Глава XII. Линейное пространство над произвольным полем

В главе IV мы рассматривали вещественные линейные пространства. Вещественными они назывались потому, что операция умножения на число предусматривала умножение вектора на вещественные числа. В §43 мы уже отмечали, что все свойства линейного пространства остаются в силе, если поле вещественных чисел $\mathbb{R}$ заменить произвольным полем P. В этой главе воспроизводится известное определение линейного пространства для случая произвольного поля, напоминаются уже известные факты и рассматриваются новые.

§60. Определение и основные свойства

Пусть дано поле P. Непустое множество V называется ***линейным*** или ***векторным пространством над полем*** P, если на этом множестве определены внутренний закон композиции $V \times V \to V$, называемый сложением, и внешний закон композиции $P \times V \to V$, называемый умножением на число из поля P, удовлетворяющие следующим аксиомам: для любых $a, b, c \in V$ и $\alpha, \beta \in P$

1) $a + b = b + a$;
2) $(a + b) + c = a + (b + c)$;
3) существует элемент $\theta \in V$ такой, что $a + \theta = \theta + a = a$;
4) для любого элемента $a \in V$ существует элемент $-a \in V$ такой, что $a + (-a) = (-a) + a = \theta$;
5) $1 \cdot a = a$;
6) $\alpha(\beta a) = (\alpha\beta)a$;
7) $(\alpha + \beta)a = \alpha a + \beta a$;
8) $\alpha(a + b) = \alpha a + \alpha b$.

Линейное пространство над полем $\mathbb{R}$ называется ***вещественным линейным пространством***, а над полем $\mathbb{C}$ – ***комплексным***.

Примеры. 1. Арифметическое пространство P^n – линейное пространство над полем P. Как и в случае $\mathbb{R}^n$, здесь $P^n = \{x = (x_1, \ldots, x_n) \,|\, x_k \in P,\ k = \overline{1,n}\}$ с операциями: если $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$, $\alpha \in P$, то

$$x + y = (x_1 + y_1, \ldots, x_n + y_n), \ \alpha x = (\alpha x_1, \ldots, \alpha x_n).$$

2. В частности, арифметическое пространство $\mathbb{C}^n$ есть комплексное линейное пространство. В таком пространстве арифметические векторы умножаются на комплексные числа. Однако можно было бы ограничиться умножением только на вещественные числа, при этом результат умножения, несомненно, останется арифметическим вектором из $\mathbb{C}^n$. Это означает, что внешний закон композиции можно

вводить и как умножение на вещественные числа, т.е. как отображение $\mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$. Символом $\mathbb{C}^n_{\mathbb{R}}$ будем обозначать множество $\mathbb{C}^n$, в котором внешний закон композиции определен как $\mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$. Нетрудно проверить, что $\mathbb{C}^n_{\mathbb{R}}$ – вещественное линейное пространство. В частности, само поле $\mathbb{C}$ можно рассматривать как вещественное линейное пространство $\mathbb{C}_{\mathbb{R}}$.

3. Аксиомы линейного пространства выявляют алгебраические свойства многих классов функций, часто встречающихся в математическом анализе: например, $C[0,1]$ – множество непрерывных функций на отрезке $[0,1]$; $D[0,1]$ – множество всех дифференцируемых функций на отрезке $[0,1]$; $C^1[0,1]$ – множество всех функций, непрерывных вместе с первой производной на отрезке $[0,1]$. Хотя у каждого из перечисленных множеств имеются и другие замечательные свойства, все они охватываются единой алгебраической схемой, аксиоматизируемой определением линейного пространства. Нетрудно проверить, что эти множества являются вещественными линейными пространствами относительно обычных операций сложения функций и умножения функции на число. Линейные пространства функций принято называть ***функциональными пространствами***.

§61. Линейная зависимость. Ранг и база системы векторов

Понятие линейной зависимости и связанные с ней утверждения рассматривались в главе IV (§ 14–16) и относились к вещественным линейным пространствам. Все определения, а также теоремы и их доказательства остаются в силе и в произвольном линейном пространстве. Введем новые понятия. В дальнейшем, употребляя термин "линейное пространство", будем иметь в виду линейное пространство над произвольным полем P, не оговаривая это дополнительно.

Ранг и база системы векторов. Будем рассматривать конечные системы $a_1, \ldots, a_k$ векторов линейного пространства.

Линейно независимая подсистема системы векторов, через которую линейно выражается любой вектор системы, называется ***базой этой системы векторов***.

Примеры. 1. В системе из нулевых векторов нет ни одной базы, так как любая ее подсистема линейно зависима.

2. Базисные строки матрицы согласно теореме о базисном миноре образуют базу системы строк, рассматриваемых как векторы арифметического пространства. Это же относится и к базисным столбцам.

3. В системе трех неколлинеарных векторов плоскости V_2 любая пара векторов образует базу.

Обратим внимание на то, что в этих примерах система векторов может обладать не единственной базой, но при этом все базы одной системы состоят из одинакового числа векторов.

Теорема 61.1. *Подсистема системы векторов является базой системы векторов тогда и только тогда, когда образует максимальную линейно независимую подсистему.*

$y_1 = \alpha(x_0 - x_0') + y \in L$, так как $x_0 - x_0' \in L$, $y \in L$. Следовательно, $\alpha x_0 + L \subset \alpha x_0' + L$. С другой стороны, для элемента из правого множества $\alpha x_0' + y$, $y \in L$, имеем $\alpha x_0' + y = \{x_0' = x_0 + y_1$, где $y_1 \in L\} = \alpha(x_0 + y_1) + y = \alpha x_0 + \alpha y_1 + y = \alpha x_0 + y_2$, где $y_2 = \alpha y_1 + y \in L$. Следовательно, $\alpha x_0' + L \subset \alpha x_0' + L$. Оба вложения доказывают (67.4).

Таким образом, правило (67.3) определяет внешний закон композиции на $V|L$ (рис. 3).

Теорема 67.5. *Множество $V|L$ всех линейных многообразий с направляющим подпространством L есть линейное пространство относительно законов композиции* (67.2), (67.3).

Доказательство. Справедливость аксиом группы сложения вытекает из того, что $V|L$ – фактор-группа. Впрочем, непосредственная проверка этих аксиом так же проста, как и проверка остальных аксиом линейного пространства, ее мы предоставляем читателю. ■

Линейное пространство $V|L$ называется ***фактор-пространством линейного пространства*** V ***по подпространству*** L.

Теорема 67.6. *Фактор-пространство $V|L$ изоморфно L^δ.*

Доказательство. Если $L = \{\theta\}$, то $L^\delta = V$, $V|L = V$ и, следовательно, $V|L \cong L^\delta$. Если $L = V$, то $L^\delta = \{\theta\}$, $V|L = \{\theta\}$ и $V|L \cong \{\theta\}$. Пусть $0 < \dim L < \dim V$ и L^δ – какое-нибудь дополнительное подпространство к L. Согласно теореме 67.4 в каждом многообразии $H = x + L$ существует, и притом единственный, вектор $z \in L^\delta$, так что $H = z + L$. Таким образом, $V|L = \{H = z + L \,|\, z \in L^\delta\}$. Отображение $\varphi : L^\delta \to V|L$, которое каждому вектору $z \in L^\delta$ ставит в соответствие линейное аффинное многообразие $H = z + L \in V|L$, биективно в силу теоремы 67.4, оно сохраняет законы композиции в силу (67.2), (67.3). Таким образом, φ – изоморфизм. ■

Следствие 2. $\dim V|L = \dim V - \dim L$.

Глава XIII. Евклидовы и унитарные пространства

Линейные пространства, изучавшиеся нами до сих пор, несмотря на абстрактный характер, по своим свойствам очень близки к геометрическим пространствам. Однако в них еще не нашли отражение понятия, связанные с измерениями, такие, как длина вектора и угол между векторами. В §22 мы видели, что эти величины тесно связаны с понятием скалярного произведения векторов. Они просто выражаются через скалярное произведение, а само скалярное произведение векторов оказывается при этом весьма простой с точки зрения алгебры функцией: оно линейно по каждой из двух переменных. Это дает основание при введении метрических понятий в теории линейных пространств отталкиваться от понятия скалярного произведения.

