подпространство само является линейным пространством.*

Рассмотрим для линейного пространства $\mathcal{L}$ линейные подпространства: само линейное пространство $\mathcal{L}$ и нулевое подпространство $\{\theta\}$. Эти два подпространства называются *несобственными*, а все остальные *собственными*.

Примеры линейных подпространств.

1. В линейном пространствеа V_3 свободных векторов трехмерного пространства линейное подпространство образуют:
   а) все векторы, параллельные данной плоскости;
   б) все векторы, параллельные данной прямой.
2. Любое решение однородной системы линейных алгебраических уравнений от n переменных можно рассматривать как вектор в линейном пространстве $\mathbb{R}^n$. Множествотаких векторов является линейным подпространством в $\mathbb{R}^n$.

Понятие линейной оболочки.

Рассмотрим элементы $x, y, \ldots, z \in \mathcal{L}$. *Линейной оболочкой* элементов $x, y, \ldots, z$ будем называть совокупность всех линейных комбинаций этих элементов, т.е. множество элементов вида

$$\alpha x + \beta y + \ldots + \gamma z,$$

где $\alpha, \beta, \ldots, \gamma \in \mathbb{R}(\mathbb{C})$.

Обозначение: $L(x, y, \ldots, z)$.

Из определения следует, что каждое конечномерное пространство является линейной оболочкой элементов своего базиса.

Например, линейное подпространство, заданное однородной системой уравнений, является линейной оболочкой фундаментальной системы решений (ФСР).

Линейная оболочка произвольных элементов $x, y, \ldots, z$ линейного пространства $\mathcal{L}$, очевидно, является подпространством основного линейного пространства $\mathcal{L}$.

Пример 5. *Являются ли каждое подмножество линейного пространства $\mathbb{R}^{n \times n}$ линейным подпространством:*

а) *множество симметричных матриц;*

б) *множество вырожденных матриц?*

Решение.

а) *сумма симметричных матриц есть симметричная матрица; при умножении симметричной матрицы на число получаем также симметричную матрицу, следовательно, множество симметричных матриц является линейным подпространством.*

б) *рассмотрим сумму вырожденных матриц:*

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

полученная матрица невырожденная ($\det \neq 0$). Следовательно, множество вырожденных матриц не является линейным подпространством.

Теорема 6 (о монотонности размерности). *Размерность любого подпространства $\mathcal{L}'$ n-мерного линейного пространства $\mathcal{L}$ не превосходит размерности n пространства $\mathcal{L}$. Если размерности линейного пространства и линейного подпространства совпадают, то подпространство совпадает с пространством.*

Теорема 7. *Если система элементов $e_1, e_2, \ldots, e_k$ является базисом k-мерного подпространства $\mathcal{L}'$ n-мерного линейного пространства $\mathcal{L}$, то этот базис можно дополнить элементами $e_{k+1}, e_{k+2}, \ldots, e_n \in \mathcal{L}$ так, что система $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ будет являться базисом всего пространства $\mathcal{L}$.*

Теорема 8 (о размерности линейной оболочки). *Размерность линейной оболочки $\dim(L(x, y, \ldots, z))$ равна максимальному числу линейно независимых элементов в системе $x, y, \ldots, z$. В частности, если система $x, y, \ldots, z$- линейно независимая, то размерность линейной оболочки системы $x, y, \ldots, z$ равна числу элементов в этой системе, а сами элементы $x, y, \ldots, z$ образуют базис линейной оболочки.*

Рангом системы элементов в линейном пространстве называют размерность линейной оболочки этой системы элементов.

Если в качестве системы элементов рассматривать строки (столбцы) матрицы, то получим следующее определение ранга матрицы: *ранг матрицы*-максимальное число линейно независимых строк(столбцов) матрицы.

<u>Сумма и пересечение подпространств</u>.

Пусть $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_k$ – линейные подпространства линейного пространства $\mathcal{L}$. *Суммой подпространств* $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_k$ называется множество всевозможных элементов x, представимых в виде

$$x = x_1 + x_2 + \ldots + x_k, \tag{3}$$

где $x_i \in \mathcal{L}_i, \quad i = 1, 2, \ldots, k$.

Обозначение: $\mathcal{L}_1 + \mathcal{L}_2 + \ldots + \mathcal{L}_k$.

Представление (3) элемента x называется *разложением элемента* x *по подпространствам* $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_k$.

Пересечением подпространств $\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_k$ называется множество

$$L_1 \cap L_2 \cap \ldots \cap L_k = \{x \in \mathcal{L} | x_i \in \mathcal{L}_i, i = 1, 2, \ldots, k\}.$$

<u>Замечание</u>. Пересечение подпространств не может быть пустым множеством, т.к. всегда содержит нулевой элемент θ пространства.

Примеры суммы и пересечения линейных подпространств.

Пусть V_3 – линейное пространство геометрических векторов трехмерного пространства.

$\mathcal{L}_1$ – линейное подпространство, состоящее из векторов, параллельных плоскости OXY.

$\mathcal{L}_2$ – линейное подпространство, состоящее из векторов, параллельных плоскости OXZ.

$\mathcal{L}_1 + \mathcal{L}_2$ – все пространство V_3, $\mathcal{L}_1 \cap \mathcal{L}_2$ – множество векторов, параллельных оси OX.

Теорема 9. *Сумма и пересечение подпространств линейного пространства* $\mathcal{L}$ *являются линейными подпространствами пространства* $\mathcal{L}$.

Теорема 10. *Для любых линейных подпространств* $\mathcal{L}_1$ *и* $\mathcal{L}_2$ *линейного пространства* $\mathcal{L}$ *справедливо равенство:*

$$\dim(\mathcal{L}_1 + \mathcal{L}_2) = \dim(\mathcal{L}_1) + \dim(\mathcal{L}_2) - \dim(\mathcal{L}_1 \cap \mathcal{L}_2).$$

<u>Прямая сумма подпространств</u>.

Сумма подпространств линейного пространства называется *прямой суммой*, если разложение в ней по слагаемым подпространства единственно.

Обозначение: $\mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \ldots \oplus \mathcal{L}_k$.

Пример прямой суммы линейных подпространств.

Пусть V_3 – линейное пространство геометрических векторов трехмерного пространства.

$\mathcal{L}_1$ – линейное подпространство, состоящее из векторов, параллельных оси OY.

$\mathcal{L}_2$ – линейное подпространство, состоящее из векторов, параллельных плоскости OXZ.

$L_1 \oplus L_2 = V_3$.

Теорема 11. *Для того чтобы n-мерное пространство $\mathcal{L}$ представляло собой прямую сумму подпространств $\mathcal{L}_1$ и $\mathcal{L}_2$ достаточно, чтобы пересечение $\mathcal{L}_1 \cap \mathcal{L}_2 = \{\theta\}$ и чтобы $\dim(\mathcal{L}) = \dim(\mathcal{L}_1) + \dim(L_2)$.*

1.4 Преобразование координат при преобразовании базиса

Пусть в линейном пространстве $\mathcal{L}$ заданы два базиса: $e = (e_1, e_2, \ldots, e_n)$ и $e' = (e'_1, e'_2, \ldots, e'_n),\ \dim \mathcal{L} = n$.

Выясним, как меняются координаты элемента при переходе от базиса e к e'. Так как элементы базиса e' являются элементами линейного пространства $\mathcal{L}$, то каждый из них можно разложить по базису e.

$$\begin{array}{c} e'_1 = a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n, \\ e'_2 = a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n, \\ \ldots \quad \ldots \quad \ldots \\ e'_n = a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n. \end{array} \qquad (4)$$

Коэффициенты a_{ij} в равенстве (4) образуют матрицу $A = (a_{ij}) \in \mathbb{R}^{\,n\times n}$, которая называется *матрицей перехода* от базиса e к e'.

Обозначение: $P_{e\to e'}$.

Равенства (4) в матричном виде могут быть записаны : $e' = eP_{e\to e'}$.

Теорема 12. *Матрица перехода от одного базиса к другому является невырожденной.*

Теорема 13. *Координаты элемента x в базисах e и e' связаны между собой следующим образом:*

$$x_e = P_{e\to e'} x_{e'}.$$

П р и м е р 6. *В линейном пространстве многочленов не выше второй степени с действительными коэффициентами заданы два базиса:*
$e : e_1 = 1, \quad e_2 = x, \quad e_3 = x^2;$
$e' : e'_1 = 1, \quad e'_2 = x - 1, \quad e'_3 = (x-1)^2$. *Найти* $P_{e\to e'}$.

Решение.
$1 = 1 \cdot e_1 + 0 \cdot e_2 + 0 \cdot e_3;$
$x - 1 = -1 \cdot e_1 + 1 \cdot e_2 + 0 \cdot e_3;$
$(x-1)^2 = 1 \cdot e_1 - 2 \cdot e_2 + 1 \cdot e_3. \quad P_{e\to e'} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$

П р и м е р 7. *Найти координаты элемента x в базисе e', если* $x_e = 3e_1 - 2e_2; \quad e'_1 = 5e_1 + 3e_2, \quad e'_2 = e_1 + e_2$.

Решение.
$x_{e'} = P^{-1}_{e\to e'} x_e; \quad P_{e\to e'} = \begin{pmatrix} 5 & 1 \\ 3 & 1 \end{pmatrix}; \ P^{-1}_{e\to e'} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -3 & 5 \end{pmatrix};$

$x_{e'} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 5 \\ -19 \end{pmatrix}.$

Евклидовы пространства
Понятие вещественного евклидова пространства

Вещественное линейное пространство $\mathcal{E}$ называется *вещественным евклидовым пространством*, если выполняются следующие два требования:

I. Имеется правило, посредством которого любым двум элементам этого пространства $\forall x, y \in \mathcal{E}$ ставится в соответствие вещественное число, называемое *скалярным произведением* этих элементов и обозначаемое символом (x, y).

II. Указанное правило подчинено следующим четырем аксиомам:
$\forall x, y, z \in \mathcal{E}$ и $\forall \in \mathbb{R}$

1) $(x, y) = (y, x)$;
2) $(x + y, z) = (x, z) + (y, z)$;
3) $(\lambda x, y) = \lambda(x, y)$;
4) $(x, x) \geq 0$, причем $(x, x) = 0 \Leftrightarrow x = \theta$.

Пример вещественного евклидова пространства.

Множество непрерывных на отрезке $[a, b]$ функций $C_{[a,b]}$, где скалярное произведение задано:

$$(f(x), g(x)) = \int_a^b f(x)g(x)dx.$$

П р и м е р 8. *Является ли вещественное линейное пространство $\mathbb{R}^2$ вещественным евклидовым пространством, если паре векторов $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ поставлено в соответствие число:*
$(x, y) = x_1 x_2 y_1 y_2$.

Решение. *Проверяем выполнение четырех аксиом:*

1. $(x, y) = x_1 x_2 y_1 y_2 = y_1 y_2 x_1 x_2 = (y, x)$;

2. $(x + y, z) = (x_1 + y_1)(x_2 + y_2) z_1 z_2 \neq (x, z) + (y, z)$.

вторая аксиома не выполняестся, следовательно, данное вещественное пространство не является вещественным евклидовым пространством.

Свойства скалярного произведения.

1) $(x, y + z) = (x, y) + (x, z)$;

2) $(x, \lambda y) = \lambda(x, y)$;

3) $(x, \theta) = 0$;

4) $(\sum\limits_{i=1}^{n} \alpha_i x_i, y) = \sum\limits_{i=1}^{n} \alpha_i (x_i, y),\ \alpha_i \in \mathbb{R},\ i = 1, 2, \ldots, n$;

5) если $x, y \in \mathcal{E}$ такие, что $\forall z \in \mathcal{E}$ выполняется равенство $(x, z) = (y, z)$, то $x = y$.

Теорема 14. (Неравенство Коши-Буняковского). $\forall x, y \in \mathcal{E}$ *справедливо неравенство*

$$(x, y)^2 \le (x, x)(y, y).$$

Основные метрические понятия.

Длиной элемента x евклидова пространства, будем называть арифметическое значение корня из скалярного квадрата этого элемента.

Обозначение: $|x|$.

$$|x| = \sqrt{(x, x)}.$$

Из аксиом скалярного произведения вытекают следующие факты:

- Любой элемент $x \in \mathcal{E}$ имеет длину, при этом

  $|x| \ge 0, \quad \forall x \in \mathcal{E}$ и $\quad (x, x) = 0 \Leftrightarrow x = \theta$;

- $|\alpha x| = |\alpha|\,|x|\,, \quad \forall x \in \mathcal{E}, \quad \forall \alpha \in \mathbb{R}$.

В новой терминологии неравенство Коши-Буняковского может быть записано следующим образом:

$$|(x, y)| \le |x||y|.$$

Элемент, длина которого равна единице, называется *нормированным.*

Любой ненулевой элемент можно нормировать, поделив его на длину.

Теорема 15. В евклидовом пространстве $\forall x, y \in \mathcal{E}$ *справедливы следующие неравенства (неравенства треугольника)*

$$|\,|x| - |y|\,| \le |x + y| \le |x| + |y|\,.$$

Углом в вещественном евклидовом пространстве будем называть угол $\varphi \in (0;\pi]$, который определяется по формуле:

$$\cos\varphi = \frac{(x,y)}{\sqrt{(x,x)}\sqrt{(y,y)}} = \frac{(x,y)}{|x|\,|y|}. \tag{5}$$

Корректность определения следует из неравенства Коши-Буняковского.

П р и м е р 9. *Найти в евклидовом пространстве непрерывных на [0;1] вещественно-значных функций $C_{[0,1]}$:*

1) длину элемента $f(x) = x$;

2) скалярное произведение $f(x) = x$; $g(x) = e^x$;

3) угол между элементами $g(x) = x$ и $f(x) = 1$,

если скалярное произведение задано следующим образом: $\int\limits_0^1 f(x)g(x)dx$.

Решение.

1. $(f(x), f(x)) = \int\limits_0^1 x^2 dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$, $\quad |f(x)| = \sqrt{(f(x), f(x))} = \frac{1}{\sqrt{3}}$.

2. $(f(x), g(x)) = \int\limits_0^1 x \cdot e^x dx = \left\{ \begin{array}{cc} e^x dx = dv & v = e^x \\ u = x & du = dx \end{array} \right\} = xe^x|_0^1 - \int\limits_0^1 e^x dx = (xe^x - e^x)|_0^1 = e^x(x-1)|_0^1 = 1.$

3. *Воспользуемся формулой (5).*

   - $(f(x), g(x)) = \int\limits_0^1 x dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$;
   - $(g(x), g(x)) = \int\limits_0^1 x^2 dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$; $\quad \sqrt{(g(x), g(x))} = \frac{1}{\sqrt{3}}$;
   - $(f(x), f(x)) = \int\limits_0^1 dx = x|_0^1 = 1$; $\quad \sqrt{(f(x), f(x))} = 1$;

$$\cos\varphi = \frac{(f(x), g(x))}{\sqrt{(f(x), f(x))}\sqrt{(g(x), g(x))}} = \frac{\sqrt{3}}{2}; \quad \varphi = 30°.$$

2.2 Ортогональные и ортонормированные базисы

Элементы $x, y \in \mathcal{E}$ называются *ортогональными*, если их скалярное произведение равно нулю $(x, y) = 0$.

Согласно свойству скалярного умножения нулевой элемент ортогонален любому элементу.

Система элементов $e_1, e_2, \ldots, e_n$ называется *ортогональной*, если выполняется $(e_i, e_j) = 0$ при $i \neq j$.

Систему, состоящую из одного элемента, будем считать ортогональной.

Система элементов $e_1, e_2, \ldots, e_n$ называется *ортонормированной*, если

$$(e_i, e_j) = 0, \ i \neq j;$$
$$(e_i, e_j) = 1, \ i = j.$$

Теорема 16. *Ортогональная система ненулевых элементов является линейно независимой.*

Следствие. Ортонормированная система элементов является линейно независимой.

Так как евклидово пространство является линейным пространством, то правомерно говорить о размерности и базисах этого пространства. Евклидово пространство может быть конечномерным и бесконечномерным.

Если базис евклидова пространства представляет собой ортогональную систему элементов, то этот *базис* называют *ортогональным*.

Если в линейном пространстве все базисы равноправны, то в евклидовом пространстве наличие скалярного произведения позволяет выделить ортогональный и ортонормированный базисы, которые более удобны и играют в линейной алгебре роль, аналогичную роли декартовой прямоугольной системы координат в аналитической геометрии.

Теорема 17. *Во всяком евклидовом n-мерном пространстве $\mathcal{E}_n$ существует ортонормированный базис.*

Доказательство. Рассматриваемое доказательство носит название *ортогонализации Грама-Шмидта.*

Пусть $f_1, f_2, \ldots, f_n$ – произвольный базис евклидова пространства $\mathcal{E}_n$.

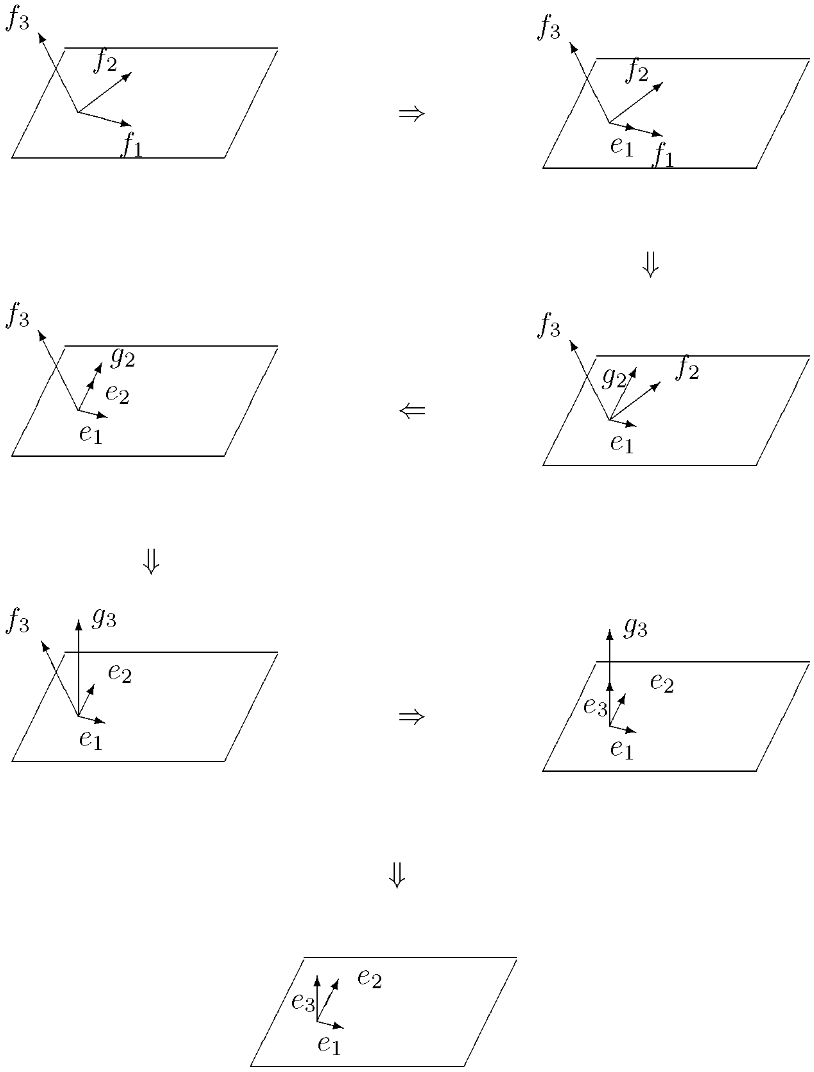
Модифицируя этот базис, мы будем строить новый $e_1, e_2, \ldots, e_n$, который будет ортонормированным.

Введем дополнительно систему элементов $g_1, g_2, \ldots, g_n$:

$$g_1 = f_1, \quad g_2 = f_2 - (f_2, e_1)e_1, \quad \ldots, \quad g_n = f_n - (f_n, e_1)e_1 - \ldots - (f_n, e_{n-1})e_{n-1};$$

$$e_1 = \frac{g_1}{|g_1|}, \quad e_2 = \frac{g_2}{|g_2|}, \ldots, e_n = \frac{g_n}{|g_n|}.$$

Рассмотрим пример при $n = 3$.



Для обоснования алгоритма нужно показать, что ни один из последовательно вычисленных элементов g_i не является нулевым (иначе процесс оборвался бы преждевременно) и что все элементы g_i попарно отогональны.

Тогда и элементы e_i образуют ортогональную систему, но при этом длина каждого из них равна единице.

Ортогональная система из n ненулевых элементов согласно теореме 14 линейно независимая и поэтому в n-мерном евклидовом пространстве является базисом.

Докажем по методу математической индукции.

1. Докажем для $n = 1$.

   $g_1 \neq 0$, т.к. $g_1 = f_1$, то систему, состоящую из одного элемента, считают ортогональной.

2. Пусть выполняется при $n = k$.

3. Докажем для $n = k + 1$.

Вычислим новый элемент g_{k+1} по формуле

$$g_{k+1} = f_{k+1} - (f_{k+1}, e_1)e_1 - \ldots - (f_{k+1}, e_k)e_k. \qquad (*)$$

Предположив, что $g_{k+1} = \theta$, получим, что

$$f_{k+1} = (f_{k+1}, e_1)e_1 + \ldots + (f_{k+1}, e_k)e_k,$$

т.е. элемент f_{k+1} является линейной комбинацией элементов $e_1, e_2, \ldots, e_k$, которые выражаются через $f_1, f_2, \ldots, f_k$.

Следовательно, f_{k+1} является линейной комбинацией $f_1, f_2, \ldots, f_k$, а следовательно, система элементов $f_1, f_2, \ldots, f_{k+1}$ линейно зависимая. Но это противоречит условию линейной независимости системы $f_1, f_2, \ldots, f_n$.

Итак, предположение о том, что $g_{k+1} = \theta$, привело к противоречию.

Покажем, что элемент g_{k+1} ортогонален каждому из элементов $e_1, e_2, \ldots, e_k$.

Умножим скалярно (*) на e_i, где $i \leq k$. Учитывая, что элементы $e_1, e_2 \ldots e_k$ образуют ортонормированную систему, получим:

$$(g_{k+1}, e_i) = (f_{k+1}, e_i) - (f_{k+1}, e_i)(e_i, e_i) = (f_{k+1}, e_i) - (f_{k+1}, e_i) = 0,$$

Следовательно, элементы $e_1, e_2, \ldots, e_k, \ldots, e_{k+1}$, где $e_{k+1} = \frac{g_{k+1}}{|g_{k+1}|}$, образуют ортонормированную систему. $\square$

В вычислениях удобны формулы, где сначала последовательно вычисляются элементы $g_1, g_2, \ldots, g_n$, а затем проводится их нормировка, приводящая к элементам $e_1, e_2, \ldots, e_n$:

$$g_1 = f_1; \quad g_2 = f_2 - \frac{(f_2, g_1)g_1}{(g_1, g_1)}; \quad \ldots \quad ; g_n = f_n - \frac{(f_n, g_1)g_1}{(g_1, g_1)} - \ldots - \frac{(f_n, g_{n-1})g_{n-1}}{(g_{n-1}, g_{n-1})};$$

$$e_1 = \frac{g_1}{|g_1|}; \quad e_2 = \frac{g_2}{|g_2|}; \ \ldots \ ; e_n = \frac{g_n}{|g_n|}.$$

П р и м е р 10. *В евклидовом пространстве дан базис:*

$$f_1 = (3, 1, 2)^T; \quad f_2 = (1, 1, 1)^T; \quad f_3 = (0, 2, 3)^T.$$

По этому базису построить ортонормированный базис. Скалярное произведение задано стандартным образом

$$(x, y) = x_1y_1 + x_2y_2 + x_3y_3.$$

Решение.

$$g_1 = f_1 = (3, 1, 2)^T; \quad g_2 = f_2 - \frac{(f_2, g_1)g_1}{(g_1, g_1)} = \frac{1}{7}(-2, 4, 1)^T;$$

$$g_3 = f_3 - \frac{(f_3, g_1)g_1}{(g_1, g_1)} - \frac{(f_3, g_2)g_2}{(g_2, g_2)} = \frac{1}{3}(-2, -2, 4)^T;$$

$$e_1 = \frac{1}{\sqrt{14}}(3, 1, 2)^T; \quad e_2 = \frac{7}{\sqrt{21}}(-2, 4, 1)^T; \quad e_3 = \frac{3}{2\sqrt{6}}(-2, -2, 4)^T.$$

Выражение скалярного произведения через компоненты сомножителей

Пусть $x, y \in \mathcal{E}$, $e_1, e_2, \ldots, e_n$ – базис этого пространства, x, y в этом базисе представляются:

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n,$$

$$y = y_1e_1 + y_2e_2 + \ldots + y_ne_n,$$

$$(x, y) = (\sum_{i=1}^{n} x_ie_i, \sum_{j=1}^{n} y_je_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_j(e_i, e_j). \qquad (6)$$

В матричном виде (6) перепишется следующим образом:

$$(x, y) = [x]_e^T \, \Gamma_e \, [y]_e \, , \qquad (7)$$

где

$$\Gamma_e = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) & \cdots & (e_1, e_n) \\ (e_2, e_1) & (e_2, e_2) & \cdots & (e_2, e_n) \\ \cdots & \cdots & \cdots & \cdots \\ (e_n, e_1) & (e_n, e_2) & \cdots & (e_n, e_n) \end{pmatrix} \text{ – матрица Грама.}$$

Матрица Грама является симметричной: $\Gamma_e = \Gamma_e^T$.

Если базис $e_1, e_2, \ldots, e_n$ ортонормированный, то $\Gamma_e = E$, а (7) перепишется следующим образом:

$$(x, y) = [x]_e^T \, [y]_e$$

Пусть в пространстве $\mathcal{E}$ заданы два базиса: $e = (e_1, e_2, \ldots, e_n)$, $e' = (e'_1, e'_2, \ldots, e'_n)$, $P_{e \to e'} = (a_{ij})$ – матрица перехода, $\Gamma_{e'} = (g'_{ij})$ – матрица Грама.

Согласно определению матрицы перехода:

$$e'_i = \sum_{k=1}^{n} a_{ki}e_k.$$

Тогда элемент матрицы *Грама* представим в виде:

$$g'_{ij} = (e'_i, e'_j) = (\sum_{k=1}^{n} a_{ki}e_k, \sum_{l=1}^{n} a_{lj}e_l) = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{ki}a_{lj}(e_k, e_l).$$

В матричной записи это эквивалентно:

$$\Gamma_{e'} = P^T_{e\to e'}\Gamma_e P_{e\to e'}.$$

Теорема 18. *Система элементов* $a_1, a_2, \dots, a_n \in \mathcal{E}$ *линейно зависима тогда и только тогда, когда определитель матрицы* Грама *этой системы равен нулю:*

$$det\Gamma_a = 0.$$

Теорема 19. *Определитель матрицы* Грама *любого базиса положителен.*

Пример 11. *Дано вещественное евклидово пространство, где скалярное произведение задано следующим образом:*

$$(x, y) = 4x_1y_1 - 2x_1y_2 - 2x_2y_1 + 4x_2y_2.$$

Вычислить матрицу Грама Γ_e *стандартного базиса*

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

и матрицу Грама Γ_f*, базиса* f*:*

$$f_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Решение.

1. *Вычислим матрицу Грама, воспользовавшись формулой:*

$$\Gamma_e = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) \\ (e_2, e_1) & (e_2, e_2) \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}.$$

2. *Найдем* Γ_f *двумя способами:*

- *Воспользуемся формулой:*

$$\Gamma_f = P^T_{e\to f}\Gamma_e P_{e\to f}$$

$P_{e\to f}$ *– это координаты базиса* f *в базисе* e *выписанные по столбцам:*

$$P_{e\to f} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$

$$\Gamma_f = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 4 & -2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 12 \end{pmatrix}.$$

- *Воспользуемся формулой:*

$$\Gamma_f = \begin{pmatrix} (f_1, f_1) & (f_1, f_2) \\ (f_2, f_1) & (f_2, f_2) \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 12 \end{pmatrix}.$$

Ортогональная матрица

Квадратная матрица Q называется *ортогональной*, если она удовлетворяет условию

$$Q^T Q = E.$$

Примеры ортогональных матриц.

1. Единичная матрица E.

2. Матрица Гивенса (вращения) $R = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$.

Свойства ортогональной матрицы.

- $\det Q = \pm 1$.
- $Q^{-1} = Q^T$.
- Если Q – ортогональная, то Q^T – тоже ортогональная матрица.
- $QQ^T = E$.
- Пусть Q_1, Q_2 – ортогональные матрицы одного порядка, тогда $Q_1 Q_2$ – ортогональная матрица.
- Пусть Q – ортогональная, тогда Q^{-1} – тоже ортогональная матрица.

Если матрица ортогональная, то удобно находить к ней обратную, например

$$R = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

$$R^{-1} = R^T = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Ортогональное дополнение

Пусть $\mathcal{H}$ – линейное подпространство $\mathcal{E}$. Элемент $x \in \mathcal{E}$ *ортогонален к линейному подпространству* $\mathcal{H}$, если x ортогонален любому элементу $y \in \mathcal{H}$.

Множество элементов из вещественного линейного евклидова пространства ортогональных линейному подпространству $\mathcal{H}$ называется *ортогональным дополнением* $\mathcal{H}^{\perp}$.

Теорема 20. *Ортогональное дополнение $\mathcal{H}^{\perp}$ является линейным подпространством.*

Теорема 21. *Пусть $\mathcal{H}$ – линейное подпространство $\mathcal{E}$, тогда*

$$\mathcal{H} \oplus \mathcal{H}^{\perp} = \mathcal{E}.$$

Следствие. Если $\mathcal{H}$ – линейное подпространство $\mathcal{E}$, то для любого $f \in \mathcal{E}$ существует и при том единственное разложение

$$f = g + h. \tag{8}$$

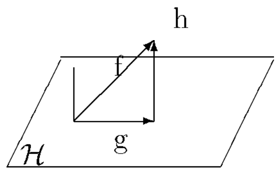
где $g \in \mathcal{H}, \quad h \in \mathcal{H}^{\perp}$.

В разложении (8)

g – ортогональная проекция элемента f на подпространство $\mathcal{H}$;

h – ортогональная составляющая элемента f.

Нахождение разложения (8) называется *задачей о перпендикуляре.* Этот термин заимствован из геометрии. Чтобы получить разложение геометрического вектора, достаточно опустить перпендикуляр из конца вектора f на плоскость.



Имея ввиду эту аналогию, называют

f – наклонной к подпространству $\mathcal{H}$;

h – перпендикуляром, опущенным на подпространство $\mathcal{H}$.

Аналогия с геометрическими векторами состоит не только в названии. Отметим несколько тех свойств g и h в разложении (8), которые имеют место и в геометрии т.к.

$$(f, f) = (g + h, g + h) = (g, g) + (h, h),$$

то

$$|f|^2 = |g|^2 + |h|^2 . \tag{9}$$

$$|h| \leq |f|.$$

Последнее неравенство свидетельствует о том, что длина перпендикуляра не превосходит длины наклонной. Равенство (9) называется *теоремой Пифагора в евклидовом пространстве*

П р и м е р 12. *В евклидовом пространстве со стандартным скалярным произведением построить* $L^{\perp}$ *для подпространства* $L(a_1, a_2)$, *где* $a_1 = (1\ \ 1\ \ 1\ \ 1)^T$, $a_2 = (1\ \ -1\ \ 1\ \ 1)^T$.

Решение.

Пусть $x = (x_1\ x_2\ x_3\ x_4)^T$ *– элемент ортогонального дополнения.*

Система линейных уравнений

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 0; \\ x_1 - x_2 + x_3 + x_4 = 0 \end{cases}$$

определяет ортогональное дополнение.