В этой главе рассматриваются только вещественные и комплексные пространства. Итак, V – линейное пространство над полем P, где P – либо $\mathbb{R}$, либо $\mathbb{C}$.

§68. Скалярное произведение

Пусть V – вещественное или комплексное линейное пространство. Отображение

$$(\,,\,) : \quad V \times V \to P$$

называется ***скалярным произведением***, если оно удовлетворяет следующим аксиомам: для любых $x, y, z \in V$ и любого $\alpha \in P$

$$\begin{aligned} &1)\ (x, y) = \overline{(y, x)};\\ &2)\ (\alpha x, y) = \alpha(x, y);\\ &3)\ (x + y, z) = (x, z) + (y, z);\\ &4)\ (x, x) \geq 0 \text{ для любого } x \in V,\\ &\quad\ (x, x) = 0 \text{ тогда и только тогда, когда } x = \theta. \end{aligned} \tag{68.1}$$

Число (x, y) называется ***скалярным произведением векторов*** x ***и*** y, аксиомы 1 – 4 называются ***аксиомами скалярного произведения***.

Замечание 1. В вещественном случае черта в первой аксиоме может быть опущена.

Замечание 2. Аксиома 4 в комплексном случае на первый взгляд кажется парадоксальной, так как для комплексных чисел знак не определен. Однако из первой аксиомы следует, что скалярный квадрат (x, x) есть вещественное число.

Вещественное линейное пространство со скалярным произведением называется ***евклидовым пространством***, а комплексное – ***унитарным***. Обозначение: E и U соответственно.

Примеры. 1. В геометрических пространствах V_n, где $n = 1, 2, 3$, было введено скалярное произведение (22.1) на основании метрики,

Доказательство. Необходимость. Пусть в системе векторов $a_1, \ldots, a_r, \ldots, a_k$ подсистема $a_1, \ldots, a_r$ образует базу (здесь для упрощения записи в качестве подсистемы взяты первые r векторов, что, как будет видно из доказательства, не нарушает общности рассуждений). Тогда любая бо́льшая подсистема будет линейно зависимой в силу теоремы 16.2, так как любой ее вектор линейно выражается через базу $a_1, \ldots, a_r$. Таким образом, база образует максимальную линейно независимую подсистему системы векторов.

Достаточность. Пусть $a_1, \ldots, a_r$ – максимальная линейно независимая подсистема системы векторов $a_1, \ldots, a_r, \ldots, a_k$. Тогда для любого вектора a_i, $i = \overline{1,k}$, подсистема $a_1, \ldots, a_r, a_i$ линейно зависима (если $i \leq r$, то как подсистема, содержащая два одинаковых вектора; если же $i > r$, то как подсистема из $r + 1 > r$ векторов). В силу теоремы 14.6 вектор a_i линейно выражается через $a_1, \ldots, a_r$. Таким образом, $a_1, \ldots, a_r$ – база. ■

Следствие 1. *Все базы одной системы векторов состоят из одинакового числа векторов, равного максимальному числу линейно независимых векторов системы.*

Число векторов базы называется ***рангом системы векторов***. Очевидно, ранг системы векторов равен максимальному числу линейно независимых векторов системы. Обозначение: $\mathrm{rg}(a_1, \ldots, a_n)$.

Две системы векторов линейного пространства называются ***эквивалентными***, если каждая из этих систем выражается через другую. Из определения следует, что ***база системы векторов эквивалентна самой системе***.

Теорема 61.2. *Если система векторов $a_1, \ldots, a_k$ линейно выражается через $b_1, \ldots, b_m$, то $\mathrm{rg}(a_1, \ldots, a_k) \leq \mathrm{rg}(b_1, \ldots, b_m)$.*

Доказательство. Пусть $a_1, \ldots, a_r$ и $b_1, \ldots, b_s$ – базы рассматриваемых систем. Из условия теоремы и транзитивности свойства "линейной выражаемости" следует, что база $a_1, \ldots, a_r$ первой системы линейно выражается через базу $b_1, \ldots, b_s$ второй. Тогда $r \leq s$, так как в противном случае, если $r > s$, система $a_1, \ldots, a_r$ была бы линейно зависимой в силу теоремы 16.2. ■

Следствие 2. *Ранги эквивалентных систем совпадают.*

Следствие 3. *Эквивалентные линейно независимые системы векторов состоят из одинакового числа векторов.*

§62. Базис и размерность

Говорят, что ***система векторов*** $e_1, \ldots, e_n$ линейного пространства V ***порождает пространство*** V, если любой вектор $x \in V$ является линейной комбинацией $e_1, \ldots, e_n$.

Упорядоченная система векторов $e_1, \ldots, e_n$ линейного пространства V называется ***базисом*** V, если она линейно независима и порождает V. Очевидно, что это определение совпадает с определением

базиса, данным в §17 для вещественного линейного пространства.

Теорема 62.1. *Любые два базиса линейного пространства состоят из одинакового числа векторов.*

Это утверждение вытекает из эквивалентности двух базисов линейного пространства и следствия 2 (§61). ■

Итак, число векторов базиса не зависит от самого базиса и однозначно определяется самим пространством. Ответ на вопрос, что представляет собой это число, дается теоремой 17.1, справедливой и в пространстве над произвольным полем.

Напомним, что число векторов базиса линейного пространства V называется ***размерностью пространства*** V и обозначается символом $\dim V$. Размерность нулевого пространства по определению считается равной нулю. Из теоремы 17.1 следует, что размерность линейного пространства совпадает с максимальным числом линейно независимых векторов этого пространства. Линейное пространство размерности n, где $n \in \mathbb{N}$, называется ***n-мерным***. Нулевое пространство и n-мерные пространства называются ***конечномерными***.

Линейное пространство называется ***бесконечномерным***, если для любого числа $k \in \mathbb{N}$ в пространстве существует k линейно независимых векторов. Пример бесконечномерного пространства дан в §17 (пример 5). В курсе линейной алгебры изучаются лишь конечномерные линейные пространства.

Теорема 62.2 (о неполном базисе). *В n-мерном пространстве любую линейно независимую систему из k, где $k < n$, векторов можно дополнить до базиса.*

Доказательство. Пусть $e_1, \ldots, e_k$ – линейно независимая система векторов пространства V. Так как $k < n$, то в силу теоремы 17.1 система векторов $e_1, \ldots, e_k$ не является базисом V и, следовательно, не порождает всего пространства V. Пусть вектор $e_{k+1} \in V$ не является линейной комбинацией $e_1, \ldots, e_k$, тогда система векторов $e_1, \ldots, e_k, e_{k+1}$ линейно независима (в силу теоремы 14.6). Если $k+1 = n$, то эта система векторов образует базис V; если же $k+1 < n$, то аналогичным образом построим линейно независимую систему из $k + 2$ векторов. За $n - k$ таких шагов мы построим искомый базис $e_1, \ldots, e_k, \ldots, e_n$. ■

Точно так же, как и в §17, определяются координаты вектора в базисе и матрица перехода от одного базиса к другому. При этом все факты, связанные с этими понятиями, остаются в силе и в общем случае, так как их доказательства опираются на те свойства вещественных чисел, которые имеют место в любом поле. Напомним эти факты.

1°. *Разложение вектора по базису единственно.*

2°. *Координаты вектора обладают свойством линейности.*

3°. *При переходе от базиса e к базису $f = eQ$ координаты вектора x изменяются по следующему закону: $x_e = Qx_f$.*

Примеры. 1. В арифметических пространствах P^n, $\mathbb{R}^n$, $\mathbb{C}^n$ еди-

Вектор единичной длины называется ***нормированным***. Любой ненулевой вектор можно нормировать, поделив его на его длину.

Теорема 69.1. *В евклидовом (унитарном) пространстве для любых векторов x, y имеют место неравенства*

$$||x| - |y|| \leq |x + y| \leq |x| + |y|. \tag{69.2}$$

Доказательство. Имеем $|x + y|^2 = (x + y, x + y) = (x, x) + (x, y) + (y, x) + (y, y)$. Применив к этой сумме числовые неравенства треугольника (которые имеют место как для вещественных, так и для комплексных чисел), с учетом неравенства (69.1) получим, что

$$|x + y|^2 \leq |x|^2 + 2|x| \cdot |y| + |y|^2 = (|x| + |y|)^2,$$
$$|x + y|^2 \geq |x|^2 - 2|x| \cdot |y| + |y|^2 = (|x| - |y|)^2.$$

Отсюда следует (69.2). ■

Неравенства (69.2) называются ***неравенствами треугольника в евклидовом (унитарном) пространстве***.

В евклидовом пространстве ***углом*** между ненулевыми векторами x и y называется угол φ, $0 \leq \varphi \leq \pi$, для которого

$$\cos\varphi = \frac{(x, y)}{|x| \cdot |y|}. \tag{69.3}$$

Корректность определения следует из неравенства Коши–Буняковского (69.1).

В унитарном пространстве понятие угла между векторами не определено. Однако и в евклидовом и в унитарном пространстве можно ввести обобщение понятия прямого угла.

§70. Ортогональные векторы

Два вектора $x, y \in E\,(U)$ называются ***ортогональными***, если $(x, y) = 0$. Из свойства 4° скалярного произведения (§68) следует, что нулевой вектор θ, и только нулевой, ортогонален любому вектору пространства.

В евклидовом пространстве вследствие (69.3) ортогональность векторов x и y означает, что либо один из них нулевой, либо угол между ними равен $\pi/2$.

Система векторов $x_1, \ldots, x_k \in E\,(U)$ называется ***ортогональной системой***, если

$$(x_i, x_j) = 0 \quad \text{при} \quad i \neq j. \tag{70.1}$$

Система векторов $x_1, \ldots, x_k \in E\,(U)$ называется ***ортонормированной системой***, если $(x_i, x_j) = \delta_{ij}$, где δ_{ij} – символ Кронекера, т.е.

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{70.2}$$

Теорема 70.1. *Ортогональная система ненулевых векторов линейно независима.*

Доказательство. Пусть $x_1, \ldots, x_k \in E\,(U)$ – ортогональная система ненулевых векторов. Умножая обе части равенства

$$\alpha_1 x_1 + \ldots + \alpha_i x_i + \ldots + \alpha_k x_k = \theta \tag{70.3}$$

скалярно на x_i, получаем в силу (70.1)

$$\alpha_i(x_i, x_i) = 0, \quad i = \overline{1,k}. \tag{70.4}$$

По условию $x_i \neq \theta$, значит, $(x_i, x_i) \neq 0$, и в силу (70.4) все коэффициенты линейной комбинации в (70.3) равны нулю. Следовательно, векторы рассматриваемой системы линейно независимы. ■

Следствие 1. *Ортонормированная система векторов линейно независима.*

Следствие 2. *В n-мерном евклидовом (унитарном) пространстве любая ортонормированная система из n векторов образует базис.*

Базис, векторы которого образуют ортонормированную систему, называется ***ортонормированным базисом***. В соответствии с (70.2) $e_1, \ldots, e_k$ – ортонормированный базис $E\,(U)$, если

$$(e_i, e_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{70.5}$$

Теорема 70.2. *В евклидовом (унитарном) пространстве координаты $x_1, \ldots, x_n$ вектора x в базисе $e = (e_1, \ldots, e_n)$ вычисляются по правилу*

$$x_i = (x, e_i), \quad i = \overline{1,n}, \tag{70.6}$$

тогда и только тогда, когда e – ортонормированный базис.

Доказательство. Необходимость. Пусть для любого вектора $x \in E\,(U)$ координаты в базисе e вычисляются согласно (70.6). Тогда по этому же правилу вычисляются координаты и базисных векторов, которые известны, так как

$$e_i = 0e_1 + \ldots + 0e_{i-1} + 1e_i + 0e_{i+1} + \ldots + 0e_n, \quad i = \overline{1,n}.$$

Сравнение координат вектора e_i с правилом (70.6) приводит к требуемым равенствам (70.5).

Достаточность. Если e – ортонормированный базис и $x = \sum_{j=1}^n x_j e_j$, то свойство линейности скалярного произведения с учетом (70.5) приводит к (70.6). ■

Теорема 70.3. *В евклидовом (унитарном) пространстве скалярное произведение векторов $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^n y_i e_i$, заданных своими координатами в базисе e, вычисляется по правилу*

$$(x, y) = \sum_{i=1}^n x_i \overline{y_i} \tag{70.7}$$

Как указывалось в §30, множество всех решений однородной системы уравнений

$$Ax = 0 \tag{64.1}$$

с n неизвестными образует линейное подпространство арифметического пространства $\mathbb{R}^n$ (или P^n в общем случае, когда $A \in P^{m \times n}$). Про это подпространство L говорят, что оно задано однородной системой (64.1), и записывают в виде

$$L : \quad Ax = 0. \tag{64.2}$$

Другой способ задания линейного подпространства дается понятием линейной оболочки.

Линейная оболочка. Пусть $a_1, \ldots, a_k$ – система векторов линейного пространства V над полем P. ***Линейной оболочкой системы векторов*** $a_1, \ldots, a_k$ называется множество всевозможных линейных комбинаций этих векторов. Говорят также, что линейная оболочка ***натянута на векторы*** $a_1, \ldots, a_k$. Обозначение: $\mathcal{L}(a_1, \ldots, a_k)$. Итак,

$$\mathcal{L}(a_1, \ldots, a_k) = \left\{ a = \sum_{i=1}^k \alpha_i a_i \,|\, \alpha_i \in P,\ i = \overline{1,k} \right\}.$$

Из определения следует, что каждое конечномерное пространство является линейной оболочкой векторов своего базиса. В частности, линейное подпространство, заданное однородной системой (64.2), является линейной оболочкой фундаментальной системы решений (§30).

Теорема 64.1. *Если $a_1, \ldots, a_k$ – векторы линейного пространства V, то $\mathcal{L}(a_1, \ldots, a_k)$ является линейным подпространством пространства V.*

Это утверждение вытекает из теоремы 18.1, так как для линейной оболочки $\mathcal{L}(a_1, \ldots, a_k)$ обе импликации имеют место. ■

Итак, линейная оболочка $\mathcal{L}(a_1, \ldots, a_k)$ – это линейное подпространство, которое порождается векторами $a_1, \ldots, a_k$.

Теорема 64.2. *Две системы векторов линейного пространства эквивалентны тогда и только тогда, когда их линейные оболочки совпадают.*

Доказательство. Необходимость. Пусть системы векторов $a_1, \ldots, a_k$ и $b_1, \ldots, b_m$ эквивалентны. Тогда $\mathcal{L}(a_1, \ldots, a_k) = \mathcal{L}(b_1, \ldots, b_m)$, так как для этих множеств, как легко проверить, имеет место двустороннее вложение. Достаточность очевидна. ■

Следствие 1. *Линейная оболочка системы векторов совпадает с линейной оболочкой своей базы.*

Следствие 2. *Размерность линейной оболочки системы векторов равна рангу этой системы:*

$$\dim \mathcal{L}(a_1, \ldots, a_k) = \mathrm{rg}(a_1, \ldots, a_k). \tag{64.3}$$

Теорема 64.3 (о монотонности размерности). *Размерность линейного подпространства не превосходит размерности пространства. Подпространство той же размерности, что и все пространство, совпадает с пространством.*

Доказательство. Пусть L – линейное подпространство V, $\dim L = k$, $\dim V = n$. Тогда $k \leq n$, так как в противном случае в n-мерном пространстве V существует k, где $k > n$, линейно независимых векторов (например, векторы базиса L). Пусть $k = n$ и $e_1, \ldots, e_n$ – базис L. Так как $\dim V = n$, то векторы $e_1, \ldots, e_n$ образуют базис V (§17, утверждение 1). Таким образом, $L = V = \mathcal{L}(e_1, \ldots, e_n)$. ■

§65. Сумма и пересечение подпространств

Пусть $L_1, \ldots, L_k$ – линейные подпространства линейного пространства V. ***Суммой подпространств*** $L_1, \ldots, L_k$ называется множество всевозможных векторов x, представимых в виде

$$x = x_1 + \ldots + x_k, \tag{65.1}$$

где $x_i \in L_i$, $i = \overline{1,k}$. Обозначение: $L_1 + \ldots + L_k$ или $\sum_{i=1}^k L_i$.