Найдем ФСР данной системы:

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

$$\begin{cases} x_1 + x_3 + x_4 = 0; \\ x_2 = 0; \end{cases} \quad \begin{cases} x_1 = -x_3 - x_4; \\ x_2 = 0; \end{cases}$$

$$X^* = c_1 \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}; \quad c_1, c_2 \in \mathbb{R}; \quad c_1^2 + c_2^2 \neq 0.$$

$$e_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}; \ \ e_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}; \ \ L^{\perp}: \ L(e_1, e_2).$$

П р и м е р 13. *Найти ортогональную проекцию* $f = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$ *на подпространство* $L(b_1, b_2)$, *где* $b_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$.

Решение.

$$f = g + h; \quad g = \alpha_1 b_1 + \alpha_2 b_2;$$

$$f = \alpha_1 b_1 + \alpha_2 b_2 + h.$$

Умножаем последнее равенство скалярно, сначала на b_1, затем на b_2. Учитывая, что $(b_1, h) = 0, \quad (b_2, h) = 0$, так как они ортогональны, получаем

$$\begin{cases} (b_1, f) = \alpha_1(b_1, b_1) + \alpha_2(b_1, b_2), \\ (b_2, f) = \alpha_1(b_2, b_1) + \alpha_2(b_2, b_2); \end{cases}$$

$$\begin{cases} -3 = 2\alpha_1 - 3\alpha_2, \\ 9 = -3\alpha_1 + 6\alpha_2; \end{cases}$$

$$\alpha_1 = \alpha_2 = 3.$$

$$g = 3b_1 + 3b_2; \quad g = 3\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix}.$$

Понятие унитарного пространства

Комплексное линейное пространство $\mathcal{U}$ называется *комплексным евклидовым пространством* или *унитарным*, если выполняются следующие два требования:

I. Имеется правило, посредством которого двум элементам $x, y \in \mathcal{U}$ ставится в соответствие комплексное число называемое скалярным произведением и обозначаемое символом (x, y).

II. Указанное правило подчинено следующим четырем аксиомам:
$\forall x, y, z \in \mathcal{U}, \quad \forall \lambda \in$

1. $(x, y) = \overline{(y, x)}$;
2. $(x + y, z) = (x, z) + (y, z)$;
3. $(\lambda x, y) = \lambda(x, y)$;
4. $(x, x) \geq 0$, причем $(x, x) = 0 \Leftrightarrow x = \theta$.

Примеры комплексных евклидовых пространств.

1. Множество комплексно-значных функций $C^*_{[a,b]} : z(t) = x(t) + iy(t)$, где $x(t), \quad y(t)$ – вещественно-значные функции, непрерывные на отрезке $[a, b]$.

   Операции сложения этих функций и умножения их на комплексные числа заимствуем из анализа. Скалярное произведение двух любых таких функций определим соотношением
$$(z_1(t), z_2(t)) = \int_a^b z_1(t)\overline{z_2(t)}dt.$$

2. Множество вектор-столбцов с комплексными координатами
$$x, y \in \mathbb{C}^n, \quad x = (x_1, x_2, \ldots, x_n)^T, \; y = (y_1, y_2, \ldots, y_n)^T.$$

   Скалярное произведение введено следующим образом
$$(x, y) = \sum_{i=1}^{n} x_i \overline{y_i}.$$

П р и м е р 14. *Можно ли в унитарном пространстве квадратных матриц 2-ого порядка ввести скалярное произведение по формуле*
$(A, B) = a_1\overline{a_2} - b_1\overline{b_2} + c_1\overline{c_2} - d_1\overline{d_2}$, *где* $A = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$, $B = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$?

<u>Решение.</u> Проверим четвертую аксиому. $(A, A) = a_1\overline{a_1} - b_1\overline{b_1} + c_1\overline{c_1} - d_1\overline{d_1}$. *Неравенство* $(A, A) \geq 0$ *выполняется только тогда, когда*
$|a_1|^2 + |c_1|^2 \geq |b_1|^2 + |d_1|^2$. *Следовательно, ввести скалярное произведение по данной формуле нельзя.*

Свойства скалярного произведения.

1. $(x, y + z) = (x, y) + (x, z)$.
2. $(x, \lambda y) = \overline{\lambda}(x, y)$.
3. $(x, \theta) = 0$.
4. $(\sum\limits_{i=1}^{n} \alpha_i x_i, y) = \sum\limits_{i=1}^{n} \alpha_i (x_i, y), \quad \alpha_i \in \mathbb{C}, \ i = 1, 2, \ldots, n.$
5. Если $x, y \in \mathcal{U} : \forall z \in \mathcal{U}$ выполняется равенство $(x, z) = (y, z)$, то $x = y$.

Теорема 22 (Неравенство Коши-Буняковского в унитарном пространстве). *Для* $\forall x, y \in \mathcal{U}$*справедливо следующее неравенство*

$$|(x, y)|^2 \leq (x, x)(y, y).$$

Длина в унитарном пространстве вводится таким же образом как в вещественном евклидовом пространстве: $|x| = \sqrt{(x, x)}$. Понятие угла в унитарном пространстве вводить не имеет смысла, т.к. $(x, y) \in \mathbb{C}$.

Понятие ортогонального и ортонормированного базиса, процесс ортогонализации системы элементов, понятие ортогонального дополнения, ортогональной проекции элемента на подпространство без изменения определений и общих схем рассуждений переносится на унитарное пространство.

Выражение скалярного произведения через компоненты сомножителей в унитарном пространстве

Пусть в унитарном пространстве $\mathcal{U}$ задан некоторый базис
$e = (e_1, e_2, \ldots, e_n)$. Рассмотрим $x, y \in \mathcal{U}, \quad x = x_1 e_1 + x_2 e_2 + \ldots + x_n e_n$,
$y = y_1 e_1 + y_2 e_2 + \ldots + y_n e_n$,

$$(x, y) = (\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{y}_j (e_i, e_j). \qquad (10)$$

Формула (10) в матричном виде запишется следующим образом:

$$(x, y) = [x]_e^T \Gamma_e [\bar{y}]_e,$$

где Γ_e – матрица Грама

$$\Gamma_e = \begin{pmatrix} (e_1, e_1) & (e_1, e_2) & \dots & (e_1, e_n) \\ (e_2, e_1) & (e_2, e_2) & \dots & (e_2, e_n) \\ \dots & \dots & \dots & \dots \\ (e_n, e_1) & (e_n, e_2) & \dots & (e_n, e_n) \end{pmatrix}.$$

Если $e_1, e_2, \dots, e_n$ – ортонормированный базис, то $\Gamma_e = E$ и

$$(x, y) = [x]_e^T [\bar{y}]_e.$$

Пусть в унитарном пространстве $\mathcal{U}$ даны два базиса $e = (e_1, e_2, \dots, e_n)$ и $e' = (e'_1, e'_2, \dots, e'_n)$, тогда справедлива формула:

$$\Gamma_{e'} = P_{e \to e'}^T \Gamma_e \bar{P}_{e \to e'}.$$

Если $e_1, e_2, \dots, e_n$ – ортонормированный базис, то $\Gamma_{e'} = P_{e \to e'}^T \bar{P}_{e \to e'}$.

П р и м е р 15. *Векторы $x, y \in \mathcal{U}$, унитарного пространства заданы в базисе e_1, e_2 координатными столбцами $[x]_e = \begin{pmatrix} 1 & i \end{pmatrix}^T$, $[y]_e = \begin{pmatrix} 1+i & 2 \end{pmatrix}^T$, соответственно, и известна матрица Грама $\Gamma_f = \begin{pmatrix} 3 & 11i \\ -11i & 41 \end{pmatrix}$ базиса $f_1 = e_1 + ie_2$, $f_2 = -3ie_1 + 4e_2$. Вычислить матрицу Грама Γ_e базиса e_1, e_2 и скалярное произведение (x, y).*

Решение.

$P_{e \to f} = \begin{pmatrix} 1 & -3i \\ i & 4 \end{pmatrix}$; $P_{f \to e} = P_{e \to f}^{-1} = \begin{pmatrix} 4 & 3i \\ -i & 1 \end{pmatrix}$;

$\Gamma_e = P_{f \to e}^T \Gamma_f \bar{P}_{f \to e}$;

$\Gamma_e = \begin{pmatrix} 4 & -i \\ 3i & 1 \end{pmatrix} \begin{pmatrix} 3 & 11i \\ -11i & 41 \end{pmatrix} \begin{pmatrix} 4 & -3i \\ i & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

$(x, y) = [x]_e^T \Gamma_e [\bar{y}]_e = \begin{pmatrix} 1 & i \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1-i \\ 2 \end{pmatrix} = 1 + 3i$.

Унитарная матрица

Матрица R называется *унитарной*, если выполняется равенство

$$\bar{R}^T R = E.$$

Свойства унитарной матрицы:

1. $\det R = \pm 1$.

2. $\bar{R}^T = R^{-1}$.

3. Если R – унитарная матрица, то $\bar{R}^T$ – тоже унитарная матрица.

4. $R\bar{R}^T = E$.

5. Если R_1 и R_2 – унитарные матрицы, то $R_1 R_2$ – унитарная матрица.

6. Если R – унитарная матрица, то R^{-1} – тоже унитарная матрица.

Линейные операторы в линейных пространствах
Определение и простейшие свойства

Пусть $\mathcal{L}_n$ и $\mathcal{L}_m$ линейные пространства размерности m и n соответственно.

Оператором, действующим из $\mathcal{L}_n$ в $\mathcal{L}_m$, называется отображение вида $\varphi : \mathcal{L}_n \to \mathcal{L}_m$, которое каждому элементу $x \in \mathcal{L}_n$ ставит в соответствие элемент $y \in \mathcal{L}_m$.

Обозначения: $\varphi(x) = y, \quad \varphi x = y$,
где y – образ элемента x, а x – прообраз элемента y.

Оператор называется *линейным*, если $\forall x_1, x_2 \in \mathcal{L}_n, \forall \lambda \in \mathbb{R}(\mathbb{C})$ выполняются соотношения:

1) $\quad \varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2);$

2) $\quad \varphi(\lambda(x_1)) = \lambda\varphi(x_1).$

Если $\mathcal{L}_m$ представляет собой множество $\mathbb{R}(\mathbb{C})$, то линейный оператор называют *линейным функционалом* или *линейной формой.*

Обозначение: $f(x)$.

Линейный оператор, действующий из $\mathcal{L}_n$ в $\mathcal{L}_m$, иногда называют *линейным отображением.*

Если пространство $\mathcal{L}_m$ совпадает с пространством $\mathcal{L}_n$, то линейный оператор $\varphi : \mathcal{L}_n \to \mathcal{L}_n$ называют *линейным преобразованием* пространства $\mathcal{L}_n$.

Два оператора φ и ψ называются *равными*, если

$$\varphi(x) = \psi(x), \quad \forall x \in \mathcal{L}_n.$$

Примеры линейных операторов.

1. Оператор (преобразование) $\varepsilon : \mathcal{L}_n \to \mathcal{L}_n$, который каждый элемент $x \in \mathcal{L}_n$ переводит в x, является линейным и называется *тождественным оператором.*

2. Оператор $\Theta : \mathcal{L}_n \to \mathcal{L}_m$, который каждый элемент $x \in \mathcal{L}_n$ переводит в нулевой элемент $\theta \in \mathcal{L}_m$, является линейным и называется *нулевым оператором.*

3. Пусть P_n – пространство вещественных многочленов степени не выше n. Оператор $\varphi : P_n \to P_{n-1}$, определенный правилом $\varphi(p(x)) = p'(x)$, где $p(x) \in P_n$, является линейным и называется *оператором дифференцирования.*

4. Изоморфизм φ линейных пространств $\mathcal{L}_n$ и $\mathcal{L}'_n$ является линейным оператором, действующим из $\mathcal{L}_n$ в $\mathcal{L}'_n$.

5. Растяжение (сжатие) элементов пространства $\mathcal{L}_n$ в одно и то же число α раз является оператором в пространстве $\mathcal{L}_n$. Такой оператор называется *оператором подобия*: $\psi : \mathcal{L}_n \to \mathcal{L}_n,\ \psi x = \alpha x,\ x \in \mathcal{L}_n$.

Простейшие свойства линейного оператора.

Из определения вытекают следующие свойства линейных операторов.

1. Линейный оператор переводит нулевой элемент в нулевой элемент: $\varphi(\theta_1) = \varphi(0 \cdot x) = 0 \cdot \varphi(x) = \theta_2;\ \theta_1 \in \mathcal{L}_n,\quad \theta_2 \in \mathcal{L}_m$.

2. Линейный оператор сохраняет линейную комбинацию, т.е. переводит линейную комбинацию элементов в линейную комбинацию образов с теми же коэффициентами:
$$\varphi\left(\sum_{i=1}^{n} \alpha_i x_i\right) = \sum_{i=1}^{n} \alpha_i \varphi x_i.$$

3. Линейный оператор переводит линейно зависимую систему элементов в линейно зависимую.

Задание линейного оператора.

Свойство 2° говорит о том, что для задания линейного оператора $\varphi : \mathcal{L}_n \to \mathcal{L}_m$ достаточно определить его только на элементах $e_1, e_2, \cdots, e_n$ некоторого базиса пространства $\mathcal{L}_n$. Зная элементы $\varphi e_1, \varphi e_2, \cdots, \varphi e_n$ можно однозначно найти образ любого элемена $x = \sum\limits_{i=1}^{n} x_i e_i \in \mathcal{L}_n$:
$$\varphi x = \sum_{i=1}^{n} x_i \varphi e_i \in \mathcal{L}_m.$$

Теорема 1. *Пусть $e_1, e_2, \cdots, e_n$ – базис линейного пространства $\mathcal{L}_n$, а $g_1, g_2, \cdots, g_n$ – произвольные элемены линейного пространства $\mathcal{L}_m$. Тогда существует единственный оператор $\varphi : \mathcal{L}_n \to \mathcal{L}_m$, который переводит элементы $e_1, e_2, \cdots, e_n$ линейного пространства $\mathcal{L}_n$ в элементы $g_1, g_2, \cdots, g_n$ линейного пространства $\mathcal{L}_m$ соответственно.*

Д о к а з а т е л ь с т в о. Построим искомый оператор, положив для каждого элемента $x = \sum\limits_{i=1}^{n} x_i e_i \in \mathcal{L}_n$
$$\varphi x = \sum_{i=1}^{n} x_i g_i. \tag{1}$$

Из единственности разложения элемента x по базису следует, что правило (1) однозначно оределяет образ элемента x, при этом, как легко проверить,

$$\varphi e_i = g_i,\ i = 1, 2, \ldots, n.$$

Линейность построенного оператора вытекает из линейности координат. Оператор φ единственный, так как если ψ любой другой линейный оператор, переводящий элемены $e_1, e_2, \ldots, e_n$ в $g_1, g_2, \ldots, g_n$, то
$\psi x = \psi(\sum\limits_{i=1}^{n} x_i e_i) = \sum\limits_{i=1}^{n} x_i \psi e_i = \sum\limits_{i=1}^{n} x_i g_i = \varphi x,\ \forall x \in \mathcal{L}_n \Rightarrow \varphi = \psi.$ $\square$

Следствие. Два оператора $\varphi,\ \psi : \mathcal{L}_n \to \mathcal{L}_m$ равны тогда и только тогда, когда они одинаково определены на элементах базиса $\mathcal{L}_n$.

Матрица линейного оператора

Пусть $e_1, e_2, \cdots, e_n$ – базис линейного пространства $\mathcal{L}_n$,
а $f_1, f_2, \cdots, f_m$ – базис линейного пространства $\mathcal{L}_m$.
По теореме из предыдущего параграфа оператор $\varphi : \mathcal{L}_n \to \mathcal{L}_m$ однозначно определяется заданием элементов $\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n)$, которые однозначно определяются своими координатами в базисе f, т.е. коэффициентами разложений

$$\begin{cases} \varphi(e_1) = a_{11}f_1 + a_{21}f_2 + \ldots + a_{m1}f_m, \\ \varphi(e_2) = a_{12}f_1 + a_{22}f_2 + \ldots + a_{m2}f_m, \\ \qquad \cdots \quad \cdots \quad \cdots \\ \varphi(e_n) = a_{1n}f_1 + a_{2n}f_2 + \ldots + a_{mn}f_m. \end{cases} \tag{2}$$

Матрица

$$[\varphi]_{fe} = A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

называется *матрицей оператора* φ в паре базисов e и f.

Координаты элемента и его образа.

Пусть $\varphi : \mathcal{L}_n \to \mathcal{L}_m$ и $e_1, e_2, \cdots, e_n$ – базис линейного пространства $\mathcal{L}_n$,
а $f_1, f_2, \cdots, f_m$ – базис линейного пространства $\mathcal{L}_m$.

Теорема 2. *Если $y = \varphi x$, то справедливо равенство*

$$y_f = [\varphi]_{fe} x_e. \tag{3}$$

Доказательство. Пусть $x = \sum\limits_{i=1}^{n} x_i e_i$, $y = \sum\limits_{i=1}^{n} y_i f_i$ и $[\varphi]_{fe} = A = (a_{ij})$.
Утверждение (3) равносильно соотношениям

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \cdots, m.$$

Докажем их. Имеем $y = \varphi x = \varphi(\sum\limits_{j=1}^{n} x_j e_j) = \sum\limits_{j=1}^{n} x_j \varphi e_j = \sum\limits_{j=1}^{n} x_j \sum\limits_{i=1}^{m} a_{ij} f_i = \sum\limits_{i=1}^{m} (\sum\limits_{j=1}^{n} a_{ij} x_j) f_i$.

Из единственности разложения элемента y по базису f следует (3).

Пример 1. *Пусть* $\varphi : P_1 \to P_2$;
$\varphi(p(x)) = (x+1)p(x)$;
$e_1 = 1,\ e_2 = x;\quad f_1 = 1,\ f_2 = x,\ f_3 = x^2$.
Найти:

1. *Матрицу линейного оператора.*

2. *Проверить* $[\varphi]_{fe} y_e = [\varphi(y)]_f, \quad \forall y = p(x) \in P_1.$

<u>*Решение.*</u>

1. $\quad \varphi e_1 = (x+1) \cdot 1 = x+1, \qquad [\varphi e_1]_f = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$

$$\varphi e_2 = (x+1) \cdot x = x^2 + x, \quad [\varphi e_2]_f = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad\qquad [\varphi e]_{fe} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

2. $\quad y = a + bx, \qquad y_e = \begin{pmatrix} a \\ b \end{pmatrix};$

$$\varphi(y) = (x+1)(a+bx) = ax + a + bx^2 + bx = bx^2 + (a+b)x + a;$$

$$[\varphi(y)]_f = \begin{pmatrix} a \\ a+b \\ b \end{pmatrix}.$$

$$[\varphi]_{fe} y_e = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ a+b \\ b \end{pmatrix} = [\varphi(y)]_f.$$

Матрицы оператора в различных базисах.

Пусть e и $e' = e \cdot P_{e \to e'}$ – два базиса в пространстве $\mathcal{L}_n$ с матрицей перехода $P_{e \to e'}$, а f и $f' = f \cdot P_{f \to f'}$ – два базиса пространства $\mathcal{L}_m$ с матрицей перехода $P_{f \to f'}$.

Одному и тому же оператору $\varphi : \mathcal{L}_n \to \mathcal{L}_m$ в паре базисов e и f соответствует матрица $[\varphi]_{fe}$, а в паре базисов e' и f' соответствует матрица $[\varphi]_{f'e'}$.

Теорема 3. *Матрицы линейного оператора в различных парах базисов связаны соотношением*

$$[\varphi]_{f'e'} = P^{-1}_{f \to f'} [\varphi]_{fe} P_{e \to e'}. \tag{4}$$

Д о к а з а т е л ь с т в о. Для произвольного элемента $x \in \mathcal{L}_n$ и его образа $y = \varphi x$ в силу $y_f = [\varphi]_{fe} x_e$ имеем

$$y_f = [\varphi]_{fe} x_e, \ y_{f'} = [\varphi]_{f'e'} x_{e'}. \tag{5}$$

В свою очередь,

$$x_e = P_{e \to e'} x_{e'}, \ y_f = P_{f \to f'} y_{f'}.$$

Подставив эти соотношения в (5), получим, что

$$P_{f \to f'} y_{f'} = [\varphi]_{fe} P_{e \to e'} x_{e'}$$

или

$$P_{f \to f'} [\varphi]_{f'e'} x_{e'} = [\varphi]_{fe} P_{e \to e'} x_{e'}.$$

Так как это соотношение имеет место для любого $x_{e'}$, то

$$P_{f \to f'} [\varphi]_{f'e'} = [\varphi]_{fe} P_{e \to e'}.$$

В силу невырожденности матрицы перехода отсюда следует (4). $\square$

Две прямоугольные одного размера матрицы A и B называются *эквивалентными*, если существуют такие две невырожденные матрицы Q и P, что

$$B = Q^{-1} A P.$$

Следствие 1. Матрицы линейного операторав в различных парах базисов являются эквивалентными.

Следствие 2. Ранг матрицы линейного оператора не зависит от выбора базисов.

Следствие 3. Если оператор действует в одном пространстве (является преобразованием), то формула (4) будет иметь вид

$$[\varphi]_{e'} = P^{-1}_{e \to e'} [\varphi]_e P_{e \to e'}.$$

Две квадратные матрицы A и B одинаковых размеров называются *подобными*, если существует невырожденная матрица P, такая что справедливо равенство

$$B = P^{-1}AP.$$

Теорема 5. *Определители подобных матриц равны.*

Д о к а з а т е л ь с т в о. Если матрицы A и B подобны, то согласно определению существует такая невырожденная матрица P, что $B = P^{-1}AP$.

Учитывая свойства определителя, получаем $\det(B) = \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P) = \det(P^{-1}P)\det(A) = \det(A)$. $\square$

Следствие 4. Матрицы линейного преобразования в различных базисах имеют равные определители.

Линейное пространство линейных операторов

Обозначим $L(\mathcal{L}_n, \mathcal{L}_m)$ множество линейных операторов, действующих из $\mathcal{L}_n$ в $\mathcal{L}_m$.

Суммой линейных операторов $\varphi, \psi \in L(\mathcal{L}_n, \mathcal{L}_m)$ будем называть оператор $\varphi + \psi : \mathcal{L}_n \to \mathcal{L}_m$, определяемый формулой

$$(\varphi + \psi)x = \varphi x + \psi x, \quad \forall x \in \mathcal{L}_n. \tag{6}$$

Произведением линейного оператора $\varphi \in L(\mathcal{L}_n, \mathcal{L}_m)$ *на число* $\alpha \in \mathbb{R}(\mathbb{C})$ будем называть оператор $\alpha\varphi : \mathcal{L}_n \to \mathcal{L}_m$, такой что

$$(\alpha\varphi)x = \alpha\varphi x, \quad \forall x \in \mathcal{L}_n.$$

Теорема 6. *Для любых операторов* $\varphi, \psi \in L(\mathcal{L}_n, \mathcal{L}_m)$ *и числа* $\alpha \in \mathbb{R}(\mathbb{C})$

$$\varphi + \psi \in L(\mathcal{L}_n, \mathcal{L}_m), \quad \alpha\varphi \in L(\mathcal{L}_n, \mathcal{L}_m).$$

Доказательство. Для $\forall x, y \in \mathcal{L}_n$ согласно (6) Имеем

$$(\varphi + \psi)(x + y) = \varphi(x + y) + \psi(x + y).$$

В силу линейности φ, ψ и аксиом линейного пространства

$$(\varphi+\psi)(x+y) = (\varphi x+\varphi y)+(\psi x+\psi y) = (\varphi x+\psi x)+(\varphi y+\psi y) = (\varphi+\psi)x+(\varphi+\psi)y.$$

Для $\forall x \in \mathcal{L}_n, \ \lambda \in \mathbb{R}(\mathbb{C})$
$(\varphi + \psi)(\lambda x) = \lambda((\varphi + \psi)x) \Rightarrow \varphi + \psi \in L(\mathcal{L}_n, \mathcal{L}_m)$.

Аналогично доказывается, что $\alpha\varphi \in L(\mathcal{L}_n, \mathcal{L}_m)$. $\square$

Теорема 7. *Множество линейных операторов* $L(\mathcal{L}_n, \mathcal{L}_m)$ *является линейным пространством относительно введенных выше операций.*

Доказательство. Достаточно проверить аксиомы линейного пространства, взяв в качестве нулевого элемента нулевое отображение $O \in L(\mathcal{L}_n, \mathcal{L}_m)$, а в качестве противоположного к оператору φ отображение $-\varphi \in L(\mathcal{L}_n, \mathcal{L}_m)$, выполняемое по правилу

$$(-\varphi)x = -\varphi x, \ \forall x \in \mathcal{L}_n.$$

Все аксиомы вытекают из соответствующих аксиом линейного пространства, примененных к $\mathcal{L}_n$ и $\mathcal{L}_m$ и проверяются по единой схеме.

Проверим, например, коммутативность. Для $\forall \varphi, \psi \in L(\mathcal{L}_n, \mathcal{L}_m)$ и $\forall x \in \mathcal{L}_n$
$(\varphi + \psi)x = \varphi x + \psi x = \psi x + \varphi x$;
$(\psi + \varphi)x = \psi x + \varphi x$.

Таким образом, $\psi + \varphi = \varphi + \psi$. $\square$

Теорема 8. *Если* $\dim(\mathcal{L}_n) = n$, *а* $\dim(\mathcal{L}_m) = m$, *то линейное пространство операторов* $L(\mathcal{L}_n, \mathcal{L}_m)$ *изоморфно пространству матриц* $\mathbb{R}^{m\times n}(\mathbb{C}^{m\times n})$.

Следствие. $\dim L(\mathcal{L}_n, \mathcal{L}_m) = \dim(\mathcal{L}_n)\cdot\dim(\mathcal{L}_m)$.

Замечание. Так как линейное пространство линейных операторов $L(\mathcal{L}_n, \mathcal{L}_m)$ изоморфно пространству матриц $\mathbb{R}^{m\times n}(\mathbb{C}^{m\times n})$, то при сложении линейных операторов их матрицы складываются, а при умножении линейного оператора на число его матрица умножается на это же число.

1.4 Умножение линейных операторов

Пусть $\varphi : \mathcal{L}_n \to \mathcal{L}_m$ и $\psi : \mathcal{L}_m \to L_k$.

Произведением линейных операторов φ, ψ будем называть оператор $\psi\varphi : \mathcal{L}_n \to L_k$, действующий по следующему правилу

$$(\psi\varphi)x = \psi(\varphi(x)), \quad \forall x \in \mathcal{L}_n.$$

Теорема 9. *Если* $\varphi \in L(\mathcal{L}_n, \mathcal{L}_m)$ *и* $\psi \in L(\mathcal{L}_m, \mathcal{L}_k)$, *то* $\psi\varphi \in L(\mathcal{L}_n, \mathcal{L}_k)$.

Д о к а з а т е л ь с т в о. Для $\forall x, y \in \mathcal{L}_n,\ \alpha \in \mathbb{R}(\mathbb{C})$
$(\psi\varphi)(x+y) = \psi(\varphi(x+y)) = \psi(\varphi x+\varphi y) = \psi(\varphi x)+\psi(\varphi y) = (\psi\varphi)x+(\psi\varphi)y$;
$(\psi\varphi)(\alpha x) = \psi(\varphi(\alpha x)) = \psi(\alpha(\varphi x) = \alpha\psi(\varphi x) = \alpha(\psi\varphi)x$. $\square$

Свойства произведения линейных операторов.

Произведение линейных операторов определено не для любой пары линейных операторов. Однако, если это произведение имеет смысл, то:

1. $\alpha(\varphi\psi) = (\alpha\varphi)\psi = \varphi(\alpha\psi), \quad \forall\alpha \in \mathbb{R}(\mathbb{C})$.
2. $(\varphi + \psi)\xi = \varphi\xi + \psi\xi$.
   $\varphi(\psi + \xi) = \varphi\psi + \varphi\xi$.
3. $(\varphi\psi)\xi = \varphi(\psi\xi)$.

Д о к а з а т е л ь с т в о.

1. Следует из определения линейного оператора на скаляр и определения произведения операторов.
2. $((\varphi+\psi)\xi)x = (\varphi+\psi)(\xi x) = \varphi(\xi x)+\psi(\xi x) = (\varphi\xi)x+(\psi\xi)x = (\varphi\xi+\psi\xi)x$.
3. Согласно определению произведение линейных операторов заключается в их последовательном действии, и поэтому операторы $(\varphi\psi)\xi$ и $\varphi(\psi\xi)$ совпадают и, следовательно, тождественны.

Умножение линейных операторов не обладает свойством коммутативности. В самом деле, о коммутативности можно говорить лишь для линейных преобразований. Но и в этом случае умножение не коммутативно.

Теорема 10. *При умножении линейных операторов их матрицы умножаются, т. е. если* e, f, g *– базисы пространств* $\mathcal{L}_n,\ \mathcal{L}_m,\ L_k$ *соответственно и* $\varphi : \mathcal{L}_n \to \mathcal{L}_m, \quad \psi : \mathcal{L}_m \to L_k$, *то*

$$[\psi\varphi]_{ge} = [\psi]_{gf}[\varphi]_{fe}. \tag{7}$$

Д о к а з а т е л ь с т в о. Пусть $[\varphi]_{fe} = (a_{ij}), \quad [\psi]_{gf} = (b_{ij}), \quad [\psi\varphi]_{ge} = (c_{ij})$, $\dim \mathcal{L}_n = n,\ \dim \mathcal{L}_m = m,\ \dim \mathcal{L}_k = k$.

Тогда

$$\psi\varphi e_j = \sum_{i=1}^{k} c_{ij} g_i. \tag{8}$$

$\psi\varphi e_j = \psi(\varphi e_j) = \psi(\sum\limits_{s=1}^{m} a_{sj} f_s) = \sum\limits_{s=1}^{m} a_{sj}(\psi f_s) = \sum\limits_{s=1}^{m} a_{sj} \sum\limits_{i=1}^{k} b_{is} g_i =$
$= \sum\limits_{s=1}^{m}\sum\limits_{i=1}^{k} a_{sj} b_{is} g_i = \sum\limits_{i=1}^{k}(\sum\limits_{s=1}^{m} b_{is} a_{sj}) g_i.$

Сравнение этого разложения с (8) приводит к равенству $c_{ij} = \sum\limits_{s=1}^{m} b_{is} a_{sj}$, которое означает (7). □

Обратный оператор

Пусть $\varphi \in L(\mathcal{L}_n, \mathcal{L}_n)$.

Отображение $\varphi^{-1} : \mathcal{L}_n \to \mathcal{L}_n$ называется *обратным оператором* к оператору φ, если

$$\varphi\varphi^{-1} = \varphi^{-1}\varphi = \varepsilon, \tag{9}$$

где ε – тождественный оператор.

Из определения обратного оператора φ^{-1} следует, что для $\forall x \in \mathcal{L}_n$ справедливо соотношение

$$\varphi^{-1}\varphi x = x.$$

Таким образом, если $\varphi^{-1}\varphi x = \theta$, то $x = \theta$, т.е., если оператор имеет обратный, то из условия $\varphi x = \theta$ следует, что $x = \theta$.