Итак,

$$\sum_{i=1}^k L_i = \{x = x_1 + \ldots + x_k \,|\, x_i \in L_i,\ i = \overline{1,k}\}.$$

Представление (65.1) вектора x называется ***разложением вектора*** x ***по подпространствам*** $L_1, \ldots, L_k$.

Пересечением подпространств $L_1, \ldots, L_k$ называется множество

$$L_1 \cap \ldots \cap L_k = \{x \in V \,|\, x_i \in L_i,\ i = \overline{1,k}\}.$$

Заметим, что пересечение подпространств не пусто, так как всегда содержит нулевой вектор θ пространства.

Теорема 65.1. *Сумма и пересечение подпространств линейного пространства V являются линейными подпространствами пространства V.*

Это утверждение вытекает из теоремы 18.1, так как для $\sum_{i=1}^k L_i$ и $\bigcap_{i=1}^k L_i$ справедливы обе импликации. ■

Заметим, что каждое из подпространств L_i, $i = \overline{1,k}$, а также их пересечение являются линейными подпространствами суммы $L_1 + \ldots + L_k$, а пересечение $L_1 \cap \ldots \cap L_k$ – линейным подпространством каждого из L_i, $i = \overline{1,k}$.

Теорема 65.2. *Сумма линейных подпространств есть линейная оболочка совокупности базисов слагаемых подпространств.*

Доказательство. Пусть $e_1, \ldots, e_m, f_1, \ldots, f_s, \ldots, g_1, \ldots, g_t$ – базисы подпространств $L_1, L_2, \ldots, L_k$ соответственно. Положим $W = \mathcal{L}(e_1, \ldots, e_m, f_1, \ldots, f_s, \ldots, g_1, \ldots, g_t)$. Тогда $W = L_1 + \ldots + L_k$, так как для этих множеств, как нетрудно проверить, имеет место двустороннее вложение. ■

с матрицей коэффициентов $G(a_1, \ldots, a_k)$. Из существования нетривиального решения полученной системы уравнений следует (теорема 28.6), что $\det G(a_1, \ldots, a_k) = 0$.

Достаточность. Пусть $\det G(a_1, \ldots, a_k) = 0$. Тогда система (71.3) имеет нетривиальное решение $\alpha_1, \ldots, \alpha_k$. Перепишем систему (71.3) в виде

$$\begin{cases} \left(\sum_{i=1}^k \alpha_i a_i, a_1\right) = 0, \\ \ldots \quad \ldots \quad \ldots \quad \ldots \\ \left(\sum_{i=1}^k \alpha_i a_i, a_k\right) = 0. \end{cases} \tag{71.4}$$

Это означает, что вектор $g = \sum_{i=1}^k \alpha_i a_i$, с одной стороны, принадлежит $\mathcal{L}(a_1, \ldots, a_k)$, а с другой стороны, ортогонален $\mathcal{L}(a_1, \ldots, a_k)$. Согласно аксиоме 4 скалярного произведения вектор g может быть только нулевым. Значит, для векторов $a_1, \ldots, a_k$ имеет место соотношение (71.2), откуда с учетом нетривиальности набора $\alpha_1, \ldots, \alpha_k$ следует линейная зависимость $a_1, \ldots, a_k$. ■

Матрица $A \in \mathbb{C}^{n \times n}$ называется ***эрмитовой матрицей***, если

$$A^H = A; \tag{71.5}$$

матрица $A \in \mathbb{R}^{n \times n}$ называется ***симметрической*** (или ***вещественной эрмитовой***) ***матрицей***, если

$$A^T = A. \tag{71.6}$$

Из (71.5) следует, что $|A| \in \mathbb{R}$. Таким образом, для эрмитовых матриц (комплексных и вещественных) можно говорить о знаке определителя.

Теорема 71.2. *Матрица Грама системы векторов евклидова (унитарного) пространства эрмитова.*

Доказательство. Пусть $G(a_1, \ldots, a_n) = G = (g_{ij})$. Из (71.1) следует, что $g_{ij} = (a_j, a_i)$, $g_{ji} = (a_i, a_j)$, т.е. $g_{ij} = \overline{g_{ji}}$. Это означает, что $G^H = G$ или, в вещественном случае, $G^T = G$. ■

Теорема 71.3. *Определитель Грама линейно независимой системы векторов в евклидовом (унитарном) пространстве положителен.*

Доказательство. Пусть $a_1, \ldots, a_k$ – линейно независимая система векторов евклидова пространства. Тогда $\dim \mathcal{L}(a_1, \ldots, a_k) = k$. Выберем ортонормированный базис $e_1, \ldots, e_k$ линейной оболочки $\mathcal{L}(a_1, \ldots, a_k)$. Составим матрицу

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ \ldots & \ldots & \ldots & \ldots \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{bmatrix} \in \mathbb{C}^{k \times k},$$

столбцами которой являются координаты векторов $a_1, \ldots, a_k$ в базисе $e = (e_1, \ldots, e_k)$. Тогда, согласно (70.7), $(a_i, a_j) = \sum_{s=1}^k a_{sj} \overline{a_{si}} =$

$\sum_{s=1}^k a^h_{is} a_{sj} = \{A^H A\}_{ij}$, $i = \overline{1,k}$, $j = \overline{1,k}$. Следовательно,

$$G(a_1, \ldots, a_k) = A^H A \tag{71.7}$$

и $\det G(a_1, \ldots, a_k) = |\det A|^2$. Таким образом,

$$\det G(a_1, \ldots, a_k) \geq 0. \tag{71.8}$$

По условию система $a_1, \ldots, a_k$ линейно независима, поэтому, согласно теореме 71.1, $\det G(a_1, \ldots, a_k) \neq 0$. Отсюда с учетом (71.8) следует, что $\det G(a_1, \ldots, a_k) > 0$. ■

Замечание. Равенство (71.7) дает компактную форму записи матрицы Грама, в частности в вещественном случае

$$G(a_1, \ldots, a_k) = A^T A.$$

§72. Ортогональное дополнение

Задача о перпендикуляре. Пусть L – линейное подпространство евклидова (унитарного) пространства $E\,(U)$. Вектор x называется ***ортогональным к подпространству*** L, если он ортогонален любому вектору $y \in L$. Обозначение: $x \perp L$.

Очевидно, что $x \perp \mathcal{L}(a_1, \ldots, a_k)$ тогда и только тогда, когда $x \perp a_i$, $i = \overline{1,k}$. Совокупность всех векторов $x \in E\,(U)$, ортогональных подпространству L, называется ***ортогональным дополнением*** к L. Обозначение: $L^\perp$.

Теорема 72.1. *Ортогональное дополнение к подпространству является линейным подпространством.*

Доказательство. Пусть $y_1, y_2 \in L^\perp$, тогда $(y_1, x) = (y_2, x) = 0$ для любого вектора $x \in L$. Складывая эти равенства, получим, что $(y_1 + y_2, x) = 0$, $\forall x \in L$, т.е. $y_1 + y_2 \in L^\perp$. Аналогично если $(y, x) = 0$, $\forall x \in L$, то $(\alpha y, x) = 0$, т.е. $\alpha y \in L$. Отсюда на основании теоремы 18.1 следует, что $L^\perp$ – линейное подпространство. ■

Теорема 72.2. *Если L – линейное подпространство $E(U)$, то*

$$L \oplus L^\perp = E\,(U). \tag{72.1}$$

Доказательство. Утверждение очевидно, если L – тривиальное подпространство. Пусть L – нетривиальное подпространство. Возьмем $e_1, \ldots, e_k$ – ортонормированный базис L, $e_{k+1}, \ldots, e_n$ – ортонормированный базис $L^\perp$. Система векторов $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ ортонормирована и, следовательно, линейно независима (теорема 70.1). Покажем, что она образует базис всего пространства $E\,(U)$. Пусть это не так. Тогда существует вектор f пространства,

$$b = b_1' + \ldots + b_i' + \ldots + b_k', \tag{66.5}$$

которые отличаются хотя бы одним слагаемым $b_i \neq b_i'$. Если вычесть из первого равенства второе и разложить каждое слагаемое $b_j - b_j'$, $j = \overline{1,n}$, по базису L_j, получим нетривиальную линейную комбинацию базисных векторов пространств $L_1, \ldots, L_k$, равную нулевому вектору. Это противоречит линейной независимости совокупности базисов $L_1, \ldots, L_k$.