Теорема 11. *Для того, чтобы линейный оператор $\varphi \in L(\mathcal{L}_n, \mathcal{L}_n)$ имел обратный, необходимо и достаточно, чтобы этот оператор действовал взаимно однозначно из $\mathcal{L}_n$ в $\mathcal{L}_n$.*

Д о к а з а т е л ь с т в о. *Необходимость.* Пусть φ имеет обратный, но не действует взаимно однозначно из $\mathcal{L}_n$ в $\mathcal{L}_n$. Это означает, что некоторым различным элементам x_1 и x_2, $x_2 - x_1 \neq \theta \in \mathcal{L}_n$ отвечает один и тот же элемент $y = \varphi x_1 = \varphi x_2$. Но тогда $\varphi(x_2 - x_1) = \theta$ и поскольку φ имеет обратный,
$x_2 - x_1 = \theta$. Но выше было отмечено, что $x_2 - x_1 \neq \theta$. Полученное противоречие доказывает необходимость условия утверждения.
Достаточность. Допустим φ действует взаимно однозначно из $\mathcal{L}_n$ в $\mathcal{L}_n$. Тогда каждому элементу $y \in \mathcal{L}_n$ отвечает элемент $x \in \mathcal{L}_n : y = \varphi x$.
Поэтому имеется оператор φ^{-1}, обладающий тем свойством, что
$\varphi^{-1}y = \varphi^{-1}(\varphi x) = x$.

Легко убедиться, что φ^{-1} линейный. Пусть $\forall y_1, y_2 \in \mathcal{L}_n$, $\exists x_1, x_2 \in \mathcal{L}_n$: $y_1 = \varphi x_1$, $y_2 = \varphi x_2$, при этом $x_1 = \varphi^{-1}y_1$, $x_2 = \varphi^{-1}y_2$.

Отсюда получим, что $\varphi^{-1}(y_1 + y_2) = \varphi^{-1}(\varphi x_1 + \varphi x_2) = \varphi^{-1}\varphi(x_1 + x_2) = x_1 + x_2 = \varphi^{-1}y_1 + \varphi^{-1}y_2$.

Аналогично $\varphi^{-1}(\alpha y_1) = \varphi^{-1}(\alpha\varphi x_1) = \varphi^{-1}\varphi(\alpha x_1) = \alpha x_1 = \alpha\varphi^{-1}y_1$,
$\forall\alpha \in \mathbb{R}(\mathbb{C})$. $\square$

Теорема 12. *Матрица обратного оператора φ^{-1} в произвольном базисе является обратной к матрице оператора φ в этом же базисе.*

Д о к а з а т е л ь с т в о. Пусть e – произвольный базис пространства $\mathcal{L}_n$ и для оператора $\varphi \in L(\mathcal{L}_n, \mathcal{L}_n)$ существует обратный оператор φ^{-1}. Перейдем в равенствах (9) к матрицам операторов в базисе e. Согласно теореме 10, получим, что $[\varphi]_e[\varphi^{-1}]_e = [\varphi^{-1}]_e[\varphi]_e = E$. Эти равенства совпадают с определением обратной матрицы для $[\varphi]_e$. $\square$

Образ и ядро линейного оператора

Образом линейного оператора $\varphi : \mathcal{L}_n \to \mathcal{L}_m$ называется множество всех элементов $y \in \mathcal{L}_m$, представляемых в виде $y = \varphi(x), \quad x \in \mathcal{L}_n$.

Обозначение: $\operatorname{im}\varphi$.

Ядром линейного оператора $\varphi : \mathcal{L}_n \to \mathcal{L}_m$ называется множество всех элементов $x \in \mathcal{L}_n$, для которых $\varphi(x) = \theta, \quad \theta \in \mathcal{L}_m$.

Обозначение: $\ker\varphi$.

Теорема 13. Если $\varphi \in L(\mathcal{L}_n, \mathcal{L}_m)$, то $\operatorname{im}\varphi$ – линейное подпространство $\mathcal{L}_n$; $\ker\varphi$ – линейное подпространство $\mathcal{L}_m$.

Д о к а з а т е л ь с т в о.

1. Докажем, что $\operatorname{im}\varphi$ – линейное подпространство $\mathcal{L}_n$. Так как $y_1 \in \operatorname{im}\varphi,\ y_2 \in \operatorname{im}\varphi \Rightarrow \exists x_1, x_2 \in \mathcal{L}_n$, что $y_1 = \varphi x_1,\ y_2 = \varphi x_2$,

$$y_1 + y_2 = \varphi x_1 + \varphi x_2 = \varphi(x_1 + x_2);$$

$$\lambda y_1 = \lambda(\varphi x_1) = \varphi(\lambda x_1).$$

2. Докажем, что $\ker\varphi$ – линейное подпространство $\mathcal{L}_m$.
$x_1 \in \ker\varphi,\ \varphi x_1 = \theta$; $x_2 \in \ker\varphi,\ \varphi x_2 = \theta$.

$$\varphi(x_1 + x_2) = \varphi x_1 + \varphi x_2 = \theta + \theta = \theta;$$

$$\varphi(\lambda x_1) = \lambda\varphi(x_1) = \lambda\theta = \theta. \ \square$$

Число $\dim(\operatorname{im}\varphi) = \operatorname{rg}\varphi$ называется *рангом линейного оператора*, а $\dim(\ker\varphi) = \operatorname{defekt}\varphi$ называется *дефектом линейного оператора.*

Нулевой оператор $\Theta x = \theta$ и тождественный оператор $\varepsilon x = x$ являются предельными с точки зрения дефекта и ранга. Нулевой оператор имеет максимальный дефект равный размерности пространства, в котором этот оператор действует и минимальный ранг. Тождественный оператор имеет минимальный дефект (нулевой) и максимальный ранг равный размерности пространства, в котором этот оператор действует.

Оператор максимального дефекта определен однозначно, а операторов минимального дефекта и максимального ранга бесконечно много.

Теорема 14. *Пусть* $\varphi : \mathcal{L}_n \to \mathcal{L}_m$. *Если* $e_1, e_2, \cdots, e_n$ *– базис в* $\mathcal{L}_n$, *то*

$$\operatorname{im}\varphi = L(\varphi e_1, \varphi e_2, \cdots, \varphi e_n). \tag{10}$$

Д о к а з а т е л ь с т в о. Достаточно показать, что для множеств (10) имеет место двусторонние вложение:

с одной стороны, если $y \in \operatorname{im}\varphi$, то $y = \varphi x$ для некоторого элемента $x \in \mathcal{L}_n$, т.е. $y = \varphi(\sum\limits_{i=1}^{n} x_i e_i) = \sum\limits_{i=1}^{n} x_i \varphi e_i \in L(\varphi e_1, \varphi e_2, \ldots, \varphi e_n)$;

с другой стороны, если $y \in L(\varphi e_1, \varphi e_2, \dots, \varphi e_n)$, то $y = \sum\limits_{i=1}^{n} x_i \varphi e_i = \varphi(\sum\limits_{i=1}^{n} x_i e_i) = \varphi x$, где $x = \sum\limits_{i=1}^{n} x_i e_i$, т.е. $y \in \operatorname{im}\varphi$. $\square$

Теорема 15. *Ранг линейного оператора равен рангу его матрицы в произвольной паре базисов.*

Доказательство. Из теоремы 14 и $\dim L(x, y, \dots, z) = \operatorname{rg}(x, y, \dots, z)$ следует, что $\operatorname{rg}\varphi = \dim \operatorname{im}\varphi = \dim L(\varphi e_1, \varphi e_2, \dots, \varphi e_n) = \operatorname{rg}(\varphi e_1, \varphi e_2, \dots, \varphi e_n)$. Ранг системы элементов $\varphi e_1, \varphi e_2, \dots, \varphi e_n$ совпадает с рангом системы элементов, состоящих из координат этих элементов в базисе f пространства $\mathcal{L}_m$, т.е. с рангом системы столбцов матрицы $[\varphi]_{fe}$. $\square$

Теорема 16. *Если $\varphi \in L(\mathcal{L}_n, \mathcal{L}_m)$, то*

$$\operatorname{rg}\varphi + \operatorname{def}\varphi = \dim(\mathcal{L}_n). \tag{11}$$

Доказательство. Пусть $e_1, e_2, \dots, e_k$ – базис $\ker\varphi$. Дополним его до базиса $e_1, e_2, \dots, e_k, e_{k+1}, \dots, e_n$ пространства $\mathcal{L}_n$. Согласно теореме 14 $\operatorname{im}\varphi = L(\varphi e_1, \varphi e_2, \cdots, \varphi e_k) = L(\varphi e_{k+1}, \varphi e_{k+2}, \cdots, \varphi e_n)$.

Докажем, что элементы $\varphi e_{k+1}, \varphi e_{k+2}, \cdots, \varphi e_n$ линейно независимы. Пусть это не так. Тогда для нетривиальной линейной комбинации этих элементов имеет место соотношение

$$\alpha_{k+1}\varphi e_{k+1} + \ldots + \alpha_n \varphi e_n = \theta;$$

$$\varphi(\alpha_{k+1} e_{k+1} + \ldots + \alpha_n e_n) = \theta.$$

Следовательно, $\alpha_{k+1} e_{k+1} + \ldots + \alpha_n e_n \in \ker\varphi$. Это означает, что элемент $\alpha_{k+1} e_{k+1} + \ldots + \alpha_n e_n$ линейно выражается через $e_1, e_2, \dots, e_k$, что невозможно в силу линейной независимости $e_1, e_2, \dots, e_k, e_{k+1}, \dots, e_n$.

Таким образом, $\dim \operatorname{im}\varphi = n - k$, $\dim\ker\varphi = k$. Отсюда следует (11). $\square$

Пример 2. *Для линейного преобразования*

$$\varphi x = (x_1 - x_2 + x_3,\ x_2 + 2x_3,\ x_1 + 3x_3)^T, \qquad x = (x_1, x_2, x_3)^T.$$

Найти:

1) $[\varphi]_e,\ e_1 = (1\ 0\ 0)^T,\ e_2 = (0\ 1\ 0)^T,\ e_3 = (0\ 0\ 1)^T$;

2) $\operatorname{defekt}\varphi, \quad \operatorname{rg}\varphi$;

3) $\ker\varphi, \quad \operatorname{im}\varphi$;

4) базисы ядра и образа.

Решение. *По условию* $\varphi(x) = \varphi \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 + x_3 \\ x_2 + 2x_3 \\ x_1 + 3x_3 \end{pmatrix}.$

1. $[\varphi(e_1)]_e = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = a_1$, $[\varphi(e_2)]_e = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = a_2$, $[\varphi(e_3)]_e = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = a_3$,
   $[\varphi]_e = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{pmatrix}.$

2. $\operatorname{defekt}\varphi + \operatorname{rg}\varphi = 3$;
   $\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \Rightarrow rg\varphi = 2 \Rightarrow \operatorname{defekt}\varphi = 3 - 2 = 1.$

3. Согласно теореме 14 $\operatorname{im}\varphi = L(\varphi e_1, \varphi e_2, \varphi e_3)$. Это означает, что $\operatorname{im}\varphi$ совпадает с линейной оболочкой системы столбцов матрицы $[\varphi]_e$ и, следовательно, за базис $\operatorname{im}\varphi$ можно взять любой из базисов системы столбцов матрицы $[\varphi]_e$, например, a_1, a_2, получим, что $\operatorname{im}\varphi = L(a_1, a_2)$.

   Аналогично, $x \in \ker\varphi$ в том и только в том случае, когда $\varphi(x) = \theta$ или в координатной форме
   $$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \qquad (12)$$

   Отсюда следует, что $\ker\varphi$ совпадает с подпространством решений однородной системы (12), и в качестве базиса в $\ker\varphi$ может быть выбрана фундаментальная система решений уравнений (12). Найдем решение. Преобразуем матрицу системы
   $\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix},$

   $\begin{cases} x_1 = -3x_3 \\ x_2 = -2x_3 \end{cases}, \quad x_3 = \alpha, \quad \alpha \in \mathbb{R}, \quad x = \begin{pmatrix} -3\alpha \\ -2\alpha \\ \alpha \end{pmatrix}$, если $\alpha = 1$, то получим

   $b_1 = \begin{pmatrix} -3 \\ -2 \\ 1 \end{pmatrix}$, $\ker\varphi = L(b_1)$, $\quad b_1$ – базисный вектор ядра.

Собственные значения и собственные векторы линейного оператора

Характеристический многочлен.

Для произвольной матрицы $A \in \mathbb{R}^{n\times n}$ рассмотрим

$$\det(A - \lambda E) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix},$$

где E – единичная матрица и $\lambda \in \mathbb{R}$.

Относительно переменной λ этот определитель является многочленом степени n и может быть записан в виде

$$f(\lambda) = \det(A - \lambda E) = \sum_{i=0}^{n} \alpha_i \lambda^i. \tag{13}$$

Многочлен $f(\lambda) = \det(A - \lambda E)$ называется *характеристическим многочленом матрицы* A, а уравнение $f(\lambda) = 0$ – *характеристическим уравнением матрицы* A,

$$\alpha_0 = f(0) = \det A,$$

$$\alpha_{n-1} = a_{11} + a_{22} + \ldots + a_{nn} = \mathrm{tr} A.$$

Теорема 17. *Характеристические многочлены (уравнения) подобных матриц совпадают.*

Доказательство. Пусть A и B подобные матрицы, т. е. $B = P^{-1}AP$, тогда в силу свойств определителей имеем

$$f_B(\lambda) = \det(B - \lambda E) = \det(P^{-1}AP - \lambda P^{-1}EP) = \det(P^{-1}(A - \lambda E)P) =$$

$$= \det P^{-1} \det(A - \lambda E) \det P = \det(A - \lambda E) = f_A(\lambda). \ \Box$$

Рассмотрим линейный оператор (преобразование) $\varphi \in L(\mathcal{L}_n \to \mathcal{L}_n)$ и тождественный оператор $\varepsilon \in L(\mathcal{L}_n \to \mathcal{L}_n)$.

Характеристическим многочленом оператора называется функция

$$f(\lambda) = \det(\varphi - \lambda\varepsilon), \ \lambda \in \mathbb{R}.$$

Так как $\det\varphi = \det[\varphi]_e$, где e – базис в $\mathcal{L}_n$, то характеристический многочлен оператора совпадает с характеристическим многочленом матрицы этого оператора в произвольном базисе. При этом коэффициенты α_k характеристического многочлена, представляемого в виде (13), также не связаны с использованным базисом, т.е. являются *инвариантами* относительно выбора базиса.

Уравнение $\det(\varphi - \lambda\varepsilon) = 0$ называется *характеристическим уравнением оператора* φ.

Пусть $\mathcal{L}'$ – подпространство n-мерного линейного пространства $\mathcal{L}_n$ и $\varphi \in L(\mathcal{L}_n, \mathcal{L}_n)$.

Линейное подпространство $\mathcal{L}'$ пространства $\mathcal{L}_n$ называется *инвариантным подпространством относительно оператора* φ, если для $\forall x \in \mathcal{L}'$ его образ $\varphi x \in L^{'}$.

Примеры инвариантных подпространств.

1. Тривиальные подпространства $\{\theta\}$ и $\mathcal{L}_n$ инвариантны относительно любого оператора $\varphi \in L(\mathcal{L}_n, \mathcal{L}_n)$.

2. Для любого линейного оператора φ инвариантными подпространствами будут $\ker\varphi$ и $\operatorname{im}\varphi$, так как если $\varphi x = \theta$, то $\varphi(\varphi x) = \varphi\theta = \theta$ и если $y = \varphi x$, то $\varphi y = \varphi(\varphi x) = \varphi x_1$, где $x_1 = \varphi x$.

Собственные значения и собственные векторы линейного оператора

Число λ называется *собственным значением линейного оператора* $\varphi \in L(\mathcal{L}_n, \mathcal{L}_n)$, если $\exists x \neq \theta$:

$$\varphi x = \lambda x. \tag{14}$$

При этом элемент x называется *собственным вектором оператора* φ.

Множество всех собственных значений линейного оператора называется *спектром линейного оператора.*

1. Каждый собственный вектор связан со своим собственным значением.

   Д о к а з а т е л ь с т в о. Если x одновременно удовлетворяет двум равенствам $\varphi x = \lambda x$ и $\varphi x = \mu x$, то

   $$\lambda x = \mu x \Rightarrow (\lambda - \mu)x = \theta \Rightarrow x = \theta,$$

   что противоречит определению собственного вектора, так как собственный вектор всегда ненулевой. $\square$

2. Каждому собственному значению соответствуют свои собственные векторы, причем таких бесконечно много.

   Д о к а з а т е л ь с т в о. Действительно, если x – собственный вектор линейного оператора φ с собственным значением λ, т.е. $\varphi x = \lambda x$, то для $\forall \alpha \in \mathbb{R} \setminus \{0\}$ имеем $\alpha x \neq \theta$ и

   $$\varphi(\alpha x) = \alpha(\varphi(x)) = \alpha\lambda x = \lambda(\alpha x).$$

   Значит, и вектор αx является для линейного оператора собственным. $\square$

Теорема 18. *Число λ является собственным значением линейного оператора $\varphi \in L(\mathcal{L}_n, \mathcal{L}_n)$ тогда и только тогда, когда оно является корнем его характеристического уравнения.*

Д о к а з а т е л ь с т в о. *Необходимость.* Пусть λ – собственное значение оператора φ,
x – собственный вектор, отвечающий этому $\lambda (x \neq \theta)$. Перепишем соотношение (14) в следующем виде

$$(\varphi - \lambda\varepsilon)x = \theta,$$

где $\varepsilon \in L(\mathcal{L}_n, \mathcal{L}_n)$ – тождественный оператор.

Так как $x \neq \theta \Rightarrow \ker(\varphi - \lambda\varepsilon) \neq \theta$, т.е. $\dim(\ker(\varphi - \lambda\varepsilon)) \geq 1$, а так как

$$\dim(\mathrm{im}(\varphi - \lambda\varepsilon)) + \dim(\ker(\varphi - \lambda\varepsilon)) = n,$$

$$\operatorname{rg}(\varphi - \lambda\varepsilon) + \operatorname{def}(\varphi - \lambda\varepsilon) = n,$$

то $\operatorname{rg}(\varphi - \lambda\varepsilon) < n$, т.е. $\det(\varphi - \lambda\varepsilon) = 0$ и $\Rightarrow \lambda$ – корень характеристического уравнения.

Достаточность. Легко убедиться, что приведенные рассуждения можно провести в обратном порядке. $\square$

Следствие. Каждые линейный оператор имеет собственное значение. Действительно, характеристическое уравнение всегда имеет корень (в силу основной теоремы алгебры).

Алгебраической кратностью собственного значения оператора будем называть кратность соответствующего корня характеристического уравнения этого оператора.

Собственное подпространство линейного оператора.

Не следует путать два термина: собственное подпространство и собственное подпространство линейного оператора.

Множество всех собственных векторов, отвечающих данному собственному значению линейного оператора, не является линейным подпространством, так как это множество не содержит θ вектора, который по определению не может быть собственным. Это формальное и легко устранимое препятствие является единственным.

Пусть $\mathrm{V}(\varphi, \lambda)$ – множество всех собственных векторов линейного оператора $\varphi \in L(\mathcal{L}_n)$, соответствующих значению λ с добавленным к этому множеству нулевым вектором.

Теорема 19. *Множество $\mathrm{V}(\varphi, \lambda)$ линейное подпространство в $\mathcal{L}_n$.*

Доказательство. Пусть $x, y \in \mathrm{V}(\varphi, \lambda)$.

$\varphi(x + y) = \varphi x + \varphi y = \lambda x + \lambda y = \lambda(x + y)$;

$\varphi(\alpha x) = \alpha\varphi x = \alpha\lambda x = \lambda\alpha x$. $\square$

Множество $\mathrm{V}(\varphi, \lambda)$ называется *собственным подпространством линейного оператора.*

Собственное подпространство является инвариантным относительно оператора φ.

Геометрическая кратность собственного значения λ – это $\dim(\mathrm{V}(\varphi, \lambda))$.

Теорема 20. *Для того, чтобы матрица $[\varphi]_e$ линейного оператора φ в данном базисе e была диагональной, необходимо и достаточно, чтобы базисные векторы e_k были собственными векторами этого оператора.*

Доказательство. Пусть базисные векторы e_k являются собственными векторами оператора φ. Тогда

$$\varphi e_k = \lambda_k e_k, \tag{15}$$

и поэтому матрица $[\varphi]_e$ имеет вид (согласно равенствам (2))

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \tag{16}$$

т.е. является диагональной.

Пусть матрица $[\varphi]_e$ диагональна, т.е. имеет вид (16). Тогда соотношения (2) примут вид (15), а это означает, что e_k – собственные векторы оператора φ. $\square$

Теорема 21. *Пусть собственные значения $\lambda_1, \lambda_2, \dots, \lambda_n$ линейного оператора φ различны, тогда отвечающие им собственные векторы $e_1, e_2, \dots, e_n$ линейно независимы.*

Д о к а з а т е л ь с т в о. Применим индукцию. Так как e_1 – ненулевой вектор, то для одного вектора ($p = 1$) утверждение справедливо (один ненулевой вектор является линейно независимым).

Пусть утверждение теоремы доказано для m векторов $e_1, e_2, \dots, e_m$. Присоединим к этим векторам e_{m+1} и допустим, что имеет место равенство

$$\sum_{k=1}^{m+1} \alpha_k e_k = 0. \tag{17}$$

Тогда, используя свойства линейного оператора, получим

$$\sum_{k=1}^{m+1} \alpha_k \varphi e_k = 0. \tag{18}$$

Так как e_k – собственные векторы, то $\varphi e_k = \lambda_k e_k$, и поэтому равенство (18) можно переписать следующим образом:

$$\sum_{k=1}^{m+1} \alpha_k \lambda_k e_k = 0. \tag{19}$$

Согласно (17) $\sum_{k=1}^{m+1} \lambda_{m+1} \alpha_k e_k = 0$. Вычитая это равенство из равенства (19), найдем

$$\sum_{k=1}^{m} (\lambda_k - \lambda_{m+1}) \alpha_k e_k = 0. \tag{20}$$

По условию все λ_k различны, т.е. $\lambda_k - \lambda_{m+1} \neq 0$. Поэтому из (20) и предположения о линейной независимости векторов $e_1, e_2, \dots, e_m$ следует, что $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$. Отсюда и из (17), а также из условия, что e_{m+1} – собственный вектор ($e_{m+1} \neq \theta$), вытекает, что $\alpha_{m+1} = 0$. Таким образом из равенства (17) мы получаем, что $\alpha_1 = \alpha_2 = \dots = \alpha_{m+1} = 0$. Это означает, что векторы $e_1, e_2, \dots, e_{m+1}$ линейно независимы. $\square$

Алгоритм нахождения собственных значений и векторов линейного оператора.

Чтобы вычислить собственные значения линейного оператора $\varphi : \mathcal{L}_n \to \mathcal{L}_n$, действующего в вещественном линейном пространстве, нужно выполнить следующие операции:

1. Выбрать в линейном пространстве базис e и сопоставить линейному оператору φ матрицу $[\varphi]_e$ в выбранном базисе e.

2. Составить характеристическое уравнение $\det([\varphi]_e - \lambda E)$ и найти всего корни.

3. Выделить только вещественные корни λ_k, так как пространство вещественное. Если действительных корней нет, то нет и собственных векторов.

4. Для каждого собственного значения λ_k найти ФСР для однорордной системы уравнений $(A - \lambda_k E)x = \theta$. Столбцы ФСР представляют собой координаты векторов некоторого базиса в собственном подпространстве $\mathrm{V}(\varphi, \lambda_k)$ линейного оператора φ. Каждому собственному вектору соответствует собственное значение $\lambda_k \in \mathrm{V}(\varphi, \lambda_k)$ и, следовательно, найденный базис в этом подпространстве позволяет представить любой собственный вектор с с собственным значением λ_k.

П р и м е р 3. *Найти собственные векторы линейного преобразования* $\varphi : \mathcal{L}_2 \to \mathcal{L}_2$, *заданного матрицей* $A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$.

Если

а) $\mathcal{L}_2$ - *вещественное линейное пространство;*

б) $\mathcal{L}_2$ - *комплексное линейное пространство.*

Решение.

$$\det(A - \lambda E) = 0; \quad \begin{vmatrix} 1-\lambda & 2 \\ -1 & -1-\lambda \end{vmatrix} = 0;$$

$$-(1-\lambda)(1+\lambda) = 0;$$

$$\lambda^2 = -1;$$

$$\lambda = \pm i.$$

а) так как λ - комплексное, то собственных значений нет.

б) $\lambda = i:$

$$\begin{pmatrix} 1-i & 2 \\ -1 & -1-i \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

$$\begin{pmatrix} 1-i & 2 \\ -1 & -1-i \end{pmatrix} \sim \begin{pmatrix} 0 & 0 \\ -1 & -1-i \end{pmatrix};$$

$$x_1 = (-1-i)x_2, \quad x_2 = 1 \quad \Rightarrow \quad X^* = \alpha \begin{pmatrix} -1-i \\ 1 \end{pmatrix}, \qquad \alpha \in \mathbb{C} \setminus \{0\}.$$

$\lambda = -i :$

$$\begin{pmatrix} 1+i & 2 \\ -1 & -1+i \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$

$$\begin{pmatrix} 1+i & 2 \\ -1 & -1+i \end{pmatrix} \sim \begin{pmatrix} 0 & 0 \\ -1 & -1+i \end{pmatrix};$$

$$x_1 = (-1+i)x_2, \; x_2 = 1 \Rightarrow \; X^* = \alpha \begin{pmatrix} -1+i \\ 1 \end{pmatrix}, \; \alpha \in \mathbb{C} \setminus \{0\}.$$

Операторы простой структуры

Оператор $\varphi : \mathcal{L}_n \to \mathcal{L}_n$ называется *оператором простой структуры*, если в $\mathcal{L}_n$ существует базис из собственных векторов этого линейного оператора.

В базисе из собственных векторов матрица оператора простой структуры имеет вид

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

где $\lambda_1, \lambda_2, \dots, \lambda_n$ – собственные значения оператора.

Если в исходном базисе $[\varphi]_e = A, \;\; [\varphi]_{e'} = \Lambda,$ и $P_{e \to e'}$ – матрица перехода, то

$$\Lambda = P^{-1}_{e \to e'} A P_{e \to e'}, \tag{21}$$

$$A = P_{e \to e'} \Lambda P^{-1}_{e \to e'}. \tag{22}$$

На матричном языке соотношение (21) означает, что матрица A приводится матрицей $P_{e \to e'}$ к диагональному виду и оператор простой структуры называется также *диагонализируемым оператором.*

Соотношение (22) называется *каноническим разложением матрицы* A, а $P_{e \to e'}$ – трансформирующей матрицей.

Теорема 22. *Оператор $\varphi : \mathcal{L}_n \to \mathcal{L}_n$ является оператором простой структуры тогда и только тогда, когда алгебраическая и геометрическая кратности его собственных значений совпадают.*

<u>*Замечание*</u>. Эта теорема в вещественном пространсте верна только для тех операторов, чьи характеристические многочлены имеют только вещественные корни.

Приведение матрицы к диагональному виду и каноническое разложение матриц используется в теории и вычислительной практике. Например, если известно каноническое разложение (22), то, если $m \in \mathbb{Z} \setminus \{0\}$,

$$A^m = P_{e \to e'} \Lambda^m P^{-1}_{e \to e'}.$$

Алгоритм нахождения трансформирующей матрицы.

1. Находим все собственные значения матрицы A.
2. При каждом собственном значении λ_k строим ФСР однородной системы уравнения $(A - \lambda_k E)x = \theta$.
3. Из решений всех построенных ФСР, как из столбцов, составляем матрицу $P_{e \to e'}$, причем в матрицу $P_{e \to e'}$ столбцами записываются решения по каждому λ_k в порядке нумерации собственных значений.

Матрица $P_{e\to e'}$ должна быть квадратной. Это будет выполняться только тогда, когда каждый корень характеристического уравнения λ_k матрицы A является ее собственным значением и для каждого λ_k его алгебраическая и геометрическая кратности совпадают. Лишь в этом случае матрица A приводится к диагональному виду.

П р и м е р 4. *Привести, если возможно, следующую матрицу к диагональному виду и найти ее трансформирующую матрицу:*

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

Решение.

$$|A-\lambda E| = \begin{vmatrix} -1-\lambda & 3 & -1 \\ -3 & 5-\lambda & -1 \\ -3 & 3 & 1-\lambda \end{vmatrix} = \begin{vmatrix} -1-\lambda & 3 & -1 \\ -3 & 5-\lambda & -1 \\ 0 & \lambda-2 & 2-\lambda \end{vmatrix} =$$

$$= (2-\lambda)(\lambda^2-3\lambda+2) = (2-\lambda)(\lambda-1)(\lambda-2) = 0$$

$$\lambda_{1,2} = 2; \qquad \overbrace{s = 2;}^{\text{алгебраическая кратность}}$$

$$\lambda_3 = 1; \qquad s = 1.$$

Найдем геометрическую кратность:

При $\lambda_{1,2} = 2:$

$$\mathrm{rg}(A-\lambda_{1,2}E) = \mathrm{rg}\begin{pmatrix} -3 & 3 & -1 \\ -3 & 3 & -1 \\ -3 & 3 & -1 \end{pmatrix} = 1 \quad\Rightarrow\quad \overbrace{k = n - \mathrm{rg}(A-\lambda_{1,2}E) = 2}^{\text{геометрическая кратность}}.$$

При $\lambda_3 = 1:$

$$\mathrm{rg}(A-\lambda_3E) = \mathrm{rg}\begin{pmatrix} -2 & 3 & -1 \\ -3 & 4 & -1 \\ -3 & 3 & 0 \end{pmatrix} = 2 \quad\Rightarrow\quad k = n - \mathrm{rg}(A-\lambda_3E) = 1.$$

Найдем собственные векторы.

При $\lambda_{1,2} = 2:$

$$-3x_1 + 3x_2 - x_3 = 0 \quad\Rightarrow\quad x_3 = 3x_2 - 3x_1.$$

$$X^* = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}, \quad \textit{где} \quad \alpha_1, \alpha_2 \in \mathbb{R} \quad \text{и} \quad \alpha_1^2 + \alpha_2^2 \neq 0.$$

При $\lambda_3 = 1$:

$$\begin{pmatrix} -2 & 3 & -1 \\ -3 & 4 & -1 \\ -3 & 3 & 0 \end{pmatrix} \sim \begin{pmatrix} -2 & 3 & -1 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\begin{cases} x_1 - x_3 = 0, \\ -x_1 + x_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1 = x_3, \\ x_1 = x_2 \end{cases} \Rightarrow X^* = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \text{ где } \quad \lambda \in \mathbb{R} \setminus \{0\}.$$

В итоге:

$$\Lambda = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad P_{e \to e'} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -3 & 3 & 1 \end{pmatrix}.$$

Жорданова нормальная форма

Итак, самый простой формой матрицы обладают только операторы простой структуры, т.е. операторы, имеющие полный набор линейно независимых собственных векторов. Как мы уже отмечали в вещественном пространстве существуют операторы, которые не имеют ни одного собственного вектора. И в комплексном пространстве не каждый линейный оператор обладает необходимым для базиса числом линейно независимых векторов.

Приведение матрицы линейного оператора к простому виду связано со структурой его собственных подпространств.

Теорема 23. *Пусть* $\mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_s$ *– инвариантные пространства линейного оператора* $\varphi : \mathcal{L}_n \to \mathcal{L}_n$*, причем* $\mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \cdots \oplus \mathcal{L}_s = \mathcal{L}$*, тогда в некотором базисе* f *матрица оператора* φ *имеет блочно-диагональный вид:*

$$[\varphi]_f = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \cdots & A_s \end{pmatrix},$$

где квадратный блок A_i *имеет порядок* $\dim \mathcal{L}_i$, $i = 1, 2, \ldots, s$*, а остальные блоки являются нулевыми.*

Д о к а з а т е л ь с т в о. Выберем в линейных подпространствах $\mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_s$ базисы

$e^{(1)} = (e_1^{(1)}, e_2^{(1)}, \ldots, e_{n_1}^{(1)})$,

$e^{(2)} = (e_1^{(2)}, e_2^{(2)}, \ldots, e_{n_2}^{(2)})$,

$\dots$

$e^{(s)} = (e_1^{(s)}, e_2^{(s)}, \ldots, e_{n_s}^{(s)})$.