2 $\iff$ 3. Это следует из теоремы 65.2.

3 $\iff$ 4. Эти утверждения отличаются только терминологией.

1 $\Rightarrow$ 5. Очевидно.

5 $\Rightarrow$ 1. Пусть $L_1 + \ldots + L_k$ – не прямая сумма. Тогда существует вектор b из этой суммы, для которого имеют место два различных разложения (66.4) и (66.5). Вычитая (66.5) из (66.4), получим нетривиальное разложение нулевого вектора (66.3). Если его сложить с разложением вектора a, то получим еще одно разложение вектора a.

1 $\Rightarrow$ 6. Пусть система векторов $a_1, \ldots, a_k$, где $a_i \in L_i$, $a_i \neq \theta$, $i = \overline{1,k}$, линейно зависима. Тогда существуют числа $\alpha_1, \ldots, \alpha_k \in P$, одновременно не равные нулю (пусть, например, $\alpha_i \neq 0$) и такие, что $\alpha_1 a_1 + \ldots + \alpha_i a_i + \ldots + \alpha_k a_k = \theta$. Это равенство дает второе разложение нулевого вектора, отличное от тривиального (так как $\alpha_i a_i \neq \theta$), что противоречит утверждению 1.

6 $\Rightarrow$ 1. Пусть $L_1 + \ldots + L_k$ – не прямая сумма, тогда существует вектор b, для которого имеют место два разложения (66.4) и (66.5). Вычитая одно из другого, получим, что $a_{i_1} + a_{i_2} + \ldots + a_{i_q} = \theta$, где $a_{i_m} = b_{i_m} - b_{i_m}' \neq \theta$, $a_{i_m} \in L_{i_m}$, $m = \overline{1,q}$. Значит, векторы $a_{i_1}, a_{i_2}, \ldots, a_{i_q}$ линейно зависимы и, следовательно, любая система ненулевых векторов, взятых по одному из каждого L_i, $i = \overline{1,k}$, содержащая эти векторы, линейно зависима (в силу теоремы 14.3). Это противоречит утверждению 6.

4 $\iff$ 7. Это следует из теоремы 65.3. ■

Теорема 66.2. *Линейное пространство V является прямой суммой своих подпространств L_1 и L_2 тогда и только тогда, когда:*

1) $\dim V = \dim L_1 + \dim L_2$;
2) $L_1 \cap L_2 = \{\theta\}$.

Доказательство. Необходимость следует из теоремы 66.1 (утверждения 4,7).

Достаточность. Из условия 2 следует, что $L_1 + L_2$ – прямая сумма. Положим $L = L_1 \oplus L_2$. Согласно теореме 66.1 (утверждение 4) $\dim L = \dim L_1 + \dim L_2$. Отсюда в силу условия 1 вытекает, что $\dim L = \dim V$. Это означает, что $L = V$ (теорема 64.3), т.е. $V = L_1 \oplus L_2$. ■

Дополнительное подпространство. Пусть L – линейное подпространство пространства V. Подпространство L^δ называется ***дополнительным подпространством к*** L (рис. 1), если $L \oplus L^\delta = V$.

Очевидно, что при этом L – дополнительное подпространство к L^δ.

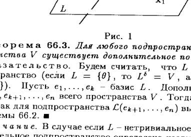


Рис. 1

Теорема 66.3. *Для любого подпространства L линейного пространства V существует дополнительное подпространство.*

Доказательство. Будем считать, что L – нетривиальное подпространство (если $L = \{\theta\}$, то $L^\delta = V$, а если $L = V$, то $L^\delta = \{\theta\}$). Пусть $e_1, \ldots, e_k$ – базис L. Дополним его до базиса $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ всего пространства V. Тогда $\mathcal{L}(e_{k+1}, \ldots, e_n) = L^\delta$, так как для подпространства $\mathcal{L}(e_{k+1}, \ldots, e_n)$ выполнены все условия теоремы 66.2. ■

Замечание. В случае если L – нетривиальное подпространство, дополнительное подпространство определено неоднозначно.

§67. Линейное многообразие

Линейное многообразие и связанные с ним понятия были определены в §18. Там же приведены некоторые элементарные утверждения, относящиеся к вещественному линейному пространству. Все они остаются в силе и в линейном пространстве над произвольным полем P (с очевидной заменой поля $\mathbb{R}$ на поле P).

Линейные многообразия $H_1 = x_1 + L_1$ и $H_2 = x_2 + L_2$ в линейном пространстве V называются ***параллельными***, если либо $L_1 \subset L_2$, либо $L_2 \subset L_1$ (рис. 1).

Теорема 67.1. *Если два линейных многообразия с непустым пересечением параллельны, то одно из них содержит другое.*

Доказательство. Пусть линейные аффинные многообразия $H_1 = x_1 + L_1$ и $H_2 = x_2 + L_2$ параллельны и пусть

$$L_1 \subset L_2. \tag{67.1}$$

По условию существует вектор $x_0 \in H_1 \cap H_2$. Так как вектором сдвига может быть любой вектор линейного многообразия (§18), то $H_1 = x_0 + L_1$, $H_2 = x_0 + L_2$. Отсюда с учетом (67.1) следует, что $H_1 \subset H_2$. ■

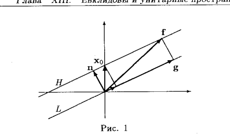


Рис. 1

Теорема 73.1. *Для любого линейного многообразия в евклидовом (унитарном) пространстве существует, и притом единственный, нормальный вектор.*

Доказательство. Рассмотрим линейное многообразие $H = x_0 + L$ (рис. 1). Все векторы из H, ортогональные L, находятся в $H \cap L^\perp$, но пересечение $H \cap L^\perp$ состоит ровно из одного вектора n, так как $L^\perp$ – дополнительное подпространство к L (теоремы 72.2 и 67.4). Этот вектор n и будет единственным нормальным вектором H. ■

Теорема 73.2. *Нормальный вектор линейного многообразия совпадает с перпендикуляром, опущенным из любого вектора линейного многообразия на направляющее подпространство.*

Доказательство. Пусть n – нормальный вектор линейного многообразия $H = x_0 + L$, тогда $H = n + L$ (§18). Следовательно, любой вектор $f \in H$ (рис. 1) может быть представлен в виде

$$f = n + g, \quad g \in L. \tag{73.1}$$

Так как $n \perp L$, то соотношение (73.1) совпадает с разложением вектора f на ортогональную проекцию g и перпендикуляр n. ■

Следствие. *Среди всех векторов линейного многообразия нормальный вектор имеет наименьшую длину.*

Уравнение гиперплоскости. Пусть $H = x_0 + L$ – гиперплоскость в $E\,(U)$, т.е. $\dim L = n - 1$, где $n = \dim E\,(\dim U)$. Тогда $L^\perp$ – одномерное подпространство и его базис состоит из одного вектора n.

Вектор $x \in H$ тогда и только тогда, когда разность $x - x_0 \in L$, т.е. когда

$$(x - x_0, n) = 0. \tag{73.2}$$

Таким образом, уравнению (73.2) удовлетворяют все векторы x гиперплоскости H, и только они. Уравнение (73.2) может быть записано в равносильной форме:

$$(x, n) = p, \tag{73.3}$$

где $p = (x_0, n)$ – фиксированное число для данной гиперплоскости.

Уравнению (73.3) можно придать координатную форму: если $(x_1, \ldots, x_n)$ и $(A_1, \ldots, A_n)$ – координаты векторов x и n в некотором ортонормированном базисе пространства, то уравнение (73.3) равносильно уравнению $A_1 x_1 + \ldots + A_n x_n = p$, которое в геометрических пространствах совпадает с общим уравнением прямой на плоскости и плоскости в пространстве.

§74. Расстояние в евклидовом (унитарном) пространстве

Множество M называется ***метрическим пространством***, если задано отображение

$$\rho : M \times M \to \mathbb{R},$$

которое каждой упорядоченной паре элементов $x, y \in M$ ставит в соответствие число $\rho(x, y) \in \mathbb{R}$ такое, что:

1) $\rho(x, y) \geq 0$, $\quad \forall x, y \in M$,
$\quad \rho(x, y) = 0 \iff x = y$;
2) $\rho(x, y) = \rho(y, x)$, $\quad \forall x, y \in M$;
3) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$, $\quad \forall x, y, z \in M$.