В совокупности эти базисы дают базис f всего пространства $\mathcal{L}$. Так как $\mathcal{L}_1$ – инвариантное подпространство линейного оператора φ, элемент $\varphi e_i^{(1)}$, $i = 1, 2, \ldots, n_1$, попадет в $\mathcal{L}_1$ и поэтому является линейной комбинацией системы элементов $e^{(1)}$. Другими словами, координаты элементов $\varphi e_i^{(1)}$ в базисе f, соответствующие $e_i^{(2)}, e_i^{(3)}, \ldots, e_i^{(s)}$, равны нулю. Аналогично координаты элементов $\varphi e_i^{(2)}$, $i = 1, 2, \ldots, n_2$ в базисе f, соответствующие $e_i^{(1)}, e_i^{(3)}, \ldots, e_i^{(s)}$, также равны нулю.

Остановимся на случае, когда характеристическое уравнение линейного оператора имеет лишь простые корни, среди которых, вообще говоря, есть и комплексные. Так как характеристическое уравнение линейного оператора имеет действительные коэффициенты, каждому комплексному корню $\alpha + i\beta$ этого уравнения соответствует комплексно сопряженный корень $\alpha - i\beta$ той же кратности.

Теорема 24. *Каждой паре комплексно-сопряженных корней характеристического уравнения линейного оператора соответствует двумерное инвариантное подпространство этого оператора.*

Доказательство. Зафиксируем в линейном пространстве $\mathcal{L}$ некоторый базис e и рассмотрим матрицу A линейного оператора φ в этом базисе.

Пусть $\lambda = \alpha + i\beta,\ \beta \neq 0$ – комплексный корень характеристического уравнения линейного оператора φ.

Тогда $\det(A - \lambda E) = 0$ и система линейных уравнений $(A - \lambda E)x = \theta$ с комплексными коэффициентами имеет ненулевое решение x, которое можно записать в виде $x = u + iv$, разделив действительные и мнимые части у элементов столбца x.

Столбец v не является нулевым, так как в противном случае $x = u$, $Au = \lambda x$. Мы видим, что действительные элементы столбца Au получаются из действительных элементов столбца u умножением на комплексное число λ, а это возможно лишь в случае, когда $u = \theta$. Но это заключение противоречит выбору столбца x.

Столбцы u и v линейно независимы. Действительно, если они линейно зависимы, то $\mu u + \nu v = 0$, где одно из чисел μ и $\nu \neq 0$. Мы можем утверждать, что $\mu \neq 0$, так как в противном случае $\nu v = \theta$. Но $v \neq \theta$, значит $\nu = 0$.

Пусть $\mu \neq 0$ и поэтому $u = kv$, где $k = -\frac{\nu}{\mu} \in \mathbb{R} \ \Rightarrow\ x = u + iv = (k + i)v$.

Так как $Ax = \lambda x$, то

$$A(k + i)v = \lambda(k + i)v,$$

$$Av = \lambda v.$$

Как мы уже знаем, для комплексных λ такое равенство невозможно.

В равенстве $Ax = \lambda x$ сделаем замены $\lambda = \alpha + i\beta,\ x = u + iv$:

$$A(u + iv) = (\alpha + i\beta)(u + iv).$$

Разделив действительные и мнимые части, получим два матричных уравнения

$$Au = \alpha u - \beta v,\ Av = \beta u + \alpha v.$$

Рассмотрим векторы x и y, которые в базисе e имеют координатные столбцы $x_e = u,\ y_e = v$, тогда

$$\varphi x = \alpha x - \beta y, \quad \varphi y = \beta x + \alpha y.$$

Векторы x и y линейно независимы, так как независимы их столбцы u и v. Полученные соотношения означают, что двумерное линейное подпространство $\mathcal{L}' = L\{x, y\}$ является инвариантным подпространством линейного оператора φ. $\square$

Для $\forall\alpha, \beta \in \mathbb{R}$ обозначим

$$C(\alpha, \beta) = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Теорема 25. *Если характеристическое уравнение линейного оператора $\varphi : \mathcal{L}_n \longrightarrow \mathcal{L}_n$ имеет p различных пар комплексно сопряженных корней $\alpha_j \pm i\beta_j$, где $j = 1, 2, \cdots, p$, и q различных действительных корней μ_j где $j = 1, 2, \cdots, q$, причем $2p + q = n$, где $\dim(\mathcal{L}_n) = n$, тогда матрица линейного оператора в некотором базисе будет иметь вид:*

$$A = \begin{pmatrix} C(\alpha_1, \beta_1) & 0 & \cdots & \cdots & 0 & 0 \\ 0 & C(\alpha_2, \beta_2) & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & C(\alpha_p, \beta_p) & \cdots & 0 & 0 \\ 0 & 0 & 0 & \mu_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu_2 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots & \mu_q \end{pmatrix}.$$

Д о к а з а т е л ь с т в о. Каждой паре комплексно сопряженных корней $\alpha_j \pm i\beta_j$ характеристического уравнения соответствует двумерное инвариантное подпространство P_j оператора φ с базисом u_j, v_j (см. доказательство т. 24). Каждому собственному значению μ_j соответствует одномерное собственное подпространство Q_j линейного оператора φ. Можно показать, что все эти подпространства образуют прямую сумму, так как пересечение любой пары таких подпространств содержит лишь θ. Учитывая, что сумма размерностей этих подпространств $2p + q = n = \dim(\mathcal{L}_n)$, заключаем, что $P_1 \oplus P_2 + \ldots + P_p \oplus Q_1 \oplus Q_2 \ldots \oplus Q_q = \mathcal{L}$. Согласно теореме 23 в некотором базисе матрица A оператора φ имеет блочно-диагональный вид, причем каждый диагональный блок представляет собой ограничения оператора φ на соответствующее инвариантное подпространство. В случае двумерного подпространства P_j в базисе u_j, v_j эта матрица равна $C(\alpha_j, \beta_j)$, а в случае одномерного инвариантного подпространства Q_j такой блок есть простое число, представляющее собой собственное значение μ_j. $\square$

Если характеристическое уравнение линейного оператора имеет кратные корни, действительные или комплексные, то инвариантные подпространства такого оператора имеют более сложную структуру.

Рассмотрим два типа специальных матриц. Для произвольного числа $\mu \in \mathbb{R}$ введем обозначение матрицы порядка s:

$$J_s(\mu) = \begin{pmatrix} \mu & 1 & 0 & \cdots & 0 \\ 0 & \mu & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \mu & 1 \\ 0 & 0 & 0 & \cdots & \mu \end{pmatrix}.$$

У этой матрицы все диагональные элементы равны μ, над главной диагональю расположены единицы, а все остальные равны нулю. В случае $s = 1$ рассматриваемая матрица сводится к единственному числу μ.

Для любого комплексного числа $\alpha + i\beta (\beta \neq 0)$ введем обозначение блочной матрицы порядка $2r$:

$$C_r(\alpha,\beta) = \begin{pmatrix} C(\alpha,\beta) & E & 0 & \cdots & 0 \\ 0 & C(\alpha,\beta) & E & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & C(\alpha,\beta) & E \\ 0 & 0 & 0 & \cdots & C(\alpha,\beta) \end{pmatrix},$$

где $C_1(\alpha,\beta) = C(\alpha,\beta) = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$. Все остальные блоки также являются матрицами второго порядка. E обозначим единичную матрицу, а O – нулевую.

Блочно-диагональную матрицу вида

$$\begin{pmatrix} C_{r_1}(\alpha_1,\beta_1) & 0 & \cdots & \cdots & 0 & 0 \\ 0 & C_{r_2}(\alpha_2,\beta_2) & \cdots & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & C_{r_l}(\alpha_l,\beta_l) & \cdots & 0 & 0 \\ 0 & 0 & 0 & J_{s_1}(\mu_1) & 0 & 0 \\ 0 & 0 & 0 & 0 & J_{s_2}(\mu_2) & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \cdots & J_{s_k}(\mu_k) \end{pmatrix},$$

где $\alpha_i, \beta_i \in \mathbb{R},\ i = 1, 2, \ldots, m,\ \mu_j \in \mathbb{R},\ j = 1, 2, \ldots, k$ называют *жордановой*. Ее диагональные блоки – *жордановыми клетками*.

Квадратичные формы.

Рассмотрим симметричную билинейную форму $f(x, y)$ в вещественном линейном пространстве $\mathcal{L}$.

Квадратичной формой (функцией , функционалом) будем называть вещественнозначную функцию $f(x, x)$, полученную из симметричной билинейной формы путем замены y на x, где $x \in \mathcal{L}$.

Соответствующую билинейную форму называют *полярной к квадратичной форме* $f(x, x)$.

Связь между квадратичной и полярной формой:

$f(x, y) = \frac{1}{2}(f(x + y, x + y) - f(y, y) - f(x, x))$.

В базисе e квадратичная форма $f(x, x)$ с матрицей $A_e = (a_{ij})$ может быть записана в следующем общем виде:

$f(x, x) = \sum\limits_{i,j=1}^{n} a_{ij}x_ix_j,$

$a_{ij} = a_{ji},$

или в компактной форме $f(x, x) = x_e^T A_e x_e$.

Рангом квадратичной формы будем называть ранг ее матрицы в произвольном базисе.

Квадратичная форма *вырожденная*, если ранг формы меньше размерности пространства, в котором она определена.

П р и м е р 2. *Составить матрицу билинейной формы и записать соответствующую ей квадратичную форму в двумерном пространстве, если билинейная форма:* $f(x, y) = 2x_1y_1 - x_1y_2 - x_2y_1 - 5x_1y_2$.

Решение. Матрица билинейной формы будет иметь вид

$$\begin{pmatrix} 2 & -1 \\ -1 & -5 \end{pmatrix}.$$

Соответствующая квадратичная форма:

$f(x, x) = 2x_1^2 - x_1x_2 - x_2x_1 - 5x_2^2 = 2x_1^2 - 2x_1x_2 - 5x_2^2$.

Виды квадратичных форм.

1. Квадратичная форма $f(x, x)$ называется *положительно (отрицательно) определенной*, если $\forall x \neq \theta$ $f(x, x) > 0$ ($f(x, x) < 0$).

2. Квадратичная форма называется *знакопеременной*, если $\exists x, y \in \mathcal{L}$, такие, что одновременно выполняются $f(x, x) > 0$ и $f(y, y) < 0$.

3. Квадратичная форма называется *положительно полуопределенной (отрицательно полуопределенной)*, если $\forall x \quad f(x, x) \geq 0$ $(f(x, x) \leq 0)$ и $\exists x \neq \theta$, при котором $f(x, x) = 0$.

Теорема 4. *Пусть $f(x, y)$– симметричная билинейная форма, полярная к положительно определенной квадратичной форме $f(x, x)$, тогда форма $f(x, y)$ определяет скалярное произведение в вещественном евклидовом пространстве.*

Д о к а з а т е л ь с т в о. Если число, называемое скалярным произведением векторов x и y, обозначить символом $f(x, y)$, то эти аксиомы запишутся следующим образом:

1) $f(x, y) = f(y, x)$;
2) $f(x + y, z) = f(x, z) + f(y, z)$;
3) $f(\lambda x, y) = \lambda f(x, y)$;
4) $f(x, x) \geq 0,\ f(x, x) > 0, x \neq \theta$.

Так как билинейная форма $f(x, y)$ полярная квадратичной форме $f(x, x)$ симметрична, то аксиома 1) выполняется. аксиомы 2) и 3) в сочетании с требованием симметрии выполнены в силу определения билинейной формы. Аксиома 4) выполняется, так как квадратичная форма $f(x, x)$ положительно определена. Значит билинейная форма определяет скалярное произведение в вещественном евклидовом пространстве. $\square$

Приведение квадратичной формы к каноническому виду.

Рассмотрим различные меоды приведения квадратичной формы к сумме квадратов, т. е. рассмотрим методы выбора такого базиса $f = (f_1, f_2, \ldots, f_n)$ в линейном пространстве $\mathcal{L}$, по отношению к которому квадратичная форма представляется в следующем *каноническом виде*:

$$f(x,x) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2, \tag{4}$$

$x_1, x_2, \ldots, x_n$ – координаты x в базисе f.

Коэффициенты $\lambda_1, \lambda_2, \ldots, \lambda_n$ в выражении (4) называются *каноническими коэффициентами.*

Так как каждому преобразованию базиса отвечает невырожденное линейное преобразование координат, а невырожденному преобразованию координат – преобразование базиса, то вопрос о приведении формы к каноническому виду можно решать путем выбора соответствующего невырожденного преобразования координат.

Метод Лагранжа.

Теорема 5. *Любая квадратичная форма $f(x,x)$, заданная в n-мерном линейном пространстве $\mathcal{L}$, с помощью невырожденного линейного преобразования координат может быть приведена к каноническому виду* (4).

Д о к а з а т е л ь с т в о. Проведем доказательство теоремы *методом Лагранжа.* Основная идея этого метода заключается в последовательном дополнении квадратного трехчлена по каждому аргументу до полного квадрата.

Будем считать, что $f(x,x) \neq 0$ (если форма $f(x,x) \equiv 0$, то ее матрица в любом базисе состоит из нулевых элементов, и поэтому такая форма по определению имеет канонический вид в любом базисе) и в данном базисе $e = (e_1, e_2, \ldots, e_n)$ имеет вид

$$f(x,x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j. \tag{5}$$

Убедимся, во-первых, что с помощью невырожденного преобразования координат форму $f(x,x)$ можно преобразовать так, что коэффициент при квадрате первой координаты вектора x будет отличен от нуля.

Если в данном базисе этот коэффициент отличен от нуля, то нужное невырожденное преобразование является тождественным.

В случае, если $a_{11} = 0$, но отличен от нуля коэффициент при квадрате какой-либо другой координаты, то с помощью перенумерации базисных векторов можно добиться требуемого результата. Ясно, что перенумерация является невырожденным преобразованием.

Если же все коэффициенты при квадратах координат равны нулю, то нужное преобразование можно получить следующим способом. Пусть, например, $a_{12} \neq 0$. (Напомним, что $f(x,x) \neq 0$ и поэтому хотя бы один коэффициент a_{ij} отличен от нуля). Рассмотрим следующее невырожденное преобразование координат (определитель матрицы этого преобразования равен 2, и поэтому это преобразование невырожденное):

$$\begin{aligned} &x'_1 = x_1 - x_2, \\ &x'_2 = x_1 + x_2, \\ &x'_i = x_i, \; i = 3, \; 4, \; \ldots, \; n. \end{aligned}$$

После этого преобразования коэффициент при x_i^2 будет равен $2a_{12}$ и поэтому отличен от нуля.

Итак, будем считать, что в соотношении (5) $a_{11} \neq 0$. Выделим в выражении (5) ту группу слагаемых, которые содержат x_1. Получим

$$f(x,x) = a_{11}x_1^2 \; + \; 2a_{12}x_1x_2 \; + \; \ldots \; + \; 2a_{1n}x_1x_n \; + \sum_{i,j \, = \, 2}^{n} a_{ij}x_ix_j. \qquad (6)$$

Преобразуем выделенную группу слагаемых следующим образом:

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n = a_{11}\left(x_1 + x_2\frac{a_{12}}{a_{11}} + \ldots + x_n\frac{a_{1n}}{a_{11}}\right)^2 \; -$$

$$-\frac{a_{12}^2}{a_{11}}x_2^2 - \ldots - \frac{a_{1n}^2}{a_{11}}x_n^2 - 2\frac{a_{12}a_{13}}{a_{11}}x_2x_3 - \ldots - 2\frac{a_{1n-1}a_{1n}}{a_{11}}x_{n-1}x_n.$$

Очевидно, выражение (6) можно теперь переписать так:

$$f(x,x) \; = \; a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)^2 \; + \sum_{i,j=2}^{n} a_{ij}^* x_ix_j, \qquad (7)$$

где a_{ij}^*–коэффициенты при x_ix_j, полученные после преобразования.

Рассмотрим следующее невырожденное преобразование координат:

$$x'_1 \; = \; x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n,$$

$$x'_2 \; = \; x_2,$$

$$\ldots$$

$$x'_n \; = \; x_n.$$

С помощью этого преобразования и представления (7) для $f(x,x)$ получим

$$f(x,x) \;=\; a_{11}(x'_1)^2 + \sum_{i,j=2}^{n} a^*_{ij}x'_ix'_j. \tag{8}$$

Итак, если форма $f(x,x) \;\neq\; 0$, то с помощью невырожденного преобразования координат эту форму можно привести к виду (8).

Обратимся теперь к квадратичнй форме $\sum\limits_{i,j=2}^{n} a^*_{ij}x'_ix'_j$. Если эта форма тождественно равна нулю, то вопрос о приведении $f(x,x)$ к каноническому виду решен. Если же форма $\sum\limits_{i,j=2}^{n} a^*_{ij}x_ix_j \;\neq\; 0$, то мы можем повторить рассуждения, рассматривая преобразования координат $x'_2,\ \ldots, x'_n$, аналогичные описанным выше, и не меняя при этом координату x'_1. Очевидно, такого типа преобразования координат $x'_1,\ x'_2,\ \ldots,\ x'_n$ будут невырожденными.

Ясно что за конечное число шагов мы приведем квадратичную форму $f(x,x)$ к каноническому виду (4).

Отметим, что нужное преобразование исходных координат $x_1, x_2, \ldots, x_n$ можно получить путем перемножения найденных в процессе рассуждений невырожденных преобразований. $\square$

Замечание 1. Базис, в котором квадратичная форма имеет канонический вид, называется *каноническим.* Отметим, что канонический базис определен неоднозначно.

Замечание 2. Если форма $f(x,x)$ приведена к каноническому виду (4), то, вообще говоря, не все канонические коэффициенты λ_i отличны от нуля. Оставляя в (4) лишь отличные от нуля λ_i и перенумеровывая их заново, получим следующее выражение для $f(x,x)$:

$$f(x,x) = \lambda_1 x^2 + \lambda_2 x_2^2 + \ldots + \lambda_r x_r^2. \tag{9}$$

Ясно, что $r \leq n$. Так как ранг квадратичной формы по определению равен рангу ее матрицы в любом базисе, то из (9) и условия $\lambda_i \neq 0$ при $i = 1, 2, \ldots, r$ вытекает, что ранг формы равен r. Таким образом, число отличных от нуля канонических коэффициентов равно рангу квадратичной формы.

Метод Якоби.

При некоторых дополнительных предположениях о квадратичной форме $f(x,x)$ можно указать явные формулы перехода от данного $e = (e_1,\ e_2, ..., e_n)$ базиса к каноническому $e' = (e'_1,\ e'_2, ..., e'_n)$ и указать явные формулы канонических коэффициентов λ_i.

Введем понятие треугольного преобразования базисных векторов.

Преобразование базисных векторов $e = (e_1,\ e_2, ..., e_n)$ называется *треугольным*, если оно имеет следующий вид:

$$\begin{array}{l} e'_1 = e_1, \\ e'_2 = \alpha_{21}e_1 + e_2, \\ e'_3 = \alpha_{31}e_1 + \alpha_{32}e_2 + e_3, \\ \quad \dots \quad \dots \quad \dots \\ e'_n = \alpha_{n1}e_1 + \alpha_{n2}e_2 + ... + e_n. \end{array} \tag{10}$$

Пусть $f(x, x)-$ квадратичная форма. И пусть $A-$ матрица квадратичной формы в базисе e.

Пусть $\Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \ \dots\ ,\ \Delta_n = |A|$ – угловые миноры матрицы A.

Теорема 6. *Пусть угловые миноры матрицы квадратичной формы $f(x, x)$ отличны от нуля. Тогда существует единственное треугольное преобразование базисных векторов $e_1,\ e_2, ..., e_n$, которое приводит эту квадратичную форму к каноническому виду.*

Доказательство. Коэффициенты b_{ij} квадратичной формы $f(x, x)$ в базисе e' вычисляются по формулам

$$b_{ij} = f(e'_i, e'_j). \tag{11}$$

Используя равенства (10) и линейное свойство квадратичной формы $f(x, x)$ по каждому аргументу, легко заметить, что соотношения (11) будут выполнены, если будут выполнены соотношения:

$$f(e_1, e'_j) = 0, \quad f(e_2, e'_j) = 0 \quad ... \quad f(e_{j-1}, e'_j) = 0, \quad j = 2, 3, ..., n. \tag{12}$$

Запишем формулы (12) в развернутом виде. Для этого подставим в левые части этих формул выражение

$$e'_j = \alpha_{j1}e_1 + \alpha_{j2}e_2 + ... + \alpha_{jj-1}e_{j-1} + e_j \tag{13}$$

из соотношений (10). Используя далее свойство линейности $f(x, x)$ по каждому аргументу и обозначение $f(e_i, e_j) = a_{ij}$, получим в результате следующую систему уравнений для неизвестных коэффициентов α_{jk}:

$$\left\{ \begin{array}{rl} \alpha_{j1}a_{11} + \alpha_{j2}a_{12} + ... + \alpha_{jj-1}a_{1j-1} + a_{1j} & = 0, \\ \alpha_{j1}a_{21} + \alpha_{j2}a_{22} + ... + \alpha_{jj-1}a_{2j-1} + a_{2j} & = 0, \\ \dots \quad \dots \quad \dots & \\ \alpha_{j1}a_{j-11} + \alpha_{j2}a_{j-12} + ... + \alpha_{jj-1}a_{j-1j-1} + a_{j-1j} & = 0, \end{array} \right. \tag{14}$$

Определитель этой системы равен Δ_{j-1}. По условию $\Delta_{j-1} \neq 0$. Следовательно, система (14) имеет единственное решение. Таким образом, можно построить единственное треугольное преобразование базисных векторов, с помощью которого квадратичная форма $f(x, x)$ приводится к каноническому виду. $\Box$

Приведем формулы, по которым можно вычислить коэффициенты α_{ji} искомого треугольного преобразования и формулы для канонических коэффициентов λ_j. Используя формулу Крамера находим выражение для коэффициентов:

$$\alpha_{ji} = \frac{(-1)^{i+j}\Delta_{j-1,i}}{\Delta_{j-1}}, \tag{15}$$

где $\Delta_{j-1,i}$ минор матрицы A, расположенный на пересечении строк этой матрицы с номерами $1, 2, ...j-1$ и столбцов с номерами $1, 2, ..., i-1, i+1, ..., j$. Так как j-й столбец должен стоять на i-ом месте, а мы приписываем его справа, то необходимо домножить на знак перестановки j-го столбца на i-е место.

Вычислим канонические коэффициенты λ_j.

$\lambda_{ij} = b_{jj} = f(e'_j, e'_j) = f(e_j, e'_j) = f(e_j, \alpha_{j1}e_1 + \alpha_{j2}e_2 + ... + \alpha_{jj-1}e_{j-1} + e_j) =$
$= \alpha_{ij}a_{j1} + \alpha_{j2}a_{j2} + ... + \alpha_{jj-1}a_{jj-1} + a_{jj}$.

Подставляя выражение (15) для α_{ji}, $\quad i = 1, 2, ..., j-1$ в правую часть последнего соотношения, найдем

$\lambda_j =$

$$= \frac{(-1)^{j+1}\Delta_{j-1,1}a_{j1} + (-1)^{j+2}\Delta_{j-1,2}a_{j2} + ... + (-1)^{i+j-1}\Delta_{j-1,j-1}a_{jj-1} + a_{jj}\Delta_{j-1}}{\Delta_{j-1}}.$$

Числитель представляет собой Δ_j. Поэтому

$\lambda_j = \frac{\Delta_j}{\Delta_{j-1}}, \quad j = 2, ..., n,$
$\lambda_1 = f(e'_1, e'_1) = f(e_1, e_1) = a_{11}$.

П р и м е р 3. *С помощью метода Якоби вычислить коэффициенты треугольного преобразования и канонические коэффициенты, если квадратичная форма имеет вид:*

$f(x, x) = 2x_1^2 + 3x_2^2 + x_3^2 - 4x_1x_2 + 2x_1x_3 - 2x_2x_3$.

Решение.

Так как форма квадратичная, то матрица ее будет симметричной:

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

$\Delta_1 = 2; \quad \Delta_2 = 2; \quad \Delta_3 = 1.$
$\lambda_1 = 2; \quad \lambda_2 = \frac{\Delta_2}{\Delta_1} = 1; \quad \lambda_3 = \frac{\Delta_3}{\Delta_2} = \frac{1}{2}.$
$\alpha_{21} = -\frac{2}{2} = -1; \quad \alpha_{31} = -\frac{1}{2}; \quad \alpha_{32} = 0.$

Закон инерции квадратичных форм.

Мы уже отмечали, что ранг квадратичной формы равен числу отличных от нуля канонических коэффициентов. Таким образом, число отличных от нуля канонических коэффициентов не зависит от выбора невырожденного преобразованияа, с помощью которого форма приводится к каноническому виду. На самом деле при любом способе приведения формы к каноническому виду не меняется число положительных и отрицательных канонических коэффициентов. Это свойство называется *законом инерции квадратичных форм.*

Теорема 7. (Закон инерции квадратичных форм) . Число положительных и отрицательных коэффициентов в каноническом виде квадратичной формы не зависит от выбора канонического базиса.

Д о к а з а т е л ь с т в о. Пусть e и f – канонические базисы квадратичной формы $f(x,x)$ ранга r и для $x=\sum\limits_{i=1}^{n} x_i e_i = \sum\limits_{i=1}^{n} y_i f_i$

$$f(x,x) = a_1x_1^2 + a_2x_2^2 + ... + a_px_p^2 - a_{p+1}x_{p+1}^2 - ... - a_rx_r^2,$$

$$f(x,x) = b_1y_1^2 + b_2y_2^2 + ... + b_{p'}y_{p'}^2 - b_{p'+1}y_{p'+1}^2 - ... - b_ry_r^2,$$

где $a_i > 0,\ b_i > 0, i = 1, 2, ..., r$. Необходимо доказать, что $p = p'$.

1) Докажем, что $p \le p'$. Предположим,что это не выполняется, т.е. $p > p'$. Рассмотрим два подпространства $\mathcal{L}_1$ и $\mathcal{L}_2$:

$\mathcal{L}_1 = \mathcal{L}(e_1, e_2, ..., e_p)$, $\mathcal{L}_2 = \mathcal{L}(f_{p'+1}, f_{p'+2}, ..., f_n)$,

$\dim(\mathcal{L}1 \bigcap \mathcal{L}_2) = \dim \mathcal{L}_1 + \dim \mathcal{L}_2 - \dim(\mathcal{L}_1 + \mathcal{L}_2) = p + (n-p') - \dim(\mathcal{L}_1 + \mathcal{L}_2)$.

Так как $\dim(\mathcal{L}_1 + \mathcal{L}_2) \le n,\ p > p'$, то $\dim(\mathcal{L}_1 \bigcap \mathcal{L}_2) > 0$.

Следовательно, существует $x_0 \ne \theta$ и $x_0 \in \dim(\mathcal{L}_1 \bigcap \mathcal{L}_2)$.

Пусть $x_0 = \alpha_1 e_1 + \alpha_2 e_2 + ...\alpha_p e_p = \beta_{p'+1} f_{p'+1} + ... + \beta_n f_n$.

Тогда

$$f(x_0, x_0) = a_1\alpha_1^2 + a_2\alpha_2^2 + ... + a_p\alpha_p^2 = -b_{p'+1}\beta_{p'+1}^2 - ... - b_r\beta_r^2. \qquad (16)$$

Так как $x_0 \ne \theta$, то $a_1\alpha_1^2 + a_2\alpha_2^2 + ... + a_p\alpha_p^2 > 0,\ -b_{p'+1}\beta_{p'+1}^2 - ... - b_r\beta_r^2 < 0$. Это противоречит (16), и значит, $p \le p'$.

2) $p \ge p'$ доказывается аналогично. $\square$

Введем обозначения:

$i_+ = p$ – положительный индекс инерции — число положительных коэффициентов в каноническом разложении квадратичной формы.

$i_- = q$ – отрицательный индекс инерции — число отрицательных коэффициентов в каноническом виде квадратичной формы.

r – ранг квадратичной формы.

$s = p - q$ – сигнатура квадратичной формы.

Вид квадратичной формы

$$f(x,x) = x_1^2 + x_2^2 + ... + x_p^2 - x_{p+1}^2 - ... - x_r^2 \tag{17}$$

называется *нормальным.*

Теорема 8. (Критерий знакоопределенности квадратичной формы). *Для того чтобы квадратичная форма, заданная в n-мерном линейном пространстве $\mathcal{L}_n$, была знакоопределенной, необходимо и достаточно, чтобы в случае положительной определенности $p = n$, а в случае отрицательной определенности $q = n$.*

Д о к а з а т е л ь с т в о. **Необходимость.**

Пусть форма $f(x,x)$ положительно определена. Тогда выражение (17) примет вид $f(x,x) = x_1^2 + x_2^2 + ... + x_p^2$.

Если при этом $p < n$, то из последнего выражения следует, что для $x \neq \theta$ с координатами

$$x_1 = 0,\ x_2 = 0,\ \dots,\ x_p = 0,\ x_{p+1} \neq 0, \dots, x_n \neq 0$$

форма $f(x,x)$ обращается в нуль, а это противоречит определению положительно определенной квадратичной формы, поэтому $p = n$.

Достаточность. Пусть $p = n$. Тогда соотношение (17) имеет вид $f(x,x) = x_1^2 + x_2^2 + ... + x_n^2$.

Ясно, что $f(x,x) \geq 0$, причем, если $f(x,x) = 0$, то $x_1 = x_2 = \ldots = x_n = 0$, т.е. $x = \theta$. Следовательно, $f(x,x)$ – положительно определенная форма.□

Замечание. Для выяснения вопроса о знакоопределенности квадратичной формы с помощью указанного признака мы должны привести эту форму к каноническому виду.

Теорема 9. (Критерий знакопеременности квадратичной формы). *Для того чтобы квадратичная форма была знакопеременной, необходимо и достаточно, чтобы $p \neq 0$ и $q \neq 0$.*

Д о к а з а т е л ь с т в о. **Необходимость.** Так как знакопеременная форма принимает как положительные, так и отрицательные значения, то ее представление (17) в нормальном виде должно содержать как положительные, так и отрицательные слагаемые (в противном случае эта форма принимала бы либо неотрицательные, либо неположительные значения). Следовательно, как положительный, так и отрицательный индексы инерции отличны от нуля.

Достаточность. Пусть $p \neq 0$ и $q \neq 0$.

Тогда для вектора $x' = (0,\ 0 \dots 0,\ x_{p+1}, \dots, x_n)$ имеем $f(x', x') < 0$, а для вектора $x'' = (x_1, x_2, \dots, x_p, 0,\ 0, \dots, 0)$ имеем $f(x'', x'') > 0$. Следовательно, форма $f(x,x)$ является знакопеременной. □

Теорема 10. (Критерий полуопределённости квадратичной формы). *Для того, чтобы квадратичная форма $f(x,x)$ была полуопределённой необходимо и достаточно, чтобы выполнялись соотношения:*

для положительной полуопределённости: $p < n$, $q = 0$;

для отрицательной полуопределённости: $q < n$, $p = 0$.

Д о к а з а т е л ь с т в о. Рассмотрим случай положительно полуопределённой квадратичной формы. Случай отрицательной полуопределенности рассматривается аналогично.

Необходимость. Пусть форма $f(x,x)$ положительно полуопределенная. Тогда, очевидно, $p < n$ и $q = 0$ (если бы $p = n$, то форма была бы положительно определённой).

Достаточность. Если $p < n$, $q = 0$, то $f(x,x) \geq 0$ и для $x = (0,\ 0, \ldots, x_{p+1}, \ldots, x_n)$ имеем $f(x,x)$=0, т.е. $f(x,x)$ – положительно полуопределенная форма. $\square$

П р и м е р 4.

Дана квадратичная форма $f(x,x) = {x_1}^2 + 4x_1x_3 + {x_2}^2 + 2x_2x_3 + 4{x_3}^2$, $(n = 2)$. *Определить вид формы.*

<u>*Решение*</u>.

${x_1}^2 + 4x_1x_3 + {x_2}^2 + 2x_2x_3 + 4{x_3}^2$ =$({x_1}^2 + 4x_1x_3 + 4{x_3}^2) + ({x_2}^2 + 2x_2x_3 + {x_3}^2) -$
$- x_3^2 = (x_1 + 2x_3)^2 + (x_2 + x_3)^2 - {x_3}^2$.

Произведём замену:

$y_1 = x_1 + 2x_3$;
$y_2 = x_2 + x_3$;
$y_3 = x_3$.

Получим: ${y_1}^2 + {y_2}^2 - {y_3}^2$.

Следовательно, форма является знакопеременной $(p = 2,\ q = 1)$.

Критерий Сильвестра.

Пусть форма $f(x,x)$ в базисе $e=(e_1,e_2,\dots,e_n)$ определяется матрицей $A_e=(a_{ij})$:

$$f(x,x)=\sum_{i,j=1}^{n} a_{ij}x_ix_j$$

и пусть

$$\Delta_1=a_{11},\ \Delta_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \cdots,\ \Delta_n=|A|.$$

Теорема 11. (Критерий Сильвестра).
Для того, чтобы квадратичная форма была положительно определённой необходимо и достаточно, чтобы были выполнены неравенства

$$\Delta_1>0,\ \Delta_2>0,\dots,\Delta_n>0.$$

Для того, чтобы квадратичноая форма была отрицательно определённой, необходимо и достаточно, чтобы знаки угловых миноров чередовались, причём $\Delta_1<0$.

Д о к а з а т е л ь с т в о. **Необходимость.** Докажем сначала, что из условия знакоопределенности квадратичной формы $f(x,x)$ следует $\Delta_i\neq 0$.

Убедимся, что предположение $\Delta_k=0$ ведет к противоречию – при этом предположении $\exists x\neq\theta$, при котором квадратичная форма обращается в ноль, что противоречит знакоопределенности формы.

Итак, пусть $\Delta_k=0$. Рассмотрим следующую квадратную однородную систему линейных уравнений:

$$\begin{cases} a_{11}x_1+a_{12}x_2+\ldots+a_{1k}x_k= & 0,\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2k}x_k= & 0,\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & \\ a_{k1}x_1+a_{k2}x_2+\ldots+a_{kk}x_k= & 0, \end{cases}$$

По предположению $\Delta_k=0$, следовательно, однородная система линейных уравнений имеет нетривиальное решение. Далее умножим последовательно первое уравнение на x_1, второе – на x_2, последнее уравнение – на x_k и сложим все k уравнений. В результате получим равенство

$$\sum_{i,j=1}^{k} a_{ij}x_ix_j=0,$$

левая часть которого представляет собой значение квадратичной формы для $x=(x_1,x_2,...,x_k,0,...0)\neq 0$. Это значение равно нулю, что противоречит знакоопределенности формы.

Итак, мы убедились, что $\Delta_i\neq 0,\ i=1,2,\dots,n$. Поэтому мы можем применить метод Якоби приведения формы к сумме квадратов и воспользоваться

формулами для вычисления канонических коэффициентов:
$\lambda_1 = \Delta_1,\ \lambda_2 = \frac{\Delta_2}{\Delta_1},\ ...,\ \lambda_n = \frac{\Delta_n}{\Delta_{n-1}}$.
Т. к. квадратичная форма положительно определённая, то $\lambda_1, \lambda_2, ..., \lambda_n > 0$, следовательно, все $\Delta_i > 0$.

Если же $f(x,x)$ – отрицательно определенная форма, то все канонические коэффициенты отрицательны и знаки угловых миноров будут чередоваться, причем $\Delta_1 < 0$.

Достаточность. Пусть все $\Delta_i > 0$, где $i = 1, 2, ..., n$, т.е. угловые миноры отличны от нуля, поэтому мы снова можем использовать метод Якоби.

$\Delta_i > 0$, следовательно, все $\lambda_i > 0$, отсюда по определению следует, что квадратичная форма будет положительно определённой.

Если же знаки Δ_i чередуются и $\Delta_1 < 0$, то все канонические коэффициенты $\lambda_i < 0$, т.е форма будет отрицательно определенной.$\square$

П р и м е р 5. *Дана квадратичная форма:*
$f(x,x) = 2{x_1}^2 + 3{x_2}^2 + {x_3}^2 - 2x_1x_2 + 2x_1x_3 - 2x_2x_3,\ (n = 3)$.
Определить, является ли эта форма знакоопределённой.

Решение.

$$A = \begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix},\ \Delta_1 = 2,\ \Delta_2 = \begin{vmatrix} 2 & 2 \\ 2 & 3 \end{vmatrix} = 2 > 0,\ \Delta_3 = |A| = 1 > 0.$$

Все $\Delta_i > 0$, следовательно, квадратичная форма положительно определённая.

Билинейные и квадратичные формы в комплексном линейном пространстве

Пусть $\mathcal{V}_n$ – комплексное линейное пространство. Комплекснозначную функцию двух аргументов $f(x, y)$, где $x, y \in \mathcal{V}$, будем называть *полуторалинейной формой*, если $\forall\ x, y, z \in \mathcal{V}_n,\ \forall \lambda \in \mathbb{C}$ выполняются соотношения:

1) $f(x + y, z) = f(x, z) + f(y, z)$;
2) $f(x, y + z) = f(x, y) + f(x, z)$;
3) $f(\lambda x, y) = \lambda f(x, y)$;
4) $f(x, \lambda y) = \overline{\lambda} f(x, y)$.

Полуторалинейную форму называют *эрмитовой*, если

$$f(x, y) = \overline{f(y, x)}\ \forall x, y \in \mathcal{V}.$$

Пусть $e_1, e_2, ..., e_n$ - базис в комплексном линейном пространстве. Рассмотрим следующее выражение:

$$f(x, y) = \sum_{i,j=1}^{n} a_{ij} x_i \overline{y}_i \tag{18}$$

где $a_{ij} = f(e_i, e_j)$.

Матрица $A_e = (a_{ij}), i, j = 1, 2, ..., n$ называется *матрицей полуторалинейной формы*, а вид (18) называется *общим видом полуторалинейной формы.*

Компактное представление полуторалинейной формы имеет вид

$$f(x, y) = [x]_e{}^T A_e \overline{[y]}_e = \overline{[y]}_e A^T [x]_e. \tag{19}$$

Ранг полуторалинейной формы – это ранг её матрицы.

Полуторалинейная форма называется *вырожденной*,
если $\mathrm{rg} f(x, y) < \dim(\mathcal{V}_n)$.

Теорема 12. *Полуторалинейная форма является эрмитовой тогда и только тогда, когда её матрица в любом базисе является эрмитовой.*

Теорема 13. *Матрицы полуторалинейной формы $f(x, y)$ в базисах e и f A_e и A_f связаны соотношением*

$$A_f = P^T{}_{e \to f} A_e \bar{P}_{e \to f}$$

Пусть $\mathcal{V}_n$ - комплексное линейное пространство, а $f(x, y)$ – эрмитовая полуторалинейная форма. Числовая вещественнозначная функция $f(x, x)$, которая получается из эрмитовой полуторалинейной формы заменой y на x, $x \in \mathcal{V}_n$, называется *эрмитовой квадратичной формой.* Соответственно $f(x, y)$ называется полярной полуторалинейной формой к эрмитовой форме.

Эрмитова форма может быть представлена в виде

$$f(x,x) = \sum_{i,j=1}^{n} a_{ij} x_i \overline{x_j},$$

$a_{ij} = \overline{a}_{ij}$.

Запись в компактном виде:

$$f(x,x) = [x]_e^T \cdot A_e \cdot \overline{[x]}_e = \overline{[x]}_e \cdot A_e{}^T \cdot [x]_e.$$

Канонический вид квадратичной формы в комплексном линейном пространстве.

$$f(x,x) = \lambda_1 |x_1|^2 + \lambda_2 |x_2|^2 + ... + \lambda_r |x_r|^2,$$

где r - ранг квадратичной формы, $\dim(\mathcal{V}_n) = n$.

В отличие от вещественного случая мы выделяем полный квадрат модуля.

Остаются справедливыми и метод Якоби, закон инерции квадратичных форм и критерий Сильвестра.

П р и м е р 6. *Составить матрицу данной эрмитовой полулинейной формы в двумерном пространстве и записать соответствующую квадратичную форму. Определить по критерию Сильвестра вид формы.*

$f(x,y) = 2x_1\overline{y_1} + (1+i)x_1\overline{y_2} + (1-i)x_2\overline{y_1} - 5x_2\overline{y_2}$.

Решение.

$\begin{pmatrix} 2 & 1+i \\ 1-i & -5 \end{pmatrix}$.

$f(x,x) = 2|x_1|^2 + (1+i)x_1\overline{x_2} + (1-i)x_2\overline{x_1} - 5|x_2|^2$,

$\Delta_1 = 2 > 0, \Delta_2 = -12 < 0$, форма является знакопеременной.

Билинейные и квадратичные формы в евклидовом пространстве

Пусть билинейная форма задана в евклидовом пространстве $\mathcal{E}_n$.

Лемма. *Пусть $f(x)$ – линейная форма, рассматриваемая в вещественном евклидовом пространстве $\mathcal{E}_n$. Тогда существует единственный элемент $h \in \mathcal{E}_n$, такой, что выполняется:*

$$f(x) = (x, h), \qquad \forall x \in \mathcal{E}_n. \tag{15}$$

Д о к а з а т е л ь с т в о. 1) Рассмотрим произвольный ортонормированный базис (ОНБ) $e_1, e_2, ..., e_n$.

$$x = \sum_{i=1}^{n} x_i e_i$$

Возьмем $h = (h_1, h_2, ..., h_n)$ и определим компоненты $h_k = f(e_k)$.

$$f(\sum_{i=1}^{n} x_i e_i) = \sum_{i=1}^{n} x_i f(e_i) = \sum_{i=1}^{n} x_i h_i,$$

h - элемент пространства, следовательно, он может быть разложен по базису:

$$h = \sum_{i=1}^{n} h_i e_i \Rightarrow f(\sum_{i=1}^{n} x_i e_i) = (x, h).$$

2) Допустим, h не единственное, т.е. существуют h_1 и h_2, такие, что $\forall x \in \mathcal{E}_n$
$(x_1, h_1) = (x_1, h_2)$,
$(x_1, h_1) - (x_1, h_2) = 0$,
$(x_1, h_1 - h_2) = 0$,
т. к. x - произвольный, предположим, что он равен $h_1 - h_2$. Получаем: $(h_1 - h_2, h_1 - h_2) = 0$. Это возможно тогда и только тогда, когда $h_1 - h_2 = \theta$, следовательно, $h_1 = h_2$. $\square$

Теорема 14. *Пусть $f(x, y)$ - билинейная квадратичная форма, определённая в евклидовом пространстве $\mathcal{E}_n$, тогда существует единственный оператор $\varphi : \mathcal{E}_n \to \mathcal{E}_n$, такой, что справедливо равенство:*

$$f(x, y) = (x, \varphi y) \qquad \forall x, y \in \mathcal{E}_n \tag{16}$$

Д о к а з а т е л ь с т в о. 1) Зафиксируем элемент y и применим лемму, рассмотренную выше. Существует h, для которого выполняется равенство $h = \varphi y$. Из свойств билинейной формы и скалярного произведения следует

данное равенство.
2) Пусть существует два таких $\varphi_1, \varphi_2, : \mathcal{E}_n \to \mathcal{E}_n$.
$\forall x, y\ (x, \varphi_1 y) = (x, \varphi_2 y)$,
$(x, \varphi_1 y) - (x, \varphi_2 y) = 0$,
$(x, (\varphi_1 - \varphi_2)y) = 0$.
Т. к. x – любой, предположим, что $x = \varphi_1 y - \varphi_2 y$.
$\varphi_1 y - \varphi_2 y, (\varphi_1 - \varphi_2)y) = 0$ справедливо при $\varphi_1 y - \varphi_2 y = 0$,
$\varphi_1 y = \varphi_2 y$ (равенство двух операторов),
$\varphi_1 = \varphi_2$. $\square$

Следствие. Наряду с равенством (16) справедливо

$$f(x, y) = (\varphi x, y) \qquad \forall x, y \in \mathcal{E}_n \tag{17}$$

Теорема 15. *Пусть $f(x, y)$ – билинейная форма, определённая в евклидовом пространстве $\mathcal{E}_n$ и пусть $[f]_e = B$ – матрица линейного оператора, фигурирующего в равенстве (17), причем e – ортонормированный базис. Тогда*

$$a_{ij} = b_{ij},\ i, j = 1, 2, ..., n,$$

где a_{ij} – элементы матрицы билинейной формы в этом базисе.

Д о к а з а т е л ь с т в о. $a_{ij} = (\varphi e_i, e_j) = (\sum\limits_{k=1}^{n} b_{ik} e_k e_j) =$ (по определению)

$$\begin{cases} \varphi e_1 = b_{11}e_1 + b_{12}e_2 + \ldots + b_{1n}e_n, \\ \varphi e_2 = b_{21}e_1 + b_{22}e_2 + \ldots + b_{2n}e_n, \\ \qquad \cdots \quad \cdots \quad \cdots \\ \varphi e_n = b_{n1}e_1 + b_{n2}e_2 + \ldots + b_{nn}e_n, \end{cases}$$

т. к. e – ортонормированный, т.е. $= b_{ij}$ при $k = j$, $a_{ij} = b_{ij}$ $\square$.

Теорема 16. $f(x, y)$ *– билинейная форма, определённая в евклидовом пространстве $\mathcal{E}_n$ – является симметричной тогда и только тогда, когда оператор φ фигурирующий в (17), является самосопряженным*

$$f(x, y) = f(y, x) \Leftrightarrow (\varphi x, y) = (x, \varphi y).$$

Д о к а з а т е л ь с т в о.
Необходимость. $(\varphi x, y) = f(x, y) = f(y, x) = (\varphi x, y) - (x, \varphi y)$.
Достаточность. $f(x, y) = (\varphi x, y) = (x, \varphi y) = (\varphi y, x) = f(y, x)$. $\square$

Приведение квадратичной формы к каноническому виду в ортонормированном базисе.

Теорема 17. *Пусть $f(x,y)$ – симметричная билинейная форма, определённая в $\mathcal{E}_n$, тогда существует ортонормированный базис $e_1, e_2, ..., e_n$ и вещественные числа $\lambda_1, \lambda_2, ..., \lambda_n$, такие, что $\forall x \in \mathcal{E}_n$ справедливо представление*

$$f(x,x) = \sum_{i=1}^{n} \lambda_i {x_i}^2,$$

где$[x]_e = (x_1, x_2, ..., x_n)^T$.

Д о к а з а т е л ь с т в о. Для билинейной формы справедливо $f(x,y) = (\varphi x, y)$; т. к. она симметричная, то по предыдущей теореме φ будет самосопряжённым. Для самосопряжённого оператора
$x = \sum\limits_{i=1}^{n} x_i e_i$,
$\varphi x = \lambda x$, следовательно, $\varphi x = \sum\limits_{i=1}^{n} \lambda_i x_i e_i$, $f(x,x) = (\varphi x, x)$, т. к. базис $e_1, e_2, ..., e_n$ – ортонормированный, то получим:

$$(\varphi x, x) = f(x,x) = \sum_{i=1}^{n} \lambda_i {x_i}^2. \ \square$$

Пусть A – матрица линейного оператора и матрица квадратичной формы в базисе e.
$A = [\varphi]_e = [f(x,x)]_e$
$\Lambda = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$,
$\Lambda = P^{-1}{}_{e\to e'} \cdot A \cdot P_{e\to e'}$, где $P_{e\to e'}$ – ортогональная, т.е. $P^{-1}{}_{e\to e'} = P^T_{e\to e'}$.

П р и м е р 7. *Найти канонический вид квадратичной формы $f(x,x) = 5x_1^2 + 8x_1x_2 + 5$, к которой она приводится ортогональным преобразованием и указать одно из таких ортогональных преобразований ($n = 2$).*

Решение.
$A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$
$\begin{vmatrix} 5-\lambda & 4 \\ 4 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - 16 = 25 - 10\lambda + \lambda^2 - 16 = \lambda^2 - 10\lambda + 9,$
$\lambda_1 = 1, \ \lambda_2 = 9,$
$f(x', x') = (x'_1)^2 + 9(x'_2)^2$

Для того, чтобы найти ортогональное преобразование, с помощью которого форма приводится к каноническому виду, необходимо найти соответствующие собственные векторы собственных значений. Проверить ортогональность и пронормировать. Соответствующие векторы выписать по столбцам (матрица перехода $e \to e'$, e – канонический базис).
$x_e = P_{e\to e'} x_e$,

$P_{e\to e'} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.
Связь x и x':
$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$ - связь координат.

Одновременное приведение двух квадратичных форм к каноническому виду.

Теорема 18. *Пусть $f(x, y)$ и $g(x, y)$ - симметричные билинейные формы, определённые в линейном пространстве, причём квадратичная форма, полученная из билинейной формы $g(x, y)$ является положительно определённой. Тогда существует базис $e_1, e_2, ..., e_n$, в котором справедливо представление*

$$f(x, x) = \sum_{i=1}^{n} \lambda x_i^2,$$

$$g(x, x) = \sum_{i=1}^{n} x_i^2.$$

Доказательство.

Так как $g(x, y)$ является полярной положительной формой, то по теореме $(x, y) = g(x, y)$. Введя таким образом скалярное произведение, мы переходим в $\mathcal{E}_n$. Для $\mathcal{E}_n$ по теореме (представление квадратичной формы в каноническом виде в ортонормированном базисе): $f(x, x) = \sum \lambda x_i^2$.

В ортонормированном базисе $(x, x) = g(x, x)$.

$$f(x, x) = \sum_{i=1}^{n} x_i^2. \ \square$$

Пусть $A = [f(x, x)]_e$, $g(x, x)$ – положительно определенная,
$\Lambda = [f(x, x)]_{e'}$,
$[g(x, y)]_e = B$, $[g(x, x)]_{e'} = E$,

$\Lambda = P^T_{e\to e'} A P_{e\to e'}$,
$A = (P^T_{e\to e'})^{-1} \Lambda (P_{e\to e'})^{-1}$,
$E = P^T_{e\to e'} B P_{e\to e'}$,
$B = (P^T_{e\to e'})^{-1} E (P_{e\to e'})^{-1}$,
$(AB)^{-1} = b^{-1} A^{-1}$,
$B^{-1} = P_{e\to e'} P^T_{e\to e'}$,
$B^{-1}A = P_{e\to e'} P^T_{e\to e'} (P^T_{e\to e'})^{-1} \Lambda (P_{e\to e'})^{-1} = P_{e\to e'} \Lambda (P_{e\to e'})^{-1}$,
$B^{-1} A P_{e\to e'} = P_{e\to e'} \Lambda$ – получили определение собственных значений собственных векторов матрицы $(B^{-1}A)$.

Λ – матрица из всех собственных значений.

$P_{e\to e'}$ – матрица, состоящая из собственных векторов.

$$|B^{-1}A - \lambda E| = 0,$$

$$|A - \lambda B| = 0.$$

Возможен вариант с очевидными исправлениями в выводе данной формулы, когда
квадратичная форма задана следующим образом: $g(x, x) = \sum \mu_i x_i^2$

$\Lambda_2 = \begin{pmatrix} \mu_1 & ... & 0 \\ ... & ... & ... \\ 0 & ... & \mu_n \end{pmatrix}$

П р и м е р 8. *Проверить, что по меньшей мере одна из двух данных форм является знакоопределённой. Найти замену координат, приводящих эти две формы к нормальному, и записать канонический вид этих форм.*
$g = x_1^2 + 2x_1x_2 + 3x_2^2$,
$f = 4x_1^2 + 16x_1x_2 + 6x_2^2$.

Решение.
$B = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$, $\quad A = \begin{pmatrix} 4 & 8 \\ 8 & 6 \end{pmatrix}$,
$B^{-1} = \frac{1}{2} \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix} \quad |A - \lambda B| = 0, \quad \lambda_1 = 5, \quad \lambda_2 = -4$,
$f(x', x') = 5{x'}_1^2 - 4{x'_2}^2$,
$g(x', x') = {x'}_1^2 + {x'_2}^2$.

Гиперповерхности второго порядка

Понятие гиперповерхности второго порядка. Пусть V — n-мерное вещественное евклидово пространство.

Ради геометрической наглядности будем называть векторы x этого пространства т о ч к а м и.

Г и п е р п о в е р х н о с т ь ю S второго порядка будем называть геометрическое место точек x, удовлетворяющих уравнению вида

$$A(x, x) + 2B(x) + c, \tag{1}$$

где $A(x, x)$ — не равная тождественно нулю квадратичная форма, $B(x, x)$ — линейная форма, а c — вещественное число.

Уравнение (1) будем называть о б щ и м   у р а в н е н и е м   г и п е р п о в е р х н о с т и   в т о р о г о   п о р я д к а.

Выделим в пространстве V какой-либо ортонормирванный базис e_k. Координаты вектора x (точки x) в этом базисе обозначим через $(x_1, x_2, ..., x_n)$. Тогда квадратичная форма $A(x, x)$ может быть представлена в виде

$$A(x, x) = \sum_{j,\ k=1}^{n} a_{jk} x_j x_k, \tag{2}$$

где

$$a_{jk} = A(e_j, e_k) \tag{3}$$

и $A(e_j, e_k)$ — значение на векторах e_j и e_k симметричной билинейной формы $A(x, y)$, полярной квадратичной форме $A(x, x)$.

Линейная форма $B(x)$ в указанном базисе e_k представляется в виде

$$B(x) = \sum_{k=1}^{n} b_k x_k. \tag{4}$$

Таким образом, *общее уравнение гиперповерхности второго порядка в евклидовом пространстве V с выделенным базисом e_k может быть представлено в следующей форме:*

$$\sum_{j,\ k=1}^{n} a_{jk} x_j x_k + 2\sum_{k=1}^{n} b_k x_k + c = 0. \tag{5}$$

Договоримся о следующей терминологии.

Слагаемое $A(x, x) = \sum\limits_{j,\ k=1}^{n} a_{jk} x_j x_k$ будем называть г р у п п о й   с т а р ш и х   ч л е н о в уравнения (1) или (5).

Группу слагаемых $B(x) + c = \sum\limits_{k=1}^{n} b_k x_k + c$ будем называть л и н е й н о й   ч а с т ь ю уравнения (1) или (5).

Мы будем рассматривать в дальнейшем матрицы

$$A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ ... & ... & ... \\ a_{n1} & ... & a_{nn} \end{pmatrix} \text{ и } B = \begin{pmatrix} a_{11} & ... & a_{1n} & b_1 \\ ... & ... & ... & ... \\ a_{n1} & ... & a_{nn} & b_n \\ b_1 & ... & b_n & c \end{pmatrix}$$

и определители $\det A$ и $\det B$ этих матриц.

Исследование гиперповерхностей второго порядка мы будем проводить с помощью метода, сходного с методом, применяемым в аналитической геометрии при исследовании кривых и поверхностей второго порядка, заданных общими уравнениями.

Идея этого метода заключается в том, что путем выбора специальной декартовой системы координат на плоскости (для кривых второго порядка) или в пространстве (для поверхностей второго порядка) достигается максимальное упрощение уравнения кривой или поверхности. Затем путем исследования этого уравнения выясняются геометрические свойства кривой или поверхности. Кроме того, перечисление всех возможных типов простейших (канонических) уравнений кривых или поверхностей второго порядка позволяет дать их классификацию.

Чтобы использовать этот метод в многомерном случае, мы сначала должны изучить такие преобразования (отображения) n-мерном евклидова пространства, которые представляют собой аналоги преобразований декартовых прямоугольных координат в случае двух и трех измерений.

Очевидно, гиперповерхность второго порядка, рассматриваемая как геометрический объект пространства V, не изменяется, если производится преобразование указанного выше вида. Ниже мы убедимся, что для каждого уравнения вида (1) (или (5)) можно выбрать такое начало координат и выбрать такой ортонормированный базис в V, что это уравнение, записанное в координатах относительно нового базиса, будет максимально простого вида, и поэтому, как и в случае двух и трех измерений, можно будет указать геометрические характеристики таких поверхностей и дать им классификацию.

2. Параллельные переносы в евклидовом пространстве. Преобразования ортонормированных базисов в ортонормированные. *Параллельным переносом* в евклидовом пространстве V мы будем называть преобразование, задаваемое формулами

$$x = x' + x^\circ, \tag{7}$$

где x — фиксированная точка, называемая новым началом координат.

Пусть точки x, x' и x° имеют координаты, соответственно равные $(x_1, x_2, ..., x_n)$, $(x'_1, x'_2 ..., x'_n)$ и $(x^\circ_1, x^\circ_2, ..., x^\circ_n)$.

Тогда в координатах параллельный перенос определяется формулами

$$x_k = x'_k + x^\circ_k, \qquad k = 1, 2, ..., n. \tag{8}$$

Отметим, что при параллельном переносе любой фиксированный базис не изменяется.

Перейдём теперь к выяснению характеристики преобразования ортонормированного базиса в ортонормированный.

Допустим, что ортонормированный базис $\{e_k\}$ преобразуется в новый ортонормированный базис $\{e'_k\}$. Разложим каждый вектор $\{e'_k\}$ по векторам $\{e_k\}$. Получим

$$\left.\begin{array}{lcl} e'_1 & = & p_{11}e_1 + p_{21}e_2 + ... + p_{n1}e_n, \\ e'_2 & = & p_{12}e_1 + p_{22}e_2 + ... + p_{n1}e_n, \\ ... & ... & \qquad\qquad\qquad\qquad\qquad ... \\ e'_n & = & p_{1n}e_1 + p_{2n}e_2 + ... + p_{nn}e_n. \end{array}\right\} \tag{9}$$

Обозначим буквой P матрицу преобразования (9):

$$P = \begin{pmatrix} p_{11} & p_{21} & ... & p_{n1} \\ p_{12} & p_{22} & ... & p_{n2} \\ ... & ... & ... & ... \\ p_{1n} & p_{2n} & ... & p_{nn} \end{pmatrix}. \tag{10}$$

Так как базисы $\{e_k\}$ и $\{e'_k\}$ ортонормированные, то из (9) путём скалярного умножения e'_j и e'_k получим

$$(e'_j, e'_k) = \sum_{m=1}^{n} p_{mj}p_{mk} = \delta_{jk} = \left\{ \begin{array}{ccccc} 1 & , & j & = & k, \\ 0 & , & j & \neq & k. \end{array}\right.$$

Рассмотрим теперь транспониролванную матрицу P', т. е. матрицу, полученную из P перестановкой строк и столбцов.

Очевидно,

$$PP' = P'P = I, \tag{12}$$

где I — единичная матрица.

Равенства (12) показывают, что матрица P' является обратной для матрицы P, т. е.

$$P^{-1} = P'. \tag{13}$$

Допустим теперь, что мы рассматриваем преобразование ортонормированног базиса $\{e_k\}$ по формулам (9), причём матрица P этого преобразования удовлетворяет условию (12) (или, что то же, (13)).

Тогда, очевидно, элементы p_{jk} матрицы P удовлетворяют условию (11), что, согласно, этим же соотношениям (11), эквивалентно условию ортонормированности базиса $\{e'_k\}$

Напомним, что в матрицу P, удовлетворяющую условию (12), мы назвали ортогональной.

Итак, *для того чтобы преобразование (9) было преобразованием ортонормированного базиса в ортонормированный, необходимо и достаточно, чтобы матрица P этого преобразования была ортогональной.*

Замечание. Обращаясь к формулам преобразования координат вектора при преобразовании базиса и учитывая, что обратная матрица для ортогональной матрицы P есть матрица P', получим следующие формулы преобразования координат точки x при переходе от ортонормированного базиса к ортонормированному:

$$\left.\begin{array}{ccccc} x_1 & = & p_{11}x'_1 + p_{12}x'_2 + & \dots & +p_{1n}x'_n, \\ x_2 & = & p_{21}x'_1 + p_{22}x'_2 + & \dots & +p_{2n}x'_n, \\ \dots & \dots & \dots & \dots & \dots \\ x_n & = & p_{n1}x'_1 + p_{n2}x'_2 + & \dots & +p_{nn}x'_n. \end{array}\right\} \tag{14}$$

Преобразование общего уравнения гиперповерхности второго порядка при параллельном переносе. Рассмотрим параллельный перенос, который определяется как преобразование пространства V по формуле (7)(или в координатах по формуле (8)).

Левая часть (1) после подстановки вместо x его выражения по формуле (7) в силу линейности квадратичной формы по первому и второму аргументу и свойств линейной формы примет вид:

$$A(x', x') + 2[A(x', x^\circ) + B(x')] + [A(x^\circ, x^\circ) + 2B(x^\circ) + c] = 0.$$

Итак, общее уравнение (1) гиперповерхности S при параллельном переносе (7) запишется в форме

$$A(x', x') + 2B'(x') + c' = 0, \tag{15}$$

где линейная форма $B'(x')$ и постоянное число c' определяются соотношениями

$$B'(x') = A(x', x^\circ) + B(x'), \tag{16}$$

$$c' = A(x^\circ, x^\circ) + 2B(x^\circ). \tag{17}$$

Запишем полученные формулы в координатах.

Пусть координаты точек x' и x° равны соответственно x'_1, x'_2, ..., x'_n и x°_1, x°_2, ..., x°_n. Так как при параллельном переносе базис $\{e_k\}$ не меняется, то квадратичная форма $A(x', x')$ запишется следующим образом:

$$A(x', x') = \sum_{k=1}^{n} a_{jk}x'_j x'_k \tag{18}$$

(отметим, что коэффициенты $a_{jk} = A(e_j, e_k)$ не меняются, так как не меняются базисные векторы e_k).