Число $\rho(x, y)$ называется ***расстоянием между*** x ***и*** y; отображение ρ – ***метрикой***, аксиомы 1–3 – ***аксиомами метрики (расстояния)***.

Расстоянием между множествами X и Y в метрическом пространстве называется число

$$\rho(X, Y) = \inf_{x \in X,\, y \in Y} \rho(x, y). \tag{74.1}$$

Теорема 74.1. *В евклидовом (унитарном) пространстве V правило*

$$\rho(x, y) = |x - y| \tag{74.2}$$

задает метрику.

Доказательство. Действительно, правило (74.2) определяет отображение $\rho : V \times V \to \mathbb{R}$, которое отвечает всем аксиомам метрики. Проверка аксиом тривиальна, отметим лишь одну из них: $\rho(x, z) = |x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z| = \rho(x, y) + \rho(y, z)$, $\forall x, y, z \in V$. ■

Итак, евклидово (унитарное) пространство является метрическим пространством относительно метрики (74.2). В дальнейшем, говоря об евклидовом (унитарном) пространстве как о метрическом, будем иметь в виду именно эту метрику.

Теорема 74.2 (о кратчайшем расстоянии). *Расстояние между вектором f и линейным подпространством L в евклидовом (унитарном) пространстве равно длине перпендикуляра, опущенного из вектора f на L.*

Доказательство. Пусть $f = g + h$, где $g \in L$, $h \in L^\perp$, и y – произвольный вектор из L. Тогда $\rho(f, y) = |f - y| = |(g + h) - y| =$

имеющейся на прямой, на плоскости и в пространстве. Оно удовлетворяет всем аксиомам скалярного произведения (теорема 22.2) и, следовательно, является скалярным произведением и в принятом выше смысле. Аксиоматическое определение скалярного произведения говорит о том, что геометрическое правило (22.1) не является единственным скалярным произведением в пространствах V_n, $n = 1, 2, 3$.

2. В арифметическом пространстве $\mathbb{R}^n$ скалярное произведение векторов $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ может быть введено следующим образом:

$$(x, y) = \sum_{i=1}^{n} x_i y_i\,, \tag{68.2}$$

а в арифметическом пространстве $\mathbb{C}^n$ –

$$(x, y) = \sum_{i=1}^{n} x_i \overline{y_i}\,. \tag{68.3}$$

Для (68.2), (68.3) аксиомы (68.1) легко проверяются непосредственно.

3. В функциональном пространстве $C[0, 1]$ скалярное произведение функций $f(x)$ и $g(x)$ может быть задано так:

$$(f, g) = \int_0^1 f(x) g(x) dx\,. \tag{68.4}$$

Справедливость аксиом (68.1) вытекает из свойств определенного интеграла [7].

Итак, операции (68.2)–(68.4) определены корректно и превращают пространства V_n, $\mathbb{R}^n$, $C[0, 1]$ в евклидовы пространства, а пространство $\mathbb{C}^n$ – в унитарное.

Из определения скалярного произведения вытекают следующие *простейшие свойства* этой операции.

1°. $(x, y + z) = (x, y) + (x, z)$, $\forall x, y, z \in E\ (U)$.

2°. $(x, \alpha x) = \overline{\alpha}(x, y)$, $\forall x, y \in E\ (U)$, $\forall \alpha \in \mathbb{R}$ (соответственно $\mathbb{C}$).

3°. $(\theta, x) = (x, \theta) = 0$, $\forall x \in E\ (U)$.

4°. $(x, y) = 0$ для любого вектора $y \in E\ (U)$ тогда и только тогда, когда $x = \theta$.

5°. Любое подпространство L евклидова (унитарного) пространства является евклидовым (соответственно унитарным) пространством.

Отметим, что скалярное произведение (x, y) векторов x, y линейно по первому аргументу, а в евклидовом пространстве – линейно по обоим аргументам (свойства 1°, 2°).

Теорема 68.1. *Для любых векторов $x, y \in E\ (U)$ имеет место неравенство*

$$|(x, y)|^2 \le (x, x)(y, y) \tag{68.5}$$

или, в другой форме,

$$\begin{vmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{vmatrix} \ge 0\,.$$

Доказательство. Будем считать, что $x \ne \theta$ (для $x = \theta$ неравенство (68.5), очевидно, обращается в равенство). Для любых векторов $x, y \in E\ (U)$ и любого числа $\alpha \in \mathbb{R}\ (\mathbb{C})$, согласно (68.1) и свойствам 1°–4°, имеем

$$0 \le (\alpha x - y, \alpha x - y) = (\alpha x, \alpha x) - (\alpha x, y) - (y, \alpha x) + (y, y) =$$
$$= |\alpha|^2 (x, x) - \alpha (x, y) - \overline{\alpha}(y, x) + (y, y)\,. \tag{68.6}$$

Так как $x \ne \theta$, то $(x, x) \ne 0$. Возьмем $\alpha = (y, x)/(x, x)$. Тогда из (68.6) имеем $0 \le \dfrac{|(y,x)|^2}{(x,x)^2}(x,x) - \dfrac{(y,x)(x,y)}{(x,x)} - \dfrac{(x,y)(y,x)}{(x,x)} + (y,y)$, откуда после очевидных преобразований получим (68.5). ■

Неравенство (68.5) называется *неравенством Коши–Буняковского.*

Замечание 3. В евклидовом пространстве неравенство Коши–Буняковского может быть записано в виде $(x, y)^2 \le (x, x)(y, y)$.

Теорема 68.2. *Неравенство Коши–Буняковского обращается в равенство тогда и только тогда, когда векторы x и y коллинеарны.*

Доказательство этой теоремы фактически содержится в доказательстве теоремы 68.1. Действительно, если $x = \theta$, то $x = 0y$. Если $x \ne \theta$, то неравенство (68.6) обращается в равенство тогда и только тогда, когда $0 = (\alpha x - y, \alpha x - y)$, т.е. когда $y = \alpha x$. ■

§69. Основные метрические понятия

В евклидовом и унитарном пространствах *длиной* вектора x называется арифметическое значение квадратного корня из его скалярного квадрата. Обозначение: $|x|$. Итак, $|x| = \sqrt{(x, x)}$.

Из аксиом скалярного произведения вытекают следующие факты.

1°. Любой вектор x евклидова (и унитарного) пространства имеет длину, при этом $|x| \ge 0$, $\forall x \in E\ (U)$ и $|x| = 0 \Leftrightarrow x = \theta$.

2°. $|\alpha x| = |\alpha||x|$, $\forall x \in E\ (U)$, $\forall \alpha \in \mathbb{R}$ (соответственно $\mathbb{C}$).

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

$$|(x, y)| \le |x| \cdot |y|\,. \tag{69.1}$$

Доказательство. Необходимость. Пусть в системе векторов $a_1, \ldots, a_r, \ldots, a_k$ подсистема $a_1, \ldots, a_r$ образует базу (здесь для упрощения записи в качестве подсистемы взяты первые r векторов, что, как будет видно из доказательства, не нарушает общности рассуждений). Тогда любая бо́льшая подсистема будет линейно зависимой в силу теоремы 16.2, так как любой ее вектор линейно выражается через базу $a_1, \ldots, a_r$. Таким образом, база образует максимальную линейно независимую подсистему системы векторов.

Достаточность. Пусть $a_1, \ldots, a_r$ – максимальная линейно независимая подсистема системы векторов $a_1, \ldots, a_r, \ldots, a_k$. Тогда для любого вектора a_i, $i = \overline{1, k}$, подсистема $a_1, \ldots, a_r, a_i$ линейно зависима (если $i \le r$, то как подсистема, содержащая два одинаковых вектора; если же $i > r$, то как подсистема из $r + 1 > r$ векторов). В силу теоремы 14.6 вектор a_i линейно выражается через $a_1, \ldots, a_r$. Таким образом, $a_1, \ldots, a_r$ – база. ■

Следствие 1. *Все базы одной системы векторов состоят из одинакового числа векторов, равного максимальному числу линейно независимых векторов системы.*

Число векторов базы называется *рангом системы векторов*. Очевидно, ранг системы векторов равен максимальному числу линейно независимых векторов системы. Обозначение: $\mathrm{rg}(a_1, \ldots, a_n)$.