Следоватльно, мы можем сделать важный вывод: *при параллельном переносе группа старших членов сохраняет свой вид.*

Займёмся теперь формулами (16) и (17). Так как

$$A(x', x^\circ) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{jk} x_j^\circ \right) x_k^\circ,$$

$$B(x') = \sum_{k=1}^{n} b_k x'_k,$$

$$A(x^\circ, x^\circ) = \sum_{j,k=1}^{n} a_{jk} x_j^\circ x_k^\circ,$$

$$B(x^\circ) = \sum_{k=1}^{n} b_k x_k^\circ,$$

то формула (16) примет вид

$$B'(x') = \sum_{k=1}^{n} b'_k x'_k = \sum_{k=1}^{n} \left[\sum \ limits_{j=1}^{n} a_{jk} x_j^\circ + b_k \right] x_k^\circ, \tag{7.19}$$

а формула (17) запишется следующим образом:

$$c' = \sum_{j,k=1}^{n} a_{jk} x_j^\circ x_k^\circ + 2 \sum_{k=1}^{n} b_k x_k^\circ + c. \tag{20}$$

Таким образом, уравнение (15) в координатах будет иметь следующий вид:

$$\sum_{j,k=1}^{n} a_{jk} x'_j x'_k + 2 \sum_{k=1}^{n} b_k x'_k + c' = 0. \tag{21}$$

Нам понадобится несколько иное, чем (20), выражение для c'. Запишем (20) в следующей форме:

$$c' = \sum_{k=1}^{n} \left[\sum_{j=1}^{n} a_{jk} x_j^\circ + b_k \right] x_k^\circ + \sum_{k=1}^{n} b_k x_k^\circ + c. \tag{22}$$

Учитывая, что коэффициенты b'_k выражаются, как это следует из (19), по формулам

$$b'_k = \sum_{j=1}^{n} a_{jk} x_j^\circ + b_k, \tag{23}$$

мы получим из (22) нужное нам выражение для c':

$$b'_k = \sum_{k=1}^{n} (b'_k + b_k) x_k^\circ + c. \tag{24}$$

Преобразование общего уравнения гиперповерхности второго порядка при переходе от ортонормированного базиса к ортонормированному. Пусть ортонормированный базис $\{e_k\}$ преобразуется в новый ортонормированный базис $\{e'_k\}$ по формулам (9) и P— ортогоналбная матрица этого преобразования (см. (10)). Тогда, согласно замечанию в п. 2 этого параграфа, координаты x_k и x'_k точки в базисах $\{e_k\}$ и $\{e'_k\}$ связаны соотношениями (5). Подставляя выражение для x_k из (14) в левую часть уравнения и учитывая, что вследствие однородности соотношений (14) группа старших членов и линейная часть уравнения (5) преобразуется автономно, получим следующее выражение для общего уравнения гиперповерхности второго порядка в координатах x'_k точек в преобразованном базисе $\{e'_k\}$:

$$\sum_{j,k=1}^{n} a'_{jk}x'_jx'_k + 2\sum_{k=1}^{n} b'_kx'_k + c' = 0. \qquad (15)$$

Согласно отмеченной выше автономности преобразования группы старших членов, справедливы равенства

$$\left.\begin{array}{c} \sum\limits_{j,k=1}^{n} a'_{jk}x'_jx'_k = \sum\limits_{j,k=1}^{n} a_{jk}x_jx_k, \\ \sum\limits_{k=1}^{n} b'_kx'_k = \sum\limits_{k=1}^{n} b_kx_k, \\ c' = c. \end{array}\right\} \qquad (26)$$

Обращаясь к первой из формул (26), мы видим, что для определения коэффициентов a'_{jk} можно воспользоваться правилом преобразования коэффициентов квадратичной формы при переходе к новому базису. Именно, если обозначим буквой A' матрицу квадратичной формы $A(x,x)$ в базисе $\{e'_k\}$ то, согласно теореме 7.2 и соотношению $P' = P^{-1}$, *получим следующую связь между матрицами* A и A' *формы* $A(x,x)$ в базисах e_k и e'_k:

$$A' = P^{-1}AP \qquad (27)$$

(напомним, что P — матрица ортогонального преобразования).

Будем рассматривать теперь матрицу A' как матрицу некоторого линейного оператора A в базисе $\{e'_k\}$, а матрицу P^{-1} как матрицу перехода от базиса $\{e'_k\}$ к $\{e_k\}$. Тогда, согласно теореме матрицу A можно рассматривать как матрицу этого линейного оператора A в базисе $\{e_k\}$.

Иными словами, *матрица квадратичной формы при преобразовании ортонормированного базиса в ортонормированный изменяется как матрица некоторого линейного оператора.*

Этот вывод мы используем в следующем пункте.

З а м е ч а н и е. Отметим, что оператор A, матрица которого в ортонормированном базисе совпадает с матрицей квадратичной формы $A(x,x)$, *самосопряжённый*

Для доказательства проведём следующие рассуждения.

Пусть $A(x,x)$ — квадратичная форма и $A(x,y)$ — симметричная билинейная форма, полярная форме $A(x,x)$. Согласно теореме 7.8 билинейная форма $A(x,y)$ может быть представлена в виде

$$A(x,y) = (Ax,y),$$

где A — *самосопряжённый* оператор.

Поэтому квадратичная форма $A(x,x)$ может быть представлена в виде

$$A(x,x) = (Ax,x).$$

Докажем, что в ортонормированном базисе $\{e_k\}$ матрицы оператора A и квадратичной формы совпадают. Этим будет доказано утверждение замечания.

Пусть a_{jk} — элементы матрицы формы $A(x,x)$ и $\tilde{a}_{jk}$ — элементы матрицы оператора A в базисе $\{e_k\}$. Согласно п. 2 §1 этой главы

$$a_{jk} = A(e_j,e_k),$$

а элементы $\tilde{a}_{jk}$ могут быть найдены из равенств

$$Ae_j = \sum_{p=1}^{n} \tilde{a}_{jp}e_p.$$

Умножим обе части последнего соотношения скалярно на e_k. Тогда, учитывая ортонормированность базиса $\{e_k\}$, получим

$$(Ae_j, e_k) = \tilde{a}_{jk}.$$

Так как $A(e_j, e_k) = (Ae_j, e_k)$, то $a_{jk} = \tilde{a}_{jk}$. Утверждение замечания доказано.

5. Инварианты общего уравнения гиперповерхности второго порядка. Назовём и н в а р и а н т о м общего уравнения (1) (или (5)) гиперповерхности второго порядка относительно параллельных переносов и преобразований ортогональных базисов в ортогональные такую функцию $f(a_{11}, a_{12}, ..., a_{nn}, b_1, ..., b_n, c)$ коэффициентов этого уравнения, значение которой не меняется при указаных преобразованиях пространства.

Докажем следующее утверждение:

Теорема 7.11. *Инвариантами общего уравнения (1)(или (5)) гиперповерхности второго порядка являются коэффициенты характеристического многочлена матрицы A квадратичной формы $A(x, x)$ и определитель* $\det B$ *матрицы B в соотношении (6). В частности, инвариантами являются* $\det A$ *и след* $a_{11} + a_{22} + ... + a_{nn}$ *матрицы A.*

Д о к а з а т е л ь с т в о. Очевидно, инвариантность перечисленных в условии теоремы величин достаточно доказать отдельно для параллельного переноса и преобразования ортонормированного базиса в ортонормированный.

Рассмотрим сначала *параллельный перенос.* В п. 3 этого параграфа мы установили, что при этом преобразовании группа старших членов сохраняет свой вид (см. формулу (18)). Поэтому не меняется матрица A, а следовательно, и характеристический многочлен этой матрицы.

Докажем инвариантность $\det B$.

При параллельном переносе (7) (или (8)) матрица преобразуется в матрицу B', определитель которой, согласно (11), имеет вид

$$\det B' = \begin{vmatrix} a_{11} & ... & a_{1n} & b'_1 \\ ... & ... & ... & ... \\ a_{n1} & ... & a_{nn} & b'_n \\ b'_1 & ... & b'_n & c' \end{vmatrix}, \qquad (28)$$

где величины b'_k c' определяются по формулам (23) и (24).

Вычтем из элементов последней $(n+1)$-й строки определителя (28) элементы первой строки, умноженные на x_1°, затем элементы второй строки, умноженные на x_2°, и т. д., наконец, элементы n-й строки, умноженные на x_n°. Так как при таких преобразованиях определитель не меняется, то, используя (23) и (24), получим соотношение

$$\det B = \begin{vmatrix} a_{11} & ... & a_{1n} & b'_1 \\ ... & ... & ... & ... \\ a_{n1} & ... & a_{nn} & b'_n \\ b_1 & ... & b_n & \left(\sum\limits_{k=1}^{n} b_k x_k^\circ + c\right) \end{vmatrix}$$

Вычтем теперь из элементов последнего $(n+1)$-го столбца определителя (29) элементы первого столбца, умноженные на x_1°, затем элементы второго столбца, умноженные на x_2°, и т. д., наконец, элементы n-го столбца, умноженные на x_n°. Так как при таких преобразованиях определитель не меняется, то, используя соотношение $a_{jk} = a_{kj}$, вытекающее из симметричности формы $A(x, y)$, и формулу (23), мы получим в результате $\det B$. Итак, равенство $\det B' = \det B$ доказано. Следовательно, $\det B$ инвариантен относительно параллельных переносов.

Рассмотрим теперь *преобразование ортонормированного базиса в ортонормированный.*

Во-первых, убедимся, что коэффициенты характеристического многочлена матрицы A квадратичной формы являются инвариантами рассматриваемого преобразования.

В предыдущем пункте мы установили, что при переходе к новому ортонормированному базису матрица A изменяется как матрица некоторого линейного оператора. Но в таком случае, как следует из замечания 1 п. 3, §2, гл. 5, коэффициенты характеристического многочлена этой матрицы не меняются при переходе к другому базису.

В частности, определитель $\det A$ и след $a_{11} + a_{22} + ... + a_{nn}$ матрицы A, как коэффициенты характеристического многочлена, являются инвариантами.

Нам останется доказать инвариантность определителя $\det B$ при преобразовании ортонормированного базиса в ортонормированный.

Приступим к этому доказательству.

Применим следующий приём. Введём обозначения $b_k = a_{k,n+1}, k = 1, 2, ..., n, \quad c = a_{n+1,n+1}$. Тогда уравнение (5) гиперповерхности можно записать следующим образом:

$$\sum_{j,k=1}^{n+1} a_{jk} x_j x_k = 0, \tag{7.91}$$

где $x_{n+1} = 1$.

Рассмотрим преобразование переменных $x_1, x_2, ..., x_n, x_{n+1}$ в переменные $x'_1, x'_2, ..., x'_n, x'_{n+1}$, при котором первые n переменных преобразуются по формулам (14), а переменная x_{n+1} преобразуется по формуле

$$x_{n+1} = x'_{n+1}.$$

Ясно, что это преобразование переменных можно рассматривать как преобразование координат при преобразовании ортонормированного базиса $e_1, e_2, ..., e_n, e_{n+1}$ $(n + 1)$-мерного евклидова пространства, причём матрица P этого преобразования имеет вид

$$\tilde{P} = \begin{pmatrix} p_{11} & ... & p_{n1} & 0 \\ ... & ... & ... & ... \\ p_{1n} & ... & p_{nn} & 0 \\ 0 & ... & 0 & 1 \end{pmatrix}. \tag{31}$$

Легко видеть, что матрица $\tilde{P}$ удовлетворяет условию

$$\tilde{P'} = \tilde{P^{-1}}$$

и поэтому является ортогональной. Но тогда, согласно п. 2 этого параграфа, ортонормированный базис $e_1, e_2, ..., e_n, e_{n+1}$ преобразуется с помощью матрицы $\tilde{P}$ в ортонормированнный базис. Выше было выяснено, что при таком преобразовании матрицы B квадратичной формы определитель $\det B$ этой матрицы представляет собой инвариант. Теорема доказана.

З а м е ч а н и е. Из рассуждений в доказательстве теоремы следует, что инвариантами общего уравнения гиперповерхности второго порядка будут также величины $\operatorname{rang} A$ и $\operatorname{rang} B$.

6. Центр гиперповерхности второго порядка. Попытаемся найти такой параллельный перенос, при котором общее уравнение (15) не содержало бы слагаемое $2B'(x')$ (или, если обратиться к уравнению (21), то слагаемых $2\sum\limits_{k=1}^{n} b'_k x'_k$).

Иными словами, будем искать параллельный перенос (т. е. координаты x_1°, x_2°, ..., x_n° точки x), при котором обратятся в нуль все коэффициенты b_k. Обращаясь к формулам (23), найдём, что искомые координаты x_1°, x_2°, ..., x_n° точки x представляют собой решение следующей системы линейных уравнений:

$$\sum_{j=1}^{n} a_{jk} x_j^\circ + b_k = 0, \quad k = 1, 2, ..., n. \tag{32}$$

Уравнения (32) называются *уравнениями центра гиперповерхности второго порядка*, а точка x° с координатами x_1°, x_2°, ..., x_n°, где x_1°, x_2°, ..., x_n° — решение системы (32), называется ц е н т р о м этой поверхности.

Поясним наименование "центр"гиперповерхности. Пусть начало координат помещено в центр $x_\circ$, т. е. произведён искомый параллельный перенос. Тогда уравнение поверхности S примет вид

$$\sum_{j,k=1}^{n} a_{jk}x_jx_k + c' = 0. \tag{33}$$

Пусть точка x с координатами $(x'_1, x'_2, ..., x'_n)$ расположена на S. Это означает, что её координаты удовлетворяют уравнению (33). Очевидно, точка $-x$ с координатами $(-x'_1, -x'_2, ..., -x'_n)$, симметричная с точкой x относительно точки x, также расположена на S, ибо её координаты тоже удовлетворяют уравнению (33).

Таким образом, если у гиперповерхности S есть центр, то *относительно центра точки S располагаются парами.*

З а м е ч а н и е 1. Если гиперповерхности S второго порядка имеет центр, то инвариантны $\det A$, $\det B$ и свободный член c' в уравнении (33) связаны соотношением

$$\det B = c' \det A. \tag{34}$$

Действительно, для уравнения (33) получим

$$\det B = \begin{vmatrix} a_{11} & ... & a_{1n} & 0 \\ ... & ... & ... & ... \\ a_{n1} & ... & a_{nn} & 0 \\ 0 & ... & 0 & c' \end{vmatrix}.$$

Из последней формулы и вытекает (34).

Наличие центра у гиперповерхности второго порядка связано с разрешимостью уравнений центра (32).

Если уравнения центра имеют единственное решение, то гиперповерхность S будем называть ц е н т р а л ь н о й.

Так как определитель системы (32). равен $\det A$, а необходимым и достаточным условием существования единственного решения этой системы является отличие от нуля её определителя, то мы можем сделать следующий вывод: *для того чтобы гиперповерхность S* была центральной, необходимо и достаточно, чтобы $\det A \neq 0$.

З а м е ч а н и е 2. Если начало координат перенесено в центр центральной гиперповерхности S, то уравнение этой гиперповерхности будет иметь вид

$$\sum_{j,k=1}^{n} a_{jk}x_jx_k + \frac{\det B}{\det A} = 0. \tag{35}$$

Действительно, после переносе начала в центр уравнение гиперповерхности примет вид (33). Так как для центральной гиперповерхности $\det A \neq 0$, то из формулы (34) найдём $c' = \det B/detA$. Подставляя это выражение для c' в формулу (33), мы и получим уравнение (35).

Стандартное упрощение любого уравнения гиперповерхности второго порядка путём преобразования ортонормированного базиса. По теореме 7.8 существует такой ортонормированный базис, в котором квадратичная форма $A(x,x)$ записывается в виде суммы квадратов. Обозначим этот базис через $\{e_k\}$, а координаты точки x в этом базисе обозначим через $(x'_1, x'_2, ..., x'_n)$. Кроме того, буквами $\lambda_1, \lambda_2, ..., \lambda_n$ обозначим собственные значения самосопряжённого оператора A, матрица которого в ортонормированном базисе совпадает с матрицей квадратичной формы $A(x,x)$ (см. замечание в п. 4 этого параграфа).

Используя теперь выводы теорем 7.8, запишем квадратичную форму $A(x,x)$ в координатах $(x'_1, x'_2, ..., x'_n)$ точки x в базисе $\{e_k\}$ следующим образом:

$$A(x,x) = \sum_{k=1}^{n} \lambda_k^2 {x'_k}^2. \tag{36}$$

Итак, перейдём от базиса $\{e_k\}$ к базису $\{e'_k\}$. Так как формулы преобразования координат точек при таком преобразовании линейны и однородны (см. замечание п.2 этого параграфа, формулы (24)), то группа старших членов и линейная часть уравнения гиперповерхности S преобразуются автономно. На основании этого и в силу (36) уравнение гиперповерхности S в базисе $\{e'_k\}$ будет иметь следующий вид:

$$\sum_{k=1}^{n} \lambda_k {x'_k}^2 + 2\sum_{k=1}^{n} b'_k x'_k + c = 0. \qquad (37)$$

Приведение любого уравнения гиперповерхности S второго порядка к виду (37) будем называть *стандартным упрощением этого уравнения (путём преобразования ортонормированного базиса).*

8. Упрощение уравнения центральной гиперповерхности второго порядка. Классификация центральных гиперповерхностей. Выводы, сделанные в предыдущих двух пунктах, позволяют решить вопрос о классификации всех центральных гиперповерхностей второго порядка. Решение этого вопроса мы проведём по следующей схеме. Во-первых, путём переноса начала координат в центр гиперповерхности (5) мы приведём её уравнение к виду (35). После этого произведём стандартное упрощение уравнения (35). В результате, очевидно, мы получим, согласно (37), следующее уравнение центральной поверхности второго порядка:

$$\lambda_1 {x''_1}^2 + \lambda_2 {x''_2}^2 + ... + \lambda_n {x''_n}^2 + \frac{\det B}{\det A} = 0, \qquad (38)$$

в котором λ_k — собственные числа матрицы A квадратичной формы $A(x,x)$ в уравнении (1), а x''_k — координаты точки x в окончательном ортонормированном базисе $\{e_k\}$.

Отметим, во-первых, что *все собственные числа $\lambda_k, k = 1, 2, ..., n$, отличны от нуля.*

Действительно, подсчитывая $\det A$ для уравнения (38), получим

$$\det A = \lambda_1 \lambda_2 ... \lambda_n,$$

а так как для центральной поверхности $\det A \neq 0$, то, очевидно, что все $\lambda_k \neq 0$.

Договоримся далее все положительные собственные числа матрицы A нумеровать первыми индексами, а отрицательные — последующими. Таким образом, найдётся такой номер p, что

$$\begin{array}{cccc} \lambda_1 > 0 & \lambda_2 > 0 & ... & \lambda_p > 0 \\ \lambda_{p+1} < 0 & \lambda_{p+2} < 0 & ... & \lambda_n < 0. \end{array}$$

Введём теперь следующие обозначения: если $sgn\frac{\det B}{\det A} \neq 0$, то положим

$$\left.\begin{array}{ll} \left|\frac{\det A}{detB}\right| = \frac{1}{a_k^2} & , \quad k = 1, 2, ..., p, \\ \left|\frac{\det A}{detB}\right| = -\frac{1}{a_k^2} & , \quad k = p+1, ..., n; \end{array}\right\} \qquad (40)$$

Тогда, очевидно, уравнение (38) может быть переписано следующим образом (при этом мы заменим обозначение координат x''_k на x_k):

$$\frac{x_1^2}{a_1^2} + \frac{x_2}{2} + ... + \frac{x_p^2}{a_p^2} - \frac{x_{p+1}^2}{a_{p+1}^2} - ... - \frac{x_n^2}{a_n^2} + sgn\frac{\det B}{\det A} = 0. \qquad (41)$$

Уравнение (41) называется *каноническим уравнением центральной гиперповерхности второго порядка.*

Величины $a_k, k = 1, 2, ..., n$, *называются полуосями центральной гиперповерхности второго порядка.* Они могут быть вычислены по формулам (39) и (40).

. Линейные операторы в евклидовых (унитарных) пространствах
Сопряженный оператор

Линейный оператор $\varphi^* : \mathcal{E}_m \to \mathcal{E}_n(\mathcal{U}_m \to \mathcal{U}_n)$ называют *сопряженным* данному оператору $\varphi : \mathcal{E}_n \to \mathcal{E}_m(\mathcal{U}_n \to \mathcal{U}_m)$, если для $\forall x \in \mathcal{E}_n(\mathcal{U}_n)$,
$\forall y \in \mathcal{E}_m(\mathcal{U}_m),\ \forall \alpha \in \mathbb{C}$ выполняется равенство

$$(\varphi x, y) = (x, \varphi^* y). \tag{23}$$

Из определения сопряженного оператора вытекают следующие его **свойства**:

1. $(\varphi^*)^* = \varphi$;
2. $(\varphi + \psi)^* = \varphi^* + \psi^*$;
3. $(\varphi\psi)^* = \psi^*\varphi^*$;
4. $(\alpha\varphi)^* = \bar{\alpha}\varphi^*$;
5. $(\varphi^*)^{-1} = (\varphi^{-1})^*$ (если φ является преобразованием).

Все свойства доказываются однотипно.

Докажем, например, свойство 3: согласно определению произведения операторов и определения сопряженного оператора получаем

$$((\varphi\psi)x, y) = ((\varphi(\psi x), y) = (\psi x, \varphi^* y) = (x, (\psi^*\varphi^*)y).$$

Выясним, как связаны матрицы операторов φ и φ^* в базисе e в вещественном евклидовом пространстве.

Обозначим соответственно матрицы этих операторов $[\varphi]_e = A$ и $[\varphi^*]_e = A^*$ и пусть для $\forall x, y \in \mathcal{E}_n(\mathcal{U}_n)$ $\quad x_e,\ y_e$ – координатные столбцы векторов x, y в базисе e, тогда равенство (23) можно переписать с учетом, что $(x, y) = x_e^T \Gamma y_e$, где $\Gamma = \begin{pmatrix} (e_1, e_1) & \cdots & (e_1, e_n) \\ \cdots & \cdots & \cdots \\ (e_n, e_1) & \cdots & (e_n, e_n) \end{pmatrix}$ в виде

$$(Ax_e)^T \Gamma y_e = x_e^T \Gamma A^* y_e.$$

Далее $\quad x_e^T A^T \Gamma y_e = x_e^T \Gamma A^* y_e; \quad x_e^T (A^T\Gamma - \Gamma A^*) y_e = 0.$

Так как $x_e,\ y_e$ – произвольные столбцы, отсюда можно заключить, что

$$A^T\Gamma - \Gamma A^* = O,$$

где O – нулевая матрица.

Итак, матрицы операторов φ и φ^* в базисе e связаны соотношением

$$A^* = \Gamma^{-1} A^T \Gamma. \tag{24}$$

В частности, если базис ортонормированный, то $\Gamma = E$ и

$$A^* = A^T. \tag{25}$$

В унитарном пространстве, где $(x, y) = x_e^T \Gamma \bar{y}_e$, формулы (25) и (26) соответственно примут вид

$$A^* = \overline{\Gamma^{-1} A^T \Gamma};$$

$$A^* = \bar{A^T}.$$

Теорема 26. *Каждый линейный оператор в евклидовом пространстве имеет сопряженный оператор, и притом только один.*

П р и м е р 5. *Линейный оператор φ в базисе $e_1' = (1\ 1\ 1)^T$, $e_2' = (1\ 0\ 1\ 1)^T$, $e_3' = (0\ 0\ 1)^T$ имеет матрицу $[\varphi]_{e'} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}$. Найти матрицу $[\varphi^*]_{e'}$, если векторы e_1', e_2', e_3' заданы координатами в ортонормированном базисе e_1, e_2, e_3.*

<u>*Решение.*</u> *Найдем матрицу*

$$\Gamma_{e'} = \begin{pmatrix} (e_1', e_1') & (e_1', e_2') & (e_1', e_3') \\ (e_2', e_1') & (e_2', e_2') & (e_2', e_3') \\ (e_3', e_1') & (e_3', e_2') & (e_3', e_3') \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

$[\varphi]_e' = \Gamma_{e'}^{-1} [\varphi]_{e'}^T \Gamma_{e'}$.

$$\Gamma_{e'}^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

В результате перемножения матриц получим

$$[\varphi]_e' = \begin{pmatrix} 6 & 5 & 3 \\ -3 & -2 & -1 \\ -2 & -2 & -1 \end{pmatrix}.$$

П р и м е р 6. *В трёхмерном евклидовом $\mathcal{E}_3$ пространстве выбран ортонормированный базис e_1, e_2, e_3. Найти преобразование φ^*, сопряжённое преобразованию φ пространства $\mathcal{E}_3$, если преобразование φ, задано формулой*

$$\varphi(x) = [a, x],$$

где a - фиксированный вектор из $\mathcal{E}_3$, $[a, x]$ - векторное произведение векторов a и x.

<u>*Решение.*</u> *По определению сопряженного оператора*

$(\varphi x, y) = ([a, x], y) = axy = xya = (x, -[a, y]) = (x, \varphi^* y) \Rightarrow \varphi^* = -\varphi.$

Областью значений φ^* является подпространство, ортогональное к ядру оператора φ. Это следует из того, что $\forall x \in \ker\varphi,\ \forall y \in \mathrm{im}\varphi$

$$(x, \varphi^* y) = (\varphi x, y) = (0, y) = 0,$$

т.е. $\varphi^* y \perp x$.

Основное свойство сопряженного оператора.

Если некоторое подпространство $\mathcal{H}$ инвариантно относительно оператора φ, то ортогональное дополнение $\mathcal{H}^\perp$ этого подпространства инвариантно относительно сопряженного оператора φ^*.

Свойства собственных значений и собственных векторов сопряженного оператора.

1. Характеристические многочлены, а, следовательно, и собственные значения сопряженных операторов в вещественном евклидовом пространстве одинаковы. В комплексном пространстве собственные значения сопряженных операторов являются комплексно сопряженными числами.

2. Каждый собственный вектор сопряженного оператора φ^* ортогонален ко всем собственным векторам оператора φ, принадлежащим другим собственным значениям.

Нормальный оператор

Линейный оператор $\varphi : \mathcal{E}_m \to \mathcal{E}_n(\mathcal{U}_m \to \mathcal{U}_n)$ называют *нормальным*, если он перестановочен со своим сопряженным оператором φ^*, т.е. если

$$\varphi^*\varphi = \varphi\varphi^*. \tag{26}$$

Квадратная матрица A называется *нормальной матрицей*, если $A^*A = AA^*$.

Из определения и связи матриц операторов φ и φ^*, рассмотренных в пункте 2.1, следует

Теорема 27. *Оператор нормален тогда и только тогда, когда в любом ортонормированном базисе его матрица нормальна.*

Теорема 28. *Собственный вектор нормального оператора, отвечающий собственному значению λ, является собственным вектором сопряженного оператора, отвечающим собственному значению $\bar{\lambda}$.*

Д о к а з а т е л ь с т в о. Легко проверить, что если φ – нормальный оператор, то $\varphi - \lambda\varepsilon$ также нормален.

Пусть теперь x – собственный вектор оператора φ, отвечающий собственному значению λ, тогда $(\varphi - \lambda\varepsilon)x = \theta$ и $((\varphi - \lambda\varepsilon)x, (\varphi - \lambda\varepsilon)x) = 0$.

Согласно определению сопряженного оператора можем записать, что

$$(x, (\varphi - \lambda\varepsilon)^*(\varphi - \lambda\varepsilon)x) = 0$$

или, с учетом нормальности оператора $\varphi - \lambda\varepsilon$, что

$$(x, (\varphi - \lambda\varepsilon)(\varphi - \lambda\varepsilon)^* x) = 0,$$

т.е.

$$(\varphi - \lambda\varepsilon)^* x, ((\varphi - \lambda\varepsilon)^* x) = 0$$

и $(\varphi - \lambda\varepsilon)^* x = \theta$.

Отсюда в силу свойств сопряженного оператора следует, что

$$(\varphi^* - \bar{\lambda}\varepsilon) = \theta,$$

т.е. $\varphi^* x = \bar{\lambda} x$. $\square$

Следствие 1. Если φ – нормальный оператор, то $\ker\varphi = \ker\varphi^*$, так как нетривиальные векторы ядра являются собственными векторами, отвечающими нулевому собственному значению.

Следствие 2. Если φ – нормальный оператор, то $\ker\varphi = \mathrm{im}^\perp\varphi$. Это следует из $\ker\varphi = \mathrm{im}^\perp\varphi^*$ и предыдущего следствия.

Теорема 29. *Собственные векторы нормального оператора, отвечающие различным собственным значениям, попарно ортогональны.*

Самосопряженный оператор

Линейный оператор $\varphi : \mathcal{E}_n \to \mathcal{E}_n(\mathcal{U}_n \to \mathcal{U}_n)$ называется *самосопряженным*, если для $\forall x, y \in \mathcal{E}_n(\mathcal{U}_n)$ выполняется равенство

$$(\varphi x, y) = (x, \varphi y),$$

т.е. $\varphi = \varphi^*$. Самосопряженный оператор в унитарном пространстве называют *эрмитовым*, а в евклидовом пространстве – *симметрическим.*

Примеры самосопряженного оператора.

1. Тождественный: $(\varepsilon x, y) = (x, y) = (x, \varepsilon y)$.
2. Нулевой: $(\theta x, y) = (0, y) = 0 = (x, 0) = (x, \theta y)$.

Квадратная матрица называется *самосопряженной*, если $A = A^*$.
Из определения вытекает, что самосопряженный оператор нормален.

Теорема 35. *Оператор самосопряженный тогда и только тогда, когда в любом ортонормированном базисе он имеет самосопряженную матрицу.*

Теорема 36. *Если подпространство $\mathcal{H}$ инвариантно относительно самосопряженного оператора φ, то ортогональное дополнение $\mathcal{H}^\perp$ этого подпространства также инвариантно относительно оператора φ.*

Теорема 37 (спектральная характеристика самосопряженного оператора). *Нормальный оператор в унитарном (евклидовом) пространстве самосопряжен тогда и только тогда, когда все корни его характеристического многочлена вещественны.*

Д о к а з а т е л ь с т в о. Необходимость. В унитарном пространстве это утверждение означает, что все собственные значения эрмитова оператора вещественны, и вытекает из равенств $\varphi x = \lambda x$ и, с учетом теоремы 28, $\varphi x = \bar{\lambda} x$. Докажем утверждение для евклидова пространства . Пусть e - ортонормированный базис , тогда $[\varphi]_e$ - самосопряженная (вещественная) матрица. Рассмотрим произвольное унитарное пространство $\mathcal{U}$ той же размерности, что и пространство $\mathcal{E}$, и в нем произвольный ортонормированный базис f. Тогда матрице $[\varphi]_e$ отвечает самосопряженный оператор $\psi \in L(\mathcal{U}_n, \mathcal{U}_n)$, для которого матрица $[\varphi]_e$ является матрицей в базисе $f : [\varphi]_e = [\psi]_f$. Следовательно, характеристические многочлены операторов φ и ψ совпадают и по доказанному выше (применительно к оператору ψ) все корни характеристического многочлена оператора φ вещественны.

Достаточность. Пусть φ - нормальный оператор и все корни его характеристического многочлена вещественны. Тогда как в евклидовом, так и в унитарном пространстве существует ортонормированный базис $e_1, e_2, \ldots, e_n$

из собственных векторов оператора φ. Если $x = \sum\limits_{i=1}^{n} x_i e_i$ – любой вектор пространства, то $\varphi x = \sum\limits_{i=1}^{n} x_i \lambda_i e_i$ и $\varphi^* x = \sum\limits_{i=1}^{n} x_i \bar{\lambda}_i e_i = \sum\limits_{i=1}^{n} x_i \lambda_i e_i$, так как $\lambda_i \in \mathbb{R}$. Следовательно, $\varphi x = \varphi^* x,\ \forall x \in \mathcal{E}(\mathcal{U})$, откуда следует, что $\varphi = \varphi^*$. $\square$

Теорема 38. *Собственные векторы самосопряженного оператора, отвечающие различным собственным значениям, ортогональны.*

Основным свойством самосопряженного оператора является то, что в евклидовом пространстве существует ортонормированный базис, состоящий из собственных векторов этого оператора. Это означает, что самосопряженный оператор является оператором простой структуры, а матрица $P_{e \to e'}$ приводит матрицу A самосопряженного оператора к диагональному виду, т.е. удовлетворяет соотношению (21)

$$\Lambda = P_{e \to e'}^{-1} A P_{e \to e'}.$$

Правило построения такой матрицы остается таким же, как и в случае любых операторов простой структуры с той лишь разницей, что базис из собственных векторов матрицы A здесь еще и ортонормируют.

2.3 Ортогональный (унитарный) оператор

Линейный оператор $\varphi : \mathcal{E}_n \to \mathcal{E}_n(\mathcal{U}_n \to \mathcal{U}_n)$ называется *ортогональным (унитарным)*, если он сохраняет скалярное произведение в $\mathcal{E}_n(\mathcal{U}_n)$, т.е. для $\forall x, y \in \mathcal{E}_n(\mathcal{U}_n)$ выполняется равенство

$$(\varphi x, \varphi y) = (x, y).$$

Полагая в этом равенстве $x = y$, получаем $|\varphi x|^2 = |x|^2$. Это означает, что ортогональный (унитарный) оператор сохраняет длины векторов.

Теорема 30. *Ортогональный (унитарный) оператор φ переводит любой ортонормированный базис $\mathcal{E}_n(\mathcal{U}_n)$ в ортонормированный базис.*

Доказательство. Пусть $e = (e_1, e_2, \dots, e_n)$ – произвольный ортонормированный базис в $\mathcal{E}_n(\mathcal{U}_n)$. В силу ортогональности оператора φ имеем

$$(\varphi e_i, \varphi e_j) = (e_i, e_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

Видим, что различные векторы φe_i и φe_j ортогональны, а длина каждого из них равна единице. Поэтому система векторов $\varphi e = (\varphi e_1, \varphi e_2, \dots, \varphi e_n)$ состоит из ненулевых векторов и ортогональна. Любая ортогональная система ненулевых векторов линейно независима. Количество векторов в линейно независимой системе φe равно размерности пространства $\mathcal{E}_n(\mathcal{U}_n)$, т.е. $\dim \mathcal{E}_n = n \Rightarrow$ эта система является базисом, притом ортонормированным. $\Box$

Теорема 31. *Если линейный оператор $\varphi : \mathcal{E}_n \to \mathcal{E}_n(\mathcal{U}_n \to \mathcal{U}_n)$ переводит какой-либо ортонормированный базис $e = (e_1, e_2, \dots, e_n)$ в ортонормированный базис $\varphi e = (\varphi e_1, \varphi e_2, \dots, \varphi e_n)$*, то этот оператор ортогональный (унитарный).

Теорема 32. *Оператор ортогонален (унитарен) тогда и только тогда, когда в любом ортонормированном базисе он имеет ортогональную (унитарную) матрицу.*

Теорема 33. *Собственные значения ортогонального (унитарного) оператора по абсолютной величине равны единице.*

Доказательство. Докажем для унитарного оператора. По определению можем записать $(\varphi x, \varphi x) = (x, x)$. Пусть x – собственный вектор оператора φ и λ – отвечающее ему собственное значение, $\varphi x = \lambda x$. Тогда $(\lambda x, \lambda x) = |\lambda|^2 (x, x)$;

$|\lambda|^2 (x, x) = (x, x)$;

$|\lambda|^2 = 1$. $\Box$

Теорема 34. *Собственные векторы ортогонального оператора, отвечающие различным собственным значениям, ортогональны.*

Элементы тензорной алгебры

В различных приложениях наряду со *скалярными* и *векторными* величинами активно используются и тензорные величины. Понятие тензора можно вводить по-разному. Согласно одному из подходов, говорят, что в линейном пространстве задан тензор, если каждому базису в соответствие поставлена упорядоченная система чисел (компонент тензора) и преобразование этой системы при переходе из одного базиса в другой подчиняется определенному закону. Компоненты тензора нумеруются, как правило, несколькими индексами, которые ставятся не только внизу, но и вверху буквенного обозначения. В рамках тензорного исчисления разрабатываются приемы и правила преобразований компонент тензоров при операциях над ними.

Сопряженное пространство

Определение 1. Отображение $f : \mathcal{L} \to \mathbb{R}$, которое определено на *линейном пространстве* $\mathcal{L}$ и принимает действительные значения, называют ***линейной функцией*** (также ***линейной формой***, ***линейным функционалом***), если оно удовлетворяет двум условиям:

а) $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y}), \boldsymbol{x}, \boldsymbol{y} \in \mathcal{L}$;

б) $f(\lambda\boldsymbol{x}) = \lambda f(\boldsymbol{x}), \boldsymbol{x} \in \mathcal{L}, \lambda \in \mathbb{R}$.

Сравнив данное определение с определением 4.1 *линейного оператора*, увидим много общего. Если рассматривать множество действительных чисел как одномерное линейное пространство, то можно сказать, что линейная функция — это линейный оператор, пространство образов которого одномерно.

Выберем в линейном пространстве $\mathcal{L}$ некоторый *базис* $\boldsymbol{e} = (\boldsymbol{e}_1 \ \ldots \ \boldsymbol{e}_n)$. Тогда для любого *вектора* $\boldsymbol{x} \in \mathcal{L}$ с *координатами* $x = (x_1 \ \ldots \ x_n)^{\mathrm{T}}$

$$f(\boldsymbol{x}) = f(x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n) = x_1 f(\boldsymbol{e}_1) + \ldots + x_n f(\boldsymbol{e}_n) = \\ = a_1x_1 + \ldots + a_nx_n = ax,$$

где $a = (a_1 \ \ldots \ a_n)$, $a_i = f(\boldsymbol{e}_i)$, $i = \overline{1,n}$. Поэтому линейная функция однозначно определяется своими значениями на базисных векторах. Наоборот, если функция $f(\boldsymbol{x})$ через координаты x вектора $\boldsymbol{x}$ выражается в виде $f(\boldsymbol{x}) = ax$, то эта функция линейная, а строка a составлена из значений этой функции на базисных векторах. Таким образом, между множеством линейных форм, заданных на линейном пространстве $\mathcal{L}$, и строками длины n установлено взаимно однозначное соответствие.

Линейные формы можно складывать и умножать на действительные числа согласно правилам:

$$(f+g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}), \qquad (\lambda f)(\boldsymbol{x}) = \lambda f(\boldsymbol{x}).$$

Введенные таким образом операции превращают множество линейных форм в пространстве $\mathcal{L}$ в линейное пространство. Это линейное пространство называют ***сопряженным пространством*** по отношению к линейному пространству $\mathcal{L}$ и обозначают $\mathcal{L}^*$.

Опираясь на базис $\boldsymbol{e}$, выбранный в пространстве $\mathcal{L}$, построим базис в сопряженном пространстве $\mathcal{L}^*$. Для каждого вектора $\boldsymbol{e}_i$ из базиса $\boldsymbol{e}$ рассмотрим линейную форму f^i, для которой $f^i(\boldsymbol{e}_i) = 1$ и $f^i(\boldsymbol{e}_j) = 0$ для всех векторов $\boldsymbol{e}_j$, кроме $\boldsymbol{e}_i$. Мы получим систему линейных форм $f^1, \ldots, f^n \in \mathcal{L}^*$. Покажем, что это *линейно независимая система.* Пусть некоторая *линейная комбинация* этих форм равна нулевой линейной форме $f = \alpha_1 f^1 + \ldots + \alpha_n f^n = 0$. Форма f на всех базисных векторах принимает нулевые значения. Но

$$\begin{aligned} f(\boldsymbol{e}_i) = (\alpha_1 f^1 + \ldots + \alpha_n f^n)(\boldsymbol{e}_i) = \\ \alpha_1 f^1(\boldsymbol{e}_i) + \ldots + \alpha_i f^i(\boldsymbol{e}_i) + \ldots + \alpha_n f^n(\boldsymbol{e}_i) = \\ \alpha_1 \cdot 0 + \ldots + \alpha_i \cdot 1 + \ldots + \alpha_n \cdot 0 = \alpha_i, \quad i = \overline{1, n}. \end{aligned}$$

Нулевые значения f на базисных векторах эквивалентны равенствам $\alpha_i = 0$, $i = \overline{1, n}$, и поэтому система линейных форм $f^1, \ldots, f^n$ линейно независима.

Система линейных форм $f^1, \ldots, f^n$ является базисом в сопряженном пространстве. Действительно, так как это линейно независимая система линейных форм, то достаточно доказать, что любая линейная форма из $\mathcal{L}^*$ является их линейной комбинацией. Выберем произвольную линейную форму f из $\mathcal{L}^*$ и пусть $a_1, \ldots, a_n$ — значения формы f на базисных векторах. Эти значения однозначно определяют линейную форму. Но линейная комбинация $f' = a_1 f^1 + \ldots + a_n f^n$ также является линейной формой, которая на базисных векторах принимает те же значения $a_1, \ldots, a_n$. Значит, эти две линейные формы совпадают, и мы получаем равенство $f = f' = a_1 f^1 + \ldots + a_n f^n$, т.е. разложение произвольно выбранной линейной формы по системе форм $f^1, \ldots, f^n$.

Приведенное рассуждение показывает, что сопряженное пространство $\mathcal{L}^*$ имеет ту же *размерность*, что и $\mathcal{L}$. Построенный нами базис $f^1, \ldots, f^n$ зависит от выбора базиса $\boldsymbol{e}$ в пространстве $\mathcal{L}$.

Определение 2. *Базисы* $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ и $f^1, \ldots, f^n$ *линейного пространства* $\mathcal{L}$ и сопряженного пространства $\mathcal{L}^*$ называют ***биортогональными***, или ***взаимными***, если

$$f^i(\boldsymbol{e}_j) = \delta^i_j = \begin{cases} 0, & i \neq j; \\ 1, & i = j. \end{cases}$$

Если базисы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ и $f^1, \ldots, f^n$ взаимны, то координатами произвольной формы f в базисе $f^1, \ldots, f^n$ являются значения этой формы на векторах взаимного базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$. При совместном рассмотрении линейного пространства $\mathcal{L}$ и сопряженного пространства $\mathcal{L}^*$ элементы каждого из этих пространств называют векторами, но элементы сопряженного пространства $\mathcal{L}^*$ именуют ***ковариантными векторами*** (***ковекторами***), а элементы из линейного пространства $\mathcal{L}$ — ***контравариантными векторами*** (или просто векторами). Координаты тех и других определяются преимущественно во взаимных базисах, при этом у координат контравариантных векторов индекс ставится вверху, а у ковариантных — внизу.

На запись $f(x)$ можно смотреть двояко. Зафиксировав форму f, мы варьируем вектор $\boldsymbol{x}$, получая всевозможные значения линейной формы. Но если мы зафиксируем вектор $\boldsymbol{x}$ и будем варьировать линейную форму f, то получим функцию, определенную на сопряженном пространстве $\mathcal{L}^*$. Нетрудно убедиться, что эта функция линейная, так как, согласно определению суммы линейных форм и произведения линейной формы на число,

$$(f+g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}), \qquad (\lambda f)(\boldsymbol{x}) = \lambda f(\boldsymbol{x}).$$

Итак, каждому вектору $\boldsymbol{x} \in \mathcal{L}$ соответствует линейная форма на сопряженном пространстве $\mathcal{L}$, или элемент ***двойного сопряженного пространства*** $(\mathcal{L}^*)^* = \mathcal{L}^{**}$. Мы получаем отображение $\varphi : \mathcal{L} \to \mathcal{L}^{**}$. Несложно убедиться, что это *отображение линейно* и что оно *инъективно*. Из инъективности следует, что $\dim \operatorname{im} \varphi = \dim \mathcal{L} = n$. Но сопряженное пространство $\mathcal{L}^*$ имеет ту же размерность, что и $\mathcal{L}$, а $\dim \mathcal{L}^{**} = \dim \mathcal{L}^* = \dim \mathcal{L}$. Таким образом, размерность *линейного подпространства* $\operatorname{im} \varphi$ в $\mathcal{L}^{**}$ совпадает с размерностью всего двойного сопряженного пространства. Значит, $\operatorname{im} \varphi = \mathcal{L}^{**}$ и отображение φ является *изоморфизмом*. Обратим внимание, что этот изоморфизм не связан с выбором какого-либо базиса. Поэтому естественно отождествить линейные формы, заданные на $\mathcal{L}^*$, с элементами пространства $\mathcal{L}$. Это означает, что двойное сопряженное пространство совпадает с исходным линейным пространством: $\mathcal{L}^{**} = \mathcal{L}$. Если $\mathcal{L}^*$ является сопряженным к $\mathcal{L}$, то и $\mathcal{L}$ является сопряженным к $\mathcal{L}^*$.

Взаимность линейного пространства и сопряженного к нему пространства указывает на симметричность связи между векторами и ковекторами. Поэтому вместо записи $f(x)$ более удобно использовать другую форму записи, симметричную: (f, x). Линейные формы мы также будем теперь обозначать полужирным курсивом: $(\boldsymbol{f}, x)$. Принятое обозначение похоже на обозначение *скалярного произведения*, но в отличие от последнего аргументы в новом обозначении берутся из разных пространств. Саму запись $(\boldsymbol{f}, x)$ можно рассматривать как запись отображения, определенного на множестве $\mathcal{L}^* \times \mathcal{L}$, которое паре из ковектора и вектора ставит в соответствие действительное число. При этом указанное отображение линейно по каждому из аргументов.

Теорема 1. Пусть $\boldsymbol{b}$ и $\boldsymbol{c}$ — два базиса *n-мерного линейного пространства* $\mathcal{L}$, U — матрица перехода из $\boldsymbol{b}$ в $\boldsymbol{c}$. Базисы $\boldsymbol{b}^*$ и $\boldsymbol{c}^*$ сопряженного простран-

ства $\mathcal{L}^*$, взаимные с базисами $\boldsymbol{b}$ и $\boldsymbol{c}$ соответственно, связаны между собой отношениями
$$\boldsymbol{c}^* = \boldsymbol{b}^* \left(U^{\mathrm{T}}\right)^{-1}, \qquad \boldsymbol{b}^* = \boldsymbol{c}^* U^{\mathrm{T}}.$$

◀ Координатами $f^c = (f_1^c \ \dots \ f_n^c)$ линейной формы $\boldsymbol{f}$ в базисе $\boldsymbol{c}^*$ являются значения этой формы на векторах базиса $\boldsymbol{c} = (\boldsymbol{c}_1 \ \dots \ \boldsymbol{c}_n)$. Выясним, как связаны координаты формы $\boldsymbol{f}$ в двух базисах $\boldsymbol{c}^*$ и $\boldsymbol{b}^*$.

Базисы $\boldsymbol{b}$ и $\boldsymbol{c}$ связаны между собой при помощи матрицы перехода матричным соотношением $\boldsymbol{c} = \boldsymbol{b}U$ (см. **1.8**). Это соотношение представляет собой равенство строк длины n, составленных из векторов. Из равенства строк векторов следует равенство строк значений линейной формы $\boldsymbol{f}$ на этих векторах:
$$\big((\boldsymbol{f}, \boldsymbol{c}_1) \ \dots \ (\boldsymbol{f}, \boldsymbol{c}_n)\big) = \big((\boldsymbol{f}, \boldsymbol{b}_1) \ \dots \ (\boldsymbol{f}, \boldsymbol{b}_n)\big)\, U,$$
или
$$f^c = f^b U,$$
где f^b и f^c — обозначения строк координат формы $\boldsymbol{f}$ в базисах $\boldsymbol{b}^*$ и $\boldsymbol{c}^*$ соответственно. Транспонировав это равенство, мы получим принятую форму связи координат элементов линейного пространства, в которой координаты записываются по столбцам:
$$(f^c)^{\mathrm{T}} = U^{\mathrm{T}} \left(f^b\right)^{\mathrm{T}}.$$
Это соотношение означает, что матрица U^{T} является матрицей перехода из базиса $\boldsymbol{c}^*$, играющего в формуле роль старого, в базис $\boldsymbol{b}^*$, играющий роль нового. Следовательно, $\boldsymbol{b}^* = \boldsymbol{c}^* U^{\mathrm{T}}$, откуда умножением на матрицу $\left(U^{\mathrm{T}}\right)^{-1}$ получаем $\boldsymbol{c}^* = \boldsymbol{b}^* \left(U^{\mathrm{T}}\right)^{-1}$. ▶

Если линейное пространство $\mathcal{L}$ *евклидово*, то скалярное произведение порождает изоморфизм между $\mathcal{L}$ и $\mathcal{L}^*$, не зависящий от базиса, который позволяет отождествить евклидово пространство с его сопряженным. Действительно, для любого вектора $\boldsymbol{a} \in \mathcal{L}$ отображение $\boldsymbol{x} \to (\boldsymbol{a}, \boldsymbol{x})$ представляет собой линейную форму в $\mathcal{L}$, так как скалярное произведение линейно по второму из своих аргументов. Возникает отображение ψ, которое вектору $\boldsymbol{a} \in \mathcal{L}$ ставит в соответствие линейную форму $f_a(\boldsymbol{x}) = (\boldsymbol{a}, \boldsymbol{x})$. Это отображение линейно в силу свойств скалярного произведения и инъективно. Инъективность следует из того, что если $(\boldsymbol{a}, \boldsymbol{x}) = 0$ для любого $\boldsymbol{x} \in \mathcal{L}$, то и $(\boldsymbol{a}, \boldsymbol{a}) = 0$, т.е. $\boldsymbol{a} = 0$. Так как линейные пространства $\mathcal{L}$ и $\mathcal{L}^*$ конечномерны и имеют одинаковые размерности, отображение ψ биективно и реализует изоморфизм этих пространств. Итак, для евклидова пространства $\mathcal{L}^* = \mathcal{L}$. В этом смысле евклидово пространство есть „самосопряженное“ пространство.

Полилинейные формы

Пусть $\mathcal{L}$ — *n-мерное линейное пространство* и $\mathcal{L}^*$ — *сопряженное* к нему *пространство*. Рассмотрим функцию $\varphi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p; \boldsymbol{f}^1, \ldots, \boldsymbol{f}^q)$, аргументами которой являются *p векторов* $\boldsymbol{x}_i \in \mathcal{L}$ и *q ковекторов* $\boldsymbol{f}^j \in \mathcal{L}^*$.

Определение 3. Функцию φ от p векторов и q ковекторов называют ***полилинейной формой***, если она линейна по каждому отдельно взятому аргументу. Пару чисел (p, q) называют ***типом полилинейной формы***.

Пример 1. Простейшие полилинейные формы — это *линейные функции*, зависящие от одного аргумента. Линейные функции на $\mathcal{L}$ представляют собой ковекторы, т.е. элементы сопряженного пространства $\mathcal{L}^*$. Линейные функции на $\mathcal{L}^*$ отождествляются с векторами. Таким образом, полилинейная форма типа $(1, 0)$ — это ковектор, а полилинейная форма типа $(0, 1)$ — это вектор.

Пример 2. Полилинейная форма типа $(2, 0)$ — это *билинейная форма*, определенная на линейном пространстве $\mathcal{L}$. Аналогично полилинейная форма типа $(0, 2)$ представляет собой билинейную форму на сопряженном пространстве $\mathcal{L}^*$.

Пример 3. Полилинейную форму $\varphi(\boldsymbol{x}; \boldsymbol{f})$ типа $(1, 1)$ можно ассоциировать с *линейным оператором*, действующим в линейном пространстве $\mathcal{L}$. Действительно, зафиксировав первый аргумент, мы получим линейную функцию на сопряженном пространстве $\mathcal{L}^*$, т.е. вектор. Таким образом, каждому вектору $\boldsymbol{x} \in \mathcal{L}$ поставлен в соответствие вектор, представленный в виде линейной формы на $\mathcal{L}^*$. Мы получаем отображение пространства $\mathcal{L}$ в себя. Покажем,что это отображение линейно.

Если вектору $\boldsymbol{x}$ соответствует линейная форма $\varphi(\boldsymbol{x}; \cdot)$ на $\mathcal{L}^*$ (точка обозначает меняющийся аргумент), а вектору $\boldsymbol{y}$ соответствует линейная форма $\varphi(\boldsymbol{y}; \cdot)$, то *сумме* этих *векторов* соответствует линейная форма $\varphi(\boldsymbol{x} + \boldsymbol{y}; \cdot)$, равная сумме форм:

$$\varphi(\boldsymbol{x} + \boldsymbol{y}; \cdot) = \varphi(\boldsymbol{x}; \cdot) + \varphi(\boldsymbol{y}; \cdot),$$

что следует из линейности φ по первому аргументу. Аналогично вектору $\lambda\boldsymbol{x}$ соответствует форма $\varphi(\lambda\boldsymbol{x}; \cdot)$, равная $\lambda\varphi(\boldsymbol{x}; \cdot)$.

Итак, любой полилинейной форме φ типа $(1, 1)$ соответствует линейный оператор, действующий в $\mathcal{L}$. Можно показать, что это соответствие *биективное*, и мы сможем отождествить полилинейные формы типа $(1, 1)$ с линейными операторами.

Соответствие между полилинейными формами типа $(1, 1)$ и линейными операторами использует ранее построенный *изоморфизм* между линейными пространствами $\mathcal{L}$ и $\mathcal{L}^{**}$. Обратное соответствие более простое. Каждому линейному оператору $\boldsymbol{A}$ можно поставить в соответствие полилинейную форму

$\varphi_A(\boldsymbol{x};\boldsymbol{f}) = (\boldsymbol{f};\boldsymbol{Ax})$ типа $(1,1)$. При фиксированном векторе $\boldsymbol{x}$ мы получаем линейную форму на сопряженном пространстве, причем эта форма отождествляется с вектором $\boldsymbol{Ax}$. Значит, это действительно то же соответствие, что и рассмотренное выше.

Полилинейные формы можно складывать по обычным правилам сложения функций: для каждой комбинации значений $p+q$ аргументов складываются значения функций. Полилинейные формы можно также умножать на действительные числа.

Теорема 2. Множество $\mathcal{P}_{p,q}$ полилинейных форм типа (p,q) в n-мерном линейном пространстве $\mathcal{L}$ является линейным пространством относительно обычных операций сложения функций и умножения функции на число.

◀ Операции сложения функций и умножения функции на число удовлетворяют всем *аксиомам линейного пространства.* Поэтому нам нужно лишь показать, что в результате сложения двух форм одного типа или умножения полилинейной формы на действительное число получается полилинейная форма того же типа.

Рассмотрим полилинейные формы $\varphi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q)$ и $\psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q)$. Их сумма представляет собой функцию $\chi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q)$, которая определяется равенством

$$\chi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q) = \\ = \varphi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q) + \psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q).$$

Проверим линейность этой функции, например, по первому аргументу, используя многоточия для обозначения остальных аргументов полилинейных форм:

$$\chi(\boldsymbol{x}_1+\boldsymbol{x}'_1,\ldots) = \varphi(\boldsymbol{x}_1+\boldsymbol{x}'_1,\ldots) + \psi(\boldsymbol{x}_1+\boldsymbol{x}'_1,\ldots) = \\ = \bigl(\varphi(\boldsymbol{x}_1,\ldots)+\varphi(\boldsymbol{x}'_1,\ldots)\bigr) + \bigl(\psi(\boldsymbol{x}_1,\ldots)+\psi(\boldsymbol{x}'_1,\ldots)\bigr) = \\ = \bigl(\varphi(\boldsymbol{x}_1,\ldots)+\psi(\boldsymbol{x}_1,\ldots)\bigr) + \bigl(\varphi(\boldsymbol{x}'_1,\ldots)+\psi(\boldsymbol{x}'_1,\ldots)\bigr) = \\ = \chi(\boldsymbol{x}_1,\ldots)+\chi(\boldsymbol{x}'_1,\ldots),$$

$$\chi(\lambda\boldsymbol{x}_1,\ldots) = \varphi(\lambda\boldsymbol{x}_1,\ldots) + \psi(\lambda\boldsymbol{x}_1,\ldots) = \\ = \lambda\varphi(\boldsymbol{x}_1,\ldots) + \lambda\psi(\boldsymbol{x}_1,\ldots) = \\ = \lambda\left(\varphi(\boldsymbol{x}_1,\ldots)+\psi(\boldsymbol{x}_1,\ldots)\right) = \lambda\chi(\boldsymbol{x}_1,\ldots).$$

Линейность функции по остальным аргументам проверяется точно так же.

Рассмотрим теперь функцию ξ, определенную через полилинейную форму φ равенством $\xi(\ldots) = \lambda\varphi(\ldots)$. Проверим ее линейность по первому аргументу:

$$\xi(\boldsymbol{x}_1+\boldsymbol{x}'_1,\ldots)=\lambda\varphi(\boldsymbol{x}_1+\boldsymbol{x}'_1,\ldots)=\\ \lambda\left(\varphi(\boldsymbol{x}_1,\ldots)+\varphi(\boldsymbol{x}'_1,\ldots)\right)=\\ \lambda\varphi(\boldsymbol{x}_1,\ldots)+\lambda\varphi(\boldsymbol{x}'_1,\ldots)=\xi(\boldsymbol{x}_1,\ldots)+\xi(\boldsymbol{x}'_1,\ldots),$$

$$\xi(\alpha\boldsymbol{x}_1,\ldots)\quad=\quad\lambda\varphi(\alpha\boldsymbol{x}_1,\ldots)\quad=\quad\lambda\alpha\varphi(\boldsymbol{x}_1,\ldots)\quad=\quad\alpha\xi(mbfx_1,\ldots).$$

Линейность функции ξ по остальным аргументам проверяется аналогично.▶

Рассмотрим в линейном пространстве $\mathcal{L}$ некоторый *базис* $\boldsymbol{e}=(\boldsymbol{e}_1,\ \ldots,\ \boldsymbol{e}_n)$. Используя свойства линейности полилинейной формы φ типа (p,q) по каждому аргументу, мы можем выразить ее значение на фиксированном наборе значений аргументов через ее же значения на базисных векторах. Пусть выбраны произвольные векторы $\boldsymbol{x}_i=x_i^1\boldsymbol{e}_1+\ldots+x_i^n\boldsymbol{e}_n,\ i=\overline{1,p}$ и произвольные ковекторы $\boldsymbol{f}^i=f_1^i\boldsymbol{e}^1+\ldots+f_n^i\boldsymbol{e}^n,\ i=\overline{1,q}$, разложенные по *взаимному базису* $\boldsymbol{e}^*=(\boldsymbol{e}^1\ \ldots\ \boldsymbol{e}^n)$. Тогда

$$\varphi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q)=\\ \varphi\left(\sum_{i_1=1}^n x_1^{i_1}\boldsymbol{e}_{i_1},\ldots,\sum_{i_p=1}^n x_p^{i_p}\boldsymbol{e}_{i_p};\sum_{j_1=1}^n f_{j_1}^1\boldsymbol{e}^{j_1},\ldots,\sum_{j_q=1}^n f_{j_q}^q\boldsymbol{e}^{j_q}\right)=\\ \sum_{i_1=1}^n\cdots\sum_{i_p=1}^n\sum_{j_1=1}^n\cdots\sum_{j_q=1}^n x_1^{i_1}\ldots x_p^{i_p}f_{j_1}^1\ldots f_{j_q}^q\times\\ \times\varphi\left(\boldsymbol{e}_{i_1},\ldots,\boldsymbol{e}_{i_p};\boldsymbol{e}^{j_1},\ldots,\boldsymbol{e}^{j_q}\right).$$

Мы видим, что набор чисел $\varphi_{i_1\ldots i_p}^{j_1\ldots j_q}=\varphi\left(\boldsymbol{e}_{i_1},\ldots,\boldsymbol{e}_{i_p};\boldsymbol{e}^{j_1},\ldots,\boldsymbol{e}^{j_q}\right)$ однозначно задает полилинейную форму. Ясно, что, взяв произвольный набор чисел $\varphi_{i_1\ldots i_p}^{j_1\ldots j_q}$, мы сможем определить полилинейную форму, используя эти числа как коэффициенты разложения в базисе $\boldsymbol{e}$. Таким образом, подобно тому, как каждому вектору соответствует столбец его *координат*, каждой полилинейной форме соответствует набор чисел, упорядоченный с помощью $p+q$ индексов. Эти числа назовем ***координатами полилинейной формы***. Термин оправдан, так как в линейном пространстве $\mathcal{P}_{p,q}$ несложно указать базис, в котором эти числа действительно будут координатами данной полилинейной формы. Этот базис будет состоять из n^{p+q} полилинейных форм, и поэтому $\dim\mathcal{P}_{p,q}=n^{p+q}$.

Уже по приведенной выкладке видно, что большое количество знаков суммы загромождает вычисления. Поэтому в тензорном исчислении используют

правило суммирования по умолчанию, или ***правило индексов***. Индексы в выражениях тензорной алгебры ставят вверху и внизу. Если в выражении какой-либо верхний индекс и какой-либо нижний индекс обозначены одинаково, то подразумевается, что по этому индексу проводится суммирование в пределах от единицы до размерности линейного пространства. При этом знак Σ суммирования опускается. Например, формулу разложения полилинейной формы в базисе в соответствии с правилом индексов записывают так:

$$\varphi(x_1, \ldots, x_p; f^1, \ldots, f^q) = \varphi_{i_1 \ldots i_p}^{j_1 \ldots j_q} x_1^{i_1} \ldots x_p^{i_p} f_{j_1}^1 \ldots f_{j_q}^q.$$

Далее мы будем использовать это правило.

Итак, полилинейную форму в данном базисе можно представить набором ее координат. Выясним, как изменяются эти координаты при изменении базиса. Пусть $\boldsymbol{b} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ и $\boldsymbol{c} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n)$ — два базиса в n-мерном линейном пространстве $\mathcal{L}$, $\boldsymbol{b}^* = (\boldsymbol{b}^1, \ldots, \boldsymbol{b}^n)$ и $\boldsymbol{c}^* = (\boldsymbol{c}^1, \ldots, \boldsymbol{c}^n)$ — базисы в сопряженном пространстве $\mathcal{L}^*$, взаимные с $\boldsymbol{b}$ и $\boldsymbol{c}$. Обозначим через U матрицу перехода их базиса $\boldsymbol{b}$ в базис $\boldsymbol{c}$.

Теорема 3. Координаты $\widetilde{\varphi}_{i_1 \ldots i_p}^{j_1 \ldots j_q}$ полилинейной формы φ в базисе $\boldsymbol{c}$ связаны с координатами $\varphi_{i_1 \ldots i_p}^{j_1 \ldots j_q}$ этой же формы в базисе $\boldsymbol{b}$ соотношениями

$$\widetilde{\varphi}_{i_1 \ldots i_p}^{j_1 \ldots j_q} = v_{s_1}^{j_1} \ldots v_{s_q}^{j_q} \varphi_{r_1 \ldots r_p}^{s_1 \ldots s_q} u_{i_1}^{r_1} \ldots u_{i_p}^{r_p},$$

где $\left(u_j^i\right) = U$ — *матрица перехода* из базиса $\boldsymbol{b}$ в базис $\boldsymbol{c}$; $\left(v_j^i\right) = V = U^{-1}$ — матрица обратного перехода (верхний индекс соответствует номеру строки). ◀ Координата $\widetilde{\varphi}_{i_1 \ldots i_p}^{j_1 \ldots j_q}$, соответствующая фиксированному набору индексов $i_1, \ldots, i_p, j_1, \ldots, j_q$, представляет собой значение полилинейной формы:

$$\widetilde{\varphi}_{i_1 \ldots i_p}^{j_1 \ldots j_q} = \varphi\left(\boldsymbol{c}_{i_1}, \ldots, \boldsymbol{c}_{i_p}; \boldsymbol{c}^{j_1}, \ldots, \boldsymbol{c}^{j_q}\right).$$

Используя выражение векторов нового базиса через старый при помощи матрицы перехода

$$\boldsymbol{c}_j = \boldsymbol{b}_k u_j^k, \qquad \boldsymbol{c} = v_l^i \boldsymbol{b}^l,$$

находим

$$\begin{aligned} \widetilde{\varphi}_{i_1 \ldots i_p}^{j_1 \ldots j_q} = \varphi(\boldsymbol{b}_{r_1} u_{i_1}^{r_1}, \ldots, \boldsymbol{b}_{r_p} u_{i_p}^{r_p}; v_{s_1}^{j_1} \boldsymbol{b}^{s_1}, \ldots, v_{s_q}^{j_q} \boldsymbol{b}^{s_q}) = \\ = v_{s_1}^{j_1} \ldots v_{s_q}^{j_q} u_{i_1}^{r_1} \ldots u_{i_p}^{r_p} \varphi(\boldsymbol{b}_{r_1}, \ldots, \boldsymbol{b}_{r_p}; \boldsymbol{b}^{s_1}, \ldots, \boldsymbol{b}^{s_q}) = \\ = v_{s_1}^{j_1} \ldots v_{s_q}^{j_q} \varphi_{r_1 \ldots r_p}^{s_1 \ldots s_q} u_{i_1}^{r_1} \ldots u_{i_p}^{r_p}. \qquad \blacktriangleright \end{aligned}$$

Пример 4. Для линейного оператора $\boldsymbol{A}$ с матрицей $A = (a_j^i)$, который отождествляется с полилинейной формой типа $(1, 1)$ (см. пример 10.3), формула преобразования при переходе к новому базису, согласно теореме 10.3, имеет вид $\widetilde{a}_i^j = v_s^j a_r^s u_j^r$. Эта формула является координатной записью известной формулы $\widetilde{A} = U^{-1}AU$ преобразования матрицы линейного оператора (см. теорему 4.6).