Две системы векторов линейного пространства называются *эквивалентными*, если каждая из этих систем выражается через другую. Из определения следует, что *база системы векторов эквивалентна самой системе.*

Теорема 61.2. *Если система векторов $a_1, \ldots, a_k$ линейно выражается через $b_1, \ldots, b_m$, то $\mathrm{rg}(a_1, \ldots, a_k) \le \mathrm{rg}(b_1, \ldots, b_m)$.*

Доказательство. Пусть $a_1, \ldots, a_r$ и $b_1, \ldots, b_s$ – базы рассматриваемых систем. Из условия теоремы и транзитивности свойства "линейной выражаемости" следует, что база $a_1, \ldots, a_r$ первой системы линейно выражается через базу $b_1, \ldots, b_s$ второй. Тогда $r \le s$, так как в противном случае, если $r > s$, система $a_1, \ldots, a_r$ была бы линейно зависимой в силу теоремы 16.2. ■

Следствие 2. *Ранги эквивалентных систем совпадают.*

Следствие 3. *Эквивалентные линейно независимые системы векторов состоят из одинакового числа векторов.*

§62. Базис и размерность

Говорят, что *система векторов* $e_1, \ldots, e_n$ линейного пространства V *порождает пространство* V, если любой вектор $x \in V$ является линейной комбинацией $e_1, \ldots, e_n$.

Упорядоченная система векторов $e_1, \ldots, e_n$ линейного пространства V называется *базисом* V, если она линейно независима и порождает V. Очевидно, что это определение совпадает с определением

базиса, данным в §17 для вещественного линейного пространства.

Теорема 62.1. *Любые два базиса линейного пространства состоят из одинакового числа векторов.*

Это утверждение вытекает из эквивалентности двух базисов линейного пространства и следствия 2 (§61). ■

Итак, число векторов базиса не зависит от самого базиса и однозначно определяется самим пространством. Ответ на вопрос, что представляет собой это число, дается теоремой 17.1, справедливой и в пространстве над произвольным полем.

Напомним, что число векторов базиса линейного пространства V называется *размерностью пространства* V и обозначается символом $\dim V$. Размерность нулевого пространства по определению считается равной нулю. Из теоремы 17.1 следует, что размерность линейного пространства совпадает с максимальным числом линейно независимых векторов этого пространства. Линейное пространство размерности n, где $n \in \mathbb{N}$, называется *n-мерным*. Нулевое пространство и n-мерные пространства называются *конечномерными*.

Линейное пространство называется *бесконечномерным*, если для любого числа $k \in \mathbb{N}$ в пространстве существует k линейно независимых векторов. Пример бесконечномерного пространства дан в §17 (пример 5). В курсе линейной алгебры изучаются лишь конечномерные линейные пространства.

Теорема 62.2 (о неполном базисе). *В n-мерном пространстве любую линейно независимую систему из k, где $k < n$, векторов можно дополнить до базиса.*

Доказательство. Пусть $e_1, \ldots, e_k$ – линейно независимая система векторов пространства V. Так как $k < n$, то в силу теоремы 17.1 система векторов $e_1, \ldots, e_k$ не является базисом V и, следовательно, не порождает всего пространства V. Пусть вектор $e_{k+1} \in V$ не является линейной комбинацией $e_1, \ldots, e_k$, тогда система векторов $e_1, \ldots, e_k, e_{k+1}$ линейно независима (в силу теоремы 14.6). Если $k+1 = n$, то эта система векторов образует базис V; если же $k+1 < n$, то аналогичным образом построим линейно независимую систему из $k + 2$ векторов. За $n - k$ таких шагов мы построим искомый базис $e_1, \ldots, e_k, \ldots, e_n$. ■

Точно так же, как и в §17, определяются координаты вектора в базисе и матрица перехода от одного базиса к другому. При этом все факты, связанные с этими понятиями, остаются в силе и в общем случае, так как их доказательства опираются на те свойства вещественных чисел, которые имеют место в любом поле. Напомним эти факты.

1°. *Разложение вектора по базису единственно.*

2°. *Координаты вектора обладают свойством линейности.*

3°. *При переходе от базиса e к базису $f = eQ$ координаты вектора x изменяются по следующему закону: $x_e = Q x_f$.*

Примеры. 1. В арифметических пространствах P^n, $\mathbb{R}^n$, $\mathbb{C}^n$ еди-

тогда и только тогда, когда e – ортонормированный базис.

Доказательство. Необходимость. Если скалярное произведение вычисляется согласно (70.7) для любой пары векторов, то это же верно для пары базисных векторов e_i и e_j, координаты которых известны. Применив правило (70.7) для вычисления скалярного произведения векторов e_i и e_j, получим требуемые равенства (70.5).

Достаточность. Если e – ортонормированный базис и $x = \sum_{i=1}^{n} x_i e_i$, $y = \sum_{i=1}^{n} y_i e_i$, то в силу свойства линейности скалярного произведения имеем

$$(x, y) = (\sum_{i=1}^{n} x_i e_i\,,\ \sum_{j=1}^{n} y_j e_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \overline{y_j}(e_i, e_j) = \sum_{i=1}^{n} x_i \overline{y_i}\,. \ ■$$

Замечание. В евклидовом пространстве черта в равенстве (70.7) может быть опущена: $(x, y) = \sum_{i=1}^{n} x_i y_i$.

До сих пор все базисы линейного пространства были равноправны. С помощью любого базиса можно сводить операции сложения векторов и умножения вектора на число к операциям над числами. В евклидовом и унитарном пространствах теорема 70.3 отводит ортонормированному базису особую роль: с помощью такого базиса и третья операция – скалярное произведение векторов – сводится к операциям над числами. Возникает естественный вопрос, всегда ли существует ортонормированный базис.

Теорема 70.4. *В конечномерном евклидовом (унитарном) пространстве существует ортонормированный базис.*

Доказательство. Пусть $\dim V = n$. Используем индукцию по n. При $n = 1$ утверждение очевидно: достаточно взять любой вектор $f \ne \theta$ и положить $e_1 = f/|f|$.

Пусть в любом $(n-1)$-мерном евклидовом (унитарном) пространстве существует ортонормированный базис; покажем, что ортонормированный базис существует и в n-мерном пространстве $E\ (U)$. Пусть $f_1, \ldots, f_n$ – базис $E\ (U)$. Линейная оболочка $\mathcal{L}(f_1, \ldots, f_{n-1})$ является $(n-1)$-мерным пространством, и в нем по индуктивному предположению существует ортонормированный базис $e_1, \ldots, e_{n-1}$. Так как $f_n \notin \mathcal{L}(f_1, \ldots, f_{n-1}) = \mathcal{L}(e_1, \ldots, e_{n-1})$, то вектор $g_n = f_n - \alpha_1 e_1 - \alpha_2 e_2 - \ldots - \alpha_{n-1} e_{n-1}$ отличен от нулевого вектора при любых $\alpha_i \in \mathbb{R}(\mathbb{C})$. Выберем коэффициенты α_i, $i = \overline{1, n-1}$, из условия ортогональности вектора g_n всем векторам $e_1, \ldots, e_{n-1}$: $0 = (g_n, e_i) = (f_n, e_i) - \alpha_i$, $i = \overline{1, n-1}$, или $\alpha_i = (f_n, e_i)$, $i = \overline{1, n-1}$. Тогда, положив $e_n = g_n/|g_n|$, получим ортонормированный базис $e_1, \ldots, e_n$ пространства $E\ (U)$. ■

Процесс ортогонализации. Доказательство теоремы 70.4, по существу, представляет собой алгоритм последовательного построения ортонормированного базиса по заданному базису $f_1, \ldots, f_n$.

Первый шаг. Полагая $g_1 = f_1$, находим $e_1 = g_1/|g_1|$.

k-й шаг ($k \ge 2$). Полагаем

$$g_k = f_k - \alpha_1 e_1 - \alpha_2 e_2 - \ldots - \alpha_{k-1} e_{k-1}\,, \tag{70.8}$$

где $\alpha_i = (f_k, e_i)$, $i = \overline{1, k-1}$, и находим $e_k = g_k/|g_k|$.