Тензоры

Определение 4 Говорят, что в *n-мерном линейном пространстве* $\mathcal{L}$ задан **тензор типа** $(p,q)^*$, если каждому *базису* $\boldsymbol{b}$ в $\mathcal{L}$ сопоставлена упорядоченная система чисел $a^{j_1\ldots j_q}_{i_1\ldots i_p}(\boldsymbol{b})$, называемых ***компонентами тензора***, причем системы чисел, соответствующие разным базисам $\boldsymbol{b}$ и $\boldsymbol{c}$, связаны между собой соотношениями

$$a^{j_1\ldots j_q}_{i_1\ldots i_p}(\boldsymbol{c}) = v^{j_1}_{s_1}\ldots v^{j_q}_{s_q} a^{s_1\ldots s_q}_{r_1\ldots r_p}(\boldsymbol{b}) u^{r_1}_{i_1}\ldots u^{r_p}_{i_p}, \tag{1}$$

где $U=(u^i_j)$ — *матрица перехода* из базиса $\boldsymbol{b}$ в базис $\boldsymbol{c}$; $V=(v^i_j)$ — обратная к U матрица (верхний индекс у u^i_j и v^i_j обозначает номер строки в матрице). Сумму $p+q$ называют ***валентностью тензора*** (или его ***рангом***).

Понятие тензора носит абстрактный характер: происхождение групп чисел, формирующих тензор, не играет роли. Пример тензора типа (p,q) дают координаты *полилинейной формы* типа (p,q). Действительно, закон преобразования полилинейных форм при замене базиса (см. теорему 3) и закон преобразования тензоров того же типа (см. определение 4) совпадают. Верно и обратное: любой тензор можно интерпретировать как совокупность координат некоторой полилинейной формы. Если тензору типа (p,q) с компонентами $a^{j_1\ldots j_q}_{i_1\ldots i_p}(\boldsymbol{b})$ в базисе $\boldsymbol{b}$ сопоставить полилинейную форму с координатами $a^{j_1\ldots j_q}_{i_1\ldots i_p}(\boldsymbol{b})$ в том же базисе, то в силу совпадения формул преобразования и в любом другом базисе компоненты тензора будут совпадать с координатами полилинейной формы. Эти соображения показывают, что тензоры при необходимости можно трактовать как полилинейные формы и наоборот.

Исходя из определения 4 можно предположить, что $p>0$ и $q>0$. Но это необязательно. Например, если $p=0$, то в законе преобразования 10.1 не будет использоваться матрица U, а если $q=0$, то не будет использоваться V.

Определение 5. Тензор типа $(p,0)$ называют ***ковариантным***, а тензор типа $(0,q)$ — ***контравариантным***. Тензор типа (p,q) ***смешанный***, если $p>0, q>0$. Про такой тензор говорят, что он p раз ковариантный и q раз контравариантный.

Приведем простейшие примеры тензоров.

Пример 5. Допустима ситуация, когда $p=q=0$. Это означает, что объект описывается одним числом (индексов нет), причем это число не зависит от выбора базиса (в законе преобразования нет суммирования и он приобретает вид $a(\boldsymbol{c})=a(\boldsymbol{b})$). Такой объект представляет собой тензор типа $(0,0)$. Его также называют ***инвариантом***. Можно также сказать, что инвариант — это попросту *скалярная величина*.

Пример 6. Тензор образуют координаты тензора. Это вытекает из интерпретации вектора как полилинейной формы типа $(0,1)$ (см. пример 1). Проверим непосредственно, что координаты вектора представляют собой тензор типа $(0,1)$. Если U — матрица перехода из старого базиса в новый, а V — обратная к U, то столбец $\widetilde{x}$ новых координат вектора $\boldsymbol{x}$ связан со столбцом x старых координат соотношением $\widetilde{x} = Vx$, а в тензорной записи это выглядит следующим образом: $\widetilde{x}^j = v^j_s x^s$. Видим, что это преобразование координат совпадает с преобразованием (10.1) при $p = 0, q = 1$.

Рассуждая аналогичным образом, убеждаемся, что *ковектор* (линейная форма) представляет собой тензор типа $(1,0)$.

Пример 7. Линейный оператор $\boldsymbol{A}$, действующий в линейном пространстве $\mathcal{L}$, можно ассоциировать с полилинейной формой типа $(1,1)$ (см. пример 3). Это значит, что линейный оператор можно рассматривать как тензор типа $(1,1)$. Проверим это непосредственно. В данном базисе $\boldsymbol{b}$ линейный оператор описывается своей матрицей A. При переходе в новый базис $\boldsymbol{c}$ матрица A линейного оператора преобразуется в матрицу $\widetilde{A}$ согласно формуле $\widetilde{A} = U^{-1}AU = VAU$. Записав матрицы в тензорной форме (верхний индекс соответствует номеру строки в матрице, а нижний — номеру столбца), получим $\widetilde{a}^j_i = v^j_s a^s_r u^r_i$ (см. пример 4), т.е. формулу (10.1) при $p = 1, q = 1$. Значит, элементы матрицы линейного оператора при переходе в другой базис меняются как компоненты тензора типа $(1,1)$.

Пример 8. ***Символ Кронекера*** — это тензор δ^i_j типа $(1,1)$, который в любом базисе имеет значения $\delta^i_i = 1$ для любого i и $\delta^i_j = 0$, если $i \neq j$. Этот тензор соответствует *тождественному линейному оператору*, так как компоненты символа Кронекера соответствуют элементам единичной матрицы.

Пример 9. В *евклидовом пространстве* $\mathcal{E}$ заданное *скалярное произведение* представляет собой *билинейную форму*, т.е. тензор g_{ij} типа $(2,0)$. Этот тензор называют ***ковариантным метрическим тензором***. Роль такого тензора (иначе, скалярного произведения) может играть любая *симметрическая билинейная форма*, порождающая *положительно определенную квадратичную форму*.

Евклидово пространство $\mathcal{E}$ изоморфно своему сопряженному $\mathcal{E}^*$, причем изоморфизм, определяемый скалярным произведением, не связан с выбором базиса (см. **1**). Этот изоморфизм переносит скалярное произведение из $\mathcal{E}$ в $\mathcal{E}^*$, порождая тензор g^{ij} типа $(0,2)$. Этот тензор называют ***контравариантным метрическим тензором***.

Компонентами ковариантного метрического тензора в данном базисе $\boldsymbol{e}$ являются элементы *матрицы Грама* Γ, так как, согласно определению метрического тензора, $g_{ij} = (\boldsymbol{e}_i, \boldsymbol{e}_j)$. Матрицей контравариантного метрического тензора g^{ij} в $\boldsymbol{e}$ является матрица, обратная матрице Грама. Действительно, компоненты тензора g^{ij} — это элементы матрицы Грама Γ^* в базисе $\boldsymbol{f}$, взаимном с $\boldsymbol{e}$. Пусть U — матрица перехода из базиса $\boldsymbol{e}$ в базис $\boldsymbol{f}$. Тогда $\boldsymbol{f} = eU$ ($\boldsymbol{f}$ и $\boldsymbol{e}$, как обычно, представляются в виде строк). Матрица Грама Γ^* для базиса $\boldsymbol{f}$ записывается в виде $\Gamma^* = \boldsymbol{f}^{\mathrm{T}} f$. Поэтому

$$\Gamma^* = \boldsymbol{f}^{\mathrm{T}} f = \boldsymbol{f}^{\mathrm{T}} \boldsymbol{e} U = EU = U.$$

С другой стороны,

$$E = \boldsymbol{f}^{\mathrm{T}} \boldsymbol{e} = (\boldsymbol{e}U)^{\mathrm{T}} \boldsymbol{e} = U^{\mathrm{T}} \boldsymbol{e}^{\mathrm{T}} \boldsymbol{e} = U^{\mathrm{T}} \Gamma,$$

откуда $U = (\Gamma^{\mathrm{T}})^{-1} = \Gamma^{-1}$ и $\Gamma^* = U = \Gamma^{-1}$.

Примеры тензоров можно почерпнуть в механике и физике.

Операции с тензорами

Линейные операции. Мы видели, что множество *полилинейных форм* одного типа образует *линейное пространство* относительно обычных операций сложения функций и умножения функции на число. Каждой полилинейной форме соответствует *тензор*, и, наоборот, любой тензор можно реализовать как полилинейную форму. Значит, структуру линейного пространства с множества линейных форм можно перенести на тензоры.

Определение 6. ***Суммой*** двух ***тензоров*** $\boldsymbol{A} = a^{j_1\dots j_q}_{i_1\dots i_p}$ и $\boldsymbol{B} = b^{j_1\dots j_q}_{i_1\dots i_p}$ типа (p,q) называют тензор $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} = c^{j_1\dots j_q}_{i_1\dots i_p}$ того же типа с компонентами

$$c^{j_1\dots j_q}_{i_1\dots i_p} = a^{j_1\dots j_q}_{i_1\dots i_p} + b^{j_1\dots j_q}_{i_1\dots i_p}. \tag{2}$$

Нетрудно убедиться, что набор компонентов $c^{j_1\dots j_q}_{i_1\dots i_p}$, вычисляемых в каждом базисе по формуле (10.2), определяет тензор типа (p,q). Действительно, для двух базисов, старого $\boldsymbol{b}$ и нового $\boldsymbol{c}$, с матрицей перехода $U = (u^i_j)$ и матрицей обратного перехода $V = (v^i_j)$

$$\begin{aligned}
c^{j_1\dots j_q}_{i_1\dots i_p}(\boldsymbol{c}) &= a^{j_1\dots j_q}_{i_1\dots i_p}(\boldsymbol{c}) + b^{j_1\dots j_q}_{i_1\dots i_p}(\boldsymbol{c}) = \\
&= v^{j_1}_{r_1}\dots v^{j_q}_{r_q} a^{r_1\dots r_q}_{s_1\dots s_p}(\boldsymbol{b}) u^{s_1}_{i_1}\dots u^{s_p}_{i_p} + v^{j_1}_{r_1}\dots v^{j_q}_{r_q} b^{r_1\dots r_q}_{s_1\dots s_p}(\boldsymbol{b}) u^{s_1}_{i_1}\dots u^{s_p}_{i_p} = \\
&= v^{j_1}_{r_1}\dots v^{j_q}_{r_q}\left(a^{r_1\dots r_q}_{s_1\dots s_p}(\boldsymbol{b}) + b^{r_1\dots r_q}_{s_1\dots s_p}(\boldsymbol{b})\right) u^{s_1}_{i_1}\dots u^{s_p}_{i_p} = \\
&= v^{j_1}_{r_1}\dots v^{j_q}_{r_q} c^{r_1\dots r_q}_{s_1\dots s_p}(\boldsymbol{b}) u^{s_1}_{i_1}\dots u^{s_p}_{i_p}.
\end{aligned}$$

Видно, что набор компонентов $c^{j_1\dots j_q}_{i_1\dots i_p}$ преобразуется по тензорному закону, т.е. согласно формуле (2).

Введенная операция сложения тензоров одного типа согласуется с операциями сложения объектов, являющихся частными случаями тензоров. Для *векторов* тензорное сложение совпадает со сложением векторов, для *линейных* или *билинейных форм* тензорное сложение равносильно сложению функций. Наконец, тензорное сложение *линейных операторов* и их обычное сложение — одно и то же. Для полилинейных форм тензорное сложение означает сложение форм как функций.

Определение 7. ***Произведением тензора*** $\boldsymbol{A} = a^{j_1\dots j_q}_{i_1\dots i_p}$ ***на действительное число*** λ называют тензор $\lambda\boldsymbol{A}$ с компонентами $\lambda a^{j_1\dots j_q}_{i_1\dots i_p}$.

Так же, как и в случае сложения тензоров, убеждаемся, что в результате умножения каждой компоненты тензора на число λ мы получаем тензор того же типа. Умножение тензора на число в частных случаях сводится к умножению на число вектора, линейной или билинейной формы, линейного оператора. В интерпретации тензора как полилинейной формы умножение тензора на число равносильно умножению функции на число.

Теорема 4. Множество $\mathcal{T}_{p,q}$ всех тензоров типа (p,q) в *n-мерном линейном пространстве* $\mathcal{L}$ относительно операций сложения тензоров и умножения тензора на число является линейным пространством *размерности* n^{p+q}.

◀ Проверка *аксиом линейного пространства* не представляет сложности. Докажем утверждение о размерности пространства тензоров. Для каждого возможного набора индексов ковариантных $i_1, \ldots, i_p$ и контравариантных $j_1, \ldots, j_q$ рассмотрим тензор $\boldsymbol{T}^{i_1\ldots i_p}_{j_1\ldots j_q}$ с компонентами $(t^{i_1\ldots i_p}_{j_1\ldots j_q})^{l_1\ldots l_q}_{k_1\ldots k_p}$, причем равны нулю все компоненты, кроме одной, индексы которой совпадают с индексами тензора: $(t^{i_1\ldots i_p}_{j_1\ldots j_q})^{j_1\ldots j_q}_{i_1\ldots i_p} = 1$. Тогда любой тензор $\boldsymbol{A} = a^{j_1\ldots j_q}_{i_1\ldots i_p}$ представляется в виде линейной комбинации указанных векторов:

$$\boldsymbol{A} = a^{j_1\ldots j_q}_{i_1\ldots i_p} \boldsymbol{T}^{i_1\ldots i_p}_{j_1\ldots j_q}$$

Докажем, что набор тензоров $\boldsymbol{T}^{i_1\ldots i_p}_{j_1\ldots j_q}$ образует линейно независимую систему. Возьмем линейную комбинацию этих тензоров и приравняем нулевому тензору, у которого все компоненты равны нулю: $\alpha^{j_1\ldots j_q}_{i_1\ldots i_p} \boldsymbol{T}^{i_1\ldots i_p}_{j_1\ldots j_q} = \Theta$. В левой части равенства стоит тензор, компонентами которого являются коэффициенты $\alpha^{j_1\ldots j_q}_{i_1\ldots i_p}$ линейной комбинации. Так как этот тензор является нулевым, все его компоненты, они же коэффициенты линейной комбинации, равны нулю. Итак, из равенства нулю линейной комбинации следует равенство нулю ее коэффициентов. Значит, выбранная система тензоров линейно независима и является базисом в пространстве $\mathcal{T}_{p,q}$. Подсчитаем количество тензоров в построенном базисе. Для этого необходимо определить количество всевозможных комбинаций из $p+q$ индексов. Так как индексы меняются независимо друг от друга и каждый индекс может иметь n возможных значений, то суммарное количество индексных комбинаций равно n^{p+q}. Следовательно, и базис $\boldsymbol{T}^{i_1\ldots i_p}_{j_1\ldots j_q}$ содержит n^{n+q} элементов, что равно размерности пространства $\mathcal{T}_{p,q}$. ▶

Транспонирование. У полилинейной формы можно переставить какие-либо два аргумента одного типа (два вектора или два ковектора). В результате мы получим, вообще говоря, новую полилинейную форму. Например, при перестановке аргументов билинейной формы мы получаем новую билинейную форму, матрица которой является транспонированной к матрице исходной формы. Такая операция не меняет билинейную форму лишь в том случае, когда эта форма *симметрическая*. Транспонирование матрицы в тензорной форме выглядит как перестановка местами индексов, указывающих номер строки и номер столбца.

Определение 8. Тензор $\boldsymbol{B} = b^{j_1\ldots j_q}_{i_1\ldots i_p}$, полученный из тензора $\boldsymbol{A} = a^{j_1\ldots j_q}_{i_1\ldots i_p}$

перестановкой двух первых нижних индексов:

$$b^{j_1 \dots j_q}_{i_1 i_2 \dots i_p} = a^{j_1 \dots j_q}_{i_2 i_1 \dots i_p},$$

называют ***транспонированным*** к тензору $\boldsymbol{A}$. Транспонированными называют также тензоры, полученные перестановкой любой другой пары верхних или нижних индексов.

Нужно, естественно, убедиться, что если мы переставляем два верхних или два нижних индекса, то в результате получаем набор компонент, меняющихся при замене базиса по тензорному закону (10.1). При изменении порядка индексов набор компонент тензора в данном базисе остается неизменным, но меняется порядок этих компонент. Например, компоненты тензора *валентности* 2 можно записать в матрицу. Тогда транспонирование тензора будет означать транспонирование матрицы. Но то же происходит и в общем случае. Компоненты тензора валентности $p + q$ можно рассматривать как элементы $(p + q)$-мерной матрицы (при $p + q = 2$ это обычная матрица, ассоциирующаяся с квадратом, при $p + q = 3$ компоненты матрицы записываются в ячейки куба и т.д.). Если фиксировать все индексы, кроме двух переставляемых, мы получим в многомерной матрице плоское сечение, представляющее собой обычную матрицу. Транспонирование тензора есть транспонирование каждого такого сечения.

Перестановка в тензоре одного верхнего и одного нижнего индекса не имеет какого-либо содержательного смысла, так как при этом нарушается тензорный закон изменения компонент. Это лучше всего наблюдать на простых тензорах типа $(1, 1)$: наличие дополнительных индексов усложняет выкладки, но не дает принципиальных изменений. Тензор типа $(1, 1)$ будем трактовать как линейный оператор. Тогда перестановка индексов в компонентах матрицы оператора означает транспонирование этой матрицы. Но транспонирование матрицы линейного оператора в разных базисах приводит к разным операторам. Чтобы такая операция была законной, нужно ограничиться рассмотрением только *евклидова пространства* и *ортонормированных базисов* в нем. В этом случае транспонирование матрицы означает переход к сопряженному оператору.

Операцию транспонирования можно усложнить, повторяя ее с разными парами индексов. Пусть, например, среди верхних индексов выбрана группа из s индексов. Используя различные перестановки в этой группе индексов, мы можем расставить эти s индексов в любом порядке. Количество вариантов есть количество *перестановок* из s элементов, которое равно $s!$. Операцию множественной перестановки индексов мы также будем называть транспонированием, выделяя ***элементарное транспонирование***, заключающееся в перестановке пары индексов.

Определение 9. Тензор называют ***симметрическим по группе индексов***, если он не изменяется при любой перестановке в этой группе индексов. Тензор ***кососимметрический по группе индексов*** (***антисимметрический по группе индексов***), если при перестановке любой пары индексов из группы он меняет знак.

Особый интерес в связи с этим определением представляют ковариантные и контравариантные тензоры. ***Симметрический тензор*** — это ковариантный (контравариантный) тензор, симметрический по группе всех индексов. Аналогично понятие ***кососимметрического вектора***, относящееся к ковариантным или контравариантным тензорам

Симметрирование и альтернирование. Рассмотрим группу из r верхних (нижних) индексов у тензора $\boldsymbol{A}$ типа (p, q), где $r \leqslant q$ $(r \leqslant p)$. Перестановкой этой группы индексов можно получить $r!$ тензоров $\boldsymbol{A}_\sigma$, включая исходный. Умножим сумму всех этих тензоров на число $1/r!$ Мы получим новый тензор $\boldsymbol{A}^s$:

$$\boldsymbol{A}^s = \frac{1}{r!}\sum_\sigma \boldsymbol{A}_\sigma,$$

который будет симметрическим по выделенной группе из r индексов. Описанную операцию преобразования тензора, в результате которой получается тензор, симметрический по группе индексов, называют ***симметрированием***.

В частном случае пары индексов симметрирование выглядит наиболее просто. Например, для тензора a_{ij} симметрирование состоит в получении нового симметрического тензора

$$(a^s)_{ij} = (a^s)_{ji} = \frac{1}{2}(a_{ij} + a_{ji}).$$

Существует также операция, которая позволяет из данного тензора получить тензор, кососимметрический по группе индексов. Рассмотрим группу из s верхних (нижних) индексов. Любой тензор $\boldsymbol{A}_\sigma$, получаемый из тензора $\boldsymbol{A}$ перестановкой этих индексов, можно описать перестановкой $\sigma = (i_1, \ldots, i_s)$ из s элементов, причем исходному тензору будет соответствовать тождественная перестановка $(1, 2, \ldots, s)$. Обозначим через $|\sigma|$ количество инверсий в перестановке σ и рассмотрим сумму

$$\boldsymbol{A}^a = \frac{1}{s!}\sum_\sigma (-1)^{|\sigma|}\boldsymbol{A}_\sigma, \tag{3}$$

которая берется по всем перестановкам σ из s элементов. Операцию преобразования $\boldsymbol{A} \to \boldsymbol{A}^a$ называют ***альтернированием*** по указанной группе индексов. В результате альтернирования тензора получается тензор, кососимметрический по группе индексов. Действительно, перестановка двух индексов в группе меняет четность каждой перестановки σ в сумме (3). Значит, каждое слагаемое и вся сумма в целом меняют знак.

В случае тензора a_{ij} типа $(2, 0)$ альтернирование выглядит наиболее просто:

$$(a^a)_{ij} = -(a^a)_{ji} = \frac{1}{2}(a_{ij} - a_{ji}).$$

Произведение тензоров. Две полилинейные формы можно перемножить, образуя функцию от большего числа переменных. Например, из полилинейных форм $\varphi(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p; \boldsymbol{f}^1, \ldots, \boldsymbol{f}^q)$ и $\psi(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_r; \boldsymbol{g}^1, \ldots, \boldsymbol{g}^s)$ типов (p, q)

и (r,s) можно образовать новую полилинейную форму

$$\chi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p,\boldsymbol{y}_1,\ldots,\boldsymbol{y}_r;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q,\boldsymbol{g}^1,\ldots,\boldsymbol{g}^s)= \\ =\varphi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_p;\boldsymbol{f}^1,\ldots,\boldsymbol{f}^q)\psi(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_r;\boldsymbol{g}^1,\ldots,\boldsymbol{g}^s),$$

имеющую тип $(p+r,q+s)$. Аналогичная операция существует и для тензоров.

Определение 10. ***Произведением*** двух ***тензоров*** $\boldsymbol{A}=a^{j_1\ldots j_q}_{i_1\ldots i_p}$ и $\boldsymbol{B}=b^{l_1\ldots l_s}_{k_1\ldots k_r}$ типов (p,q) и (r,s) называют тензор $\boldsymbol{C}=\boldsymbol{A}\otimes\boldsymbol{B}$ типа $(p+r,q+s)$ с компонентами

$$c^{j_1\ldots j_q l_1\ldots l_s}_{i_1\ldots i_p k_1\ldots k_r}=a^{j_1\ldots j_q}_{i_1\ldots i_p}b^{l_1\ldots l_s}_{k_1\ldots k_r}.$$

Необходимо, конечно, убедиться, что указанный набор компонент действительно представляет собой тензор, т.е. меняется при замене базиса по тензорному закону (1). Непосредственная проверка закона изменения достаточно сложна. Можно ее обойти следующим образом. Каждый тензор можно трактовать как полилинейную форму такого же типа. При этом произведению тензоров будет соответствовать произведение полилинейных форм, так как компонентам перемножаемых тензоров будут соответствовать координаты полилинейных форм, т.е. значения этих форм на различных комбинациях базисных векторов. Убедиться же в том, что при перемножении полилинейных форм получается полилинейная форма, несложно.

Произведение тензоров не является коммутативным. Действительно, при перестановке сомножителей меняется порядок индексов в их произведении. А это ведет к изменению результата. Чтобы проанализировать ситуацию подробнее, можно рассмотреть произведение двух *ковекторов* $\boldsymbol{A}=a_i$ и $\boldsymbol{B}=b_j$. Их тензорным произведением будет ковариантный тензор $\boldsymbol{C}=c_{ij}$ ранга 2 с компонентами $c_{ij}=a_ib_j$. Если изменить порядок сомножителей, то получим $c'_{ij}=b_ia_j=c_{ji}$, т.е. *тензор* $\boldsymbol{C}'$, *транспонированный* к тензору $\boldsymbol{C}$. Тензоры $\boldsymbol{C}$ и $\boldsymbol{C}'$ совпадают, если каждый из них является симметрическим. Например, если

$$f_i=\begin{cases}1, & i=1;\\ 0, & i\neq 1;\end{cases}\qquad g_j=\begin{cases}1, & j=2;\\ 0, & j\neq 2,\end{cases}$$

то $c_{12}=f_1g_2=1$, $c_{21}=f_2g_1=0$, и поэтому $f_i\otimes g_j\neq g_j\otimes f_i$.

Умножение тензоров является ассоциативным и дистрибутивным по отношению к сложению. Это напоминает свойства умножения линейных операторов. Однако отметим, что произведением двух линейных операторов является линейный оператор, а их произведением как тензоров является тензор типа $(2,2)$, который трактовать как линейный оператор нельзя. Эти две операции принципиально разные, хотя и обладают некоторыми одинаковыми алгебраическими свойствами.

Произведение тензоров открывает возможность получать новые тензоры из тензоров более низкой валентности. Отталкиваясь от векторов и ковекторов, мы можем получать тензоры любого типа. Однако не каждый тензор может быть представлен в виде произведения тензоров. Например, билинейная форма, являющаяся произведением двух ковекторов, имеет матрицу специального вида, ранг которой равен единице. В качестве контрпримера достаточно взять билинейную форму с рангом, равным 2.

Оказывается, что любой тензор типа (p, q) может быть представлен в виде линейной комбинации элементарных тензоров вида $\boldsymbol{x}^1 \otimes \ldots \otimes \boldsymbol{x}^p \otimes \boldsymbol{y}_1 \otimes \ldots \otimes \boldsymbol{y}_q$, являющихся произведением p ковекторов $\boldsymbol{x}^i$ и q векторов $\boldsymbol{y}_j$. Действительно, вспомним базис в пространстве $\mathcal{T}_{p,q}$ тензоров типа (p, q), рассмотренный в доказательстве теоремы 4. Каждый из тензоров этого базиса является элементарным. Тензор $\boldsymbol{T}^{i_1 \ldots i_p}_{j_1 \ldots j_q}$, у которого единичное значение имеет компонента с индексами нижними $i_1, \ldots, i_p$ и верхними $j_1, \ldots, j_q$, является произведением

$$\boldsymbol{T}^{i_1 \ldots i_p}_{j_1 \ldots j_q} = \boldsymbol{f}^{i_1} \otimes \ldots \otimes \boldsymbol{f}^{i_p} \otimes \boldsymbol{e}_{j_1} \otimes \ldots \otimes \boldsymbol{e}_{j_q}$$

векторов базиса $(e_1, \ldots, e_n)$ и ковекторов двойственного ему базиса $(\boldsymbol{f}^1, \ldots, \boldsymbol{f}^n)$. Поэтому любой тензор $\boldsymbol{A}^{j_1 \ldots j_q}_{i_1 \ldots i_p}$ представим в виде

$$\boldsymbol{A} = a^{j_1 \ldots j_q}_{i_1 \ldots i_p} \boldsymbol{f}^{i_1} \otimes \ldots \otimes \boldsymbol{f}^{i_p} \otimes \boldsymbol{e}_{j_1} \otimes \ldots \otimes \boldsymbol{e}_{j_q}.$$

Свертывание. Если тензорное умножение позволяет получить тензоры более высокой валентности, то следующая операция, наоборот, понижает валентность вектора.

Определение 11. ***Сверткой тензора*** $\boldsymbol{A} = a^{j_1 \ldots j_q}_{i_1 \ldots i_p}$ типа (p, q) по одному верхнему и одному нижнему индексам, например по индексами i_1 и j_1, называют тензор $\boldsymbol{B} = b^{j_2 \ldots j_q}_{i_2 \ldots i_p}$ типа $(p-1, q-1)$ с компонентами

$$b^{j_2 \ldots j_q}_{i_2 \ldots i_p} = a^{k j_2 \ldots j_q}_{k i_2 \ldots i_p}.$$

Напомним, что наличие одинаковых верхнего и нижнего индексов предусматривает суммирование по этому индексу (*правило индексов*). Можно убедиться, что введенная операция приводит к тензору, т.е. для нового набора компонент остается верным закон (1) преобразования тензоров.

Пример 11. Сверткой тензора типа $(1, 1)$ является тензор валентности 0, т.е. инвариант. Тензор типа $(1, 1)$ представляет собой совокупность элементов матрицы линейного оператора, а его свертка — это сумма диагональных элементов матрицы оператора, т.е. не что иное, как *след матрицы линейного оператора*, который от выбора базиса не зависит.

Говорят также о ***свертке двух тензоров***, подразумевая под этим свертку произведения этих тензоров, причем один из двух индексов, по которым выполняется свертка, относится к первому тензору, а второй — ко второму. Например, выражение $a_{ij}^{k}b_{k}^{rs}$ означает свертку тензоров $\boldsymbol{A}$ типа $(2,1)$ и $\boldsymbol{B}$ типа $(1,2)$, в результате которой получается тензор типа $(2,2)$.

Свертка тензора или двух тензоров может выполняться не по одной паре индексов, а по нескольким. Например, рассмотрим тензор a_i^j типа $(1,1)$. Произведением этого тензора на себя будет тензор $a_i^j a_k^l$ типа $(2,2)$. Это произведение можно свернуть по двум парам индексов. Получим инвариант (тензор типа $(0,0)$) $a_i^j a_j^i$, который называют инвариантом второго порядка, в отличие от инварианта первого порядка — следа линейного оператора. Отметим, что другой вариант свертки по двум парам индексов $a_i^i a_j^j$ дает квадрат инварианта первого порядка.

В евклидовом пространстве выделен *ковариантный метрический тензор* g_{ij} и *контравариантный метрический тензор* g^{ij}. Свертка тензора $\boldsymbol{A}$ типа (p,q) с g_{ij} приводит к тензору типа $(p+1,q-1)$ и представляет собой операцию ***опускания индекса тензора***, а свертка $\boldsymbol{A}$ с g^{ij} приводит к тензору типа $(p-1,q+1)$, т.е. к ***поднятию индекса тензора***. Компоненты метрического тензора в ортонормированном базисе составляют единичную матрицу. Поэтому в таком базисе опускание (поднятие) индекса выглядит как простая перестановка индекса сверху вниз (снизу вверх), т.е. различие между верхними и нижними индексами исчезает. Если рассматривается евклидово пространство и в нем только ортонормированные базисы, то все индексы можно записывать или только внизу, или только вверху. Возникающий при этом объект называют ***евклидовым тензором***.