Через n шагов получим ортонормированный базис $e_1, \ldots, e_n$ пространства. Описанный алгоритм называется *процессом ортогонализации Грама–Шмидта* и состоит в следующем.

Ортогональная (унитарная) матрица. Матрица $U \in \mathbb{C}^{n\times n}$ называется *унитарной*, если

$$UU^H = U^H U = I\,, \tag{70.9}$$

матрица $Q \in \mathbb{R}^{n\times n}$ называется *ортогональной*, если

$$QQ^T = Q^T Q = I\,. \tag{70.10}$$

Напомним, что в §24 уже рассматривались ортогональные матрицы в связи с ортонормированным базисом геометрических пространств. В произвольном евклидовом (унитарном) пространстве имеет место то же утверждение.

Теорема 70.5. *Матрица ортогональна (унитарна) тогда и только тогда, когда она является матрицей перехода от одного ортонормированного базиса к другому ортонормированному базису евклидова (унитарного) пространства.*

Доказательство повторяет доказательство теоремы 24.1 с учетом соотношений (70.5), (70.7), (70.9), (70.10). ■

§71. Матрица Грама

Матрицей Грама системы векторов $a_1, \ldots, a_k$ евклидова (унитарного) пространства называется матрица

$$G(a_1, \ldots, a_k) = \begin{bmatrix} (a_1, a_1) & (a_2, a_1) & \ldots & (a_k, a_1) \\ \ldots & \ldots & \ldots & \ldots \\ (a_1, a_k) & (a_2, a_k) & \ldots & (a_k, a_k) \end{bmatrix}. \tag{71.1}$$

Определитель матрицы Грама называется *определителем Грама.*

Теорема 71.1. *Система векторов $a_1, \ldots, a_k$ евклидова (унитарного) пространства линейно зависима тогда и только тогда, когда* $\det G(a_1, \ldots, a_k) = 0$.

Доказательство. Необходимость. Пусть $a_1, \ldots, a_k$ – линейно зависимая система векторов. Последовательно умножая нетривиальную линейную комбинацию

$$\alpha_1 a_1 + \ldots + \alpha_k a_k = \theta \tag{71.2}$$

скалярно на векторы $a_1, \ldots, a_k$, получим однородную систему уравнений относительно неизвестных $\alpha_1, \ldots, \alpha_k$:

$$\begin{cases} \alpha_1(a_1, a_1) + \ \ldots\ + \alpha_k(a_k, a_1) = 0, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ \alpha_k(a_1, a_k) + \ \ldots\ + \alpha_k(a_k, a_k) = 0 \end{cases} \tag{71.3}$$

ничные векторы (14.6) образуют базис (§17, пример 2). Этот базис принято называть *естественным*. Координаты любого вектора $a = (a_1, \ldots, a_n)$ в естественном базисе совпадают с компонентами $a_1, \ldots, a_n$ этого вектора.

2. В арифметическом пространстве $\mathbb{C}^n_{\mathbb{R}}$ векторы (14.6), очевидно, не могут быть базисом, так как умножением этих векторов только на вещественные числа нельзя получить любой комплексный арифметический вектор $a = (\alpha_1 + i\beta_1, \ldots, \alpha_n + i\beta_n)$, где $\alpha_k, \beta_k \in \mathbb{R}$, $k = \overline{1, n}$. Добавим к векторам (14.6) еще n векторов:

$$f_1 = (i, 0, \ldots, 0),\ f_2 = (0, i, \ldots, 0),\ \ldots,\ f_n = (0, 0, \ldots, i).$$

Покажем, что система векторов $e_1, \ldots, e_n, f_1, \ldots, f_n$ образует базис $\mathbb{C}^n_{\mathbb{R}}$. Действительно, линейная комбинация этих векторов с вещественными коэффициентами α_k, β_k имеет вид

$$\sum_{i=1}^{n} \alpha_k e_k + \sum_{i=1}^{n} \beta_k f_k = (\alpha_1 + i\beta_1, \ldots, \alpha_n + i\beta_n) \tag{62.1}$$

и равна нулевому вектору $\theta = (0, \ldots, 0)$ тогда и только тогда, когда $\alpha_k + i\beta_k = 0$, $k = \overline{1, n}$, т.е. $\alpha_k = \beta_k = 0$, $k = \overline{1, n}$. Следовательно, векторы $e_1, \ldots, e_n, f_1, \ldots, f_n$ линейно независимы. Они порождают все пространство $\mathbb{C}^n_{\mathbb{R}}$, так как любой вектор $(\alpha_1 + i\beta_1, \ldots, \alpha_n + i\beta_n)$ из $\mathbb{C}^n_{\mathbb{R}}$ согласно (62.1) является линейной комбинацией этих векторов с вещественными коэффициентами $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$.

Итак, $\dim \mathbb{C}^n_{\mathbb{R}} = 2n$.

§63. Изоморфизм линейных пространств

Два линейных пространства V_1 и V_2 над общим полем P называются *изоморфными*, если существует биективное отображение $\varphi : V_1 \to V_2$, которое сохраняет законы композиции, т.е. если для любых векторов $x, y \in V_1$ и любого числа $\alpha \in P$

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

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

Обозначение: $V_1 \cong V_2$. Само отображение φ называется *изоморфизмом линейных пространств.*

Примеры. 1. Геометрические пространства V_1, V_2, V_3 векторов на прямой, на плоскости и в пространстве изоморфны арифметическим пространствам $\mathbb{R}^1$, $\mathbb{R}^2$ и $\mathbb{R}^3$ соответственно. Действительно, поставим в соответствие каждому вектору $\mathbf{x} \in V_n$ набор его координат в каком-либо базисе e, т.е. арифметический вектор $x_e \in \mathbb{R}^n$, где $n = 1, 2, 3$. Это соответствие взаимно однозначно и сохраняет законы композиции, так как координаты вектора обладают свойством линейности (§17).

2. $V_2 \cong C_{\mathbb{R}}$.

3. $P^{m\times n} \cong P^{mn}$, и в частности $\mathbb{R}^{m\times n} \cong \mathbb{R}^{mn}$.

Отметим простейшие свойства изоморфных пространств.

1°. *Отношение изоморфизма есть отношение эквивалентности на множестве всех линейных пространств над полем P.*

2°. *В изоморфных пространствах*

а) образ (и прообраз) линейной комбинации векторов есть линейная комбинация образов (прообразов) с теми же коэффициентами;

б) образ (и прообраз) нулевого вектора есть нулевой вектор;

в) образ (и прообраз) линейно независимой системы векторов образует линейно независимую систему;

г) образ (и прообраз) базиса есть базис.

Доказательства этих свойств опираются лишь на определение и элементарные свойства рассматриваемых объектов. Мы предоставляем их читателю.

Несмотря на простоту формулировок, уже эти свойства говорят о том, что изоморфные линейные пространства, даже самой различной природы, с точки зрения свойств линейного пространства неразличимы. Это позволяет многие свойства линейных пространств изучать на линейных пространствах простой структуры. К этой возможности мы будем неоднократно прибегать, проводя все вычисления в арифметических пространствах, а "рисуя" – в геометрических.

Теорема 63.1 (критерий изоморфизма). *Два линейных пространства над общим полем изоморфны тогда и только тогда, когда их размерности совпадают.*

Доказательство. Необходимость вытекает из свойства "г" изоморфных пространств.

Достаточность. Пусть V_1 и V_2 – линейные пространства над полем P и пусть $\dim V_1 = \dim V_2 = n$. Выберем в этих пространствах по базису: пусть $e_1, \ldots, e_n$ – базис V_1, $f_1, \ldots, f_n$ – базис V_2. Построим отображение $\varphi : V_1 \to V_2$, поставив в соответствие каждому вектору $x = \sum_{i=1}^{n} \alpha_i e_i \in V_1$ вектор $y = \sum_{i=1}^{n} \alpha_i f_i \in V_2$ (т.е. вектор, который имеет те же координаты, что и вектор x). В силу единственности разложения вектора по базису отображение φ биективно. При этом φ – изоморфизм, так как координаты вектора обладают свойством линейности. ■

Следствие. Любое n-мерное вещественное пространство изоморфно арифметическому пространству $\mathbb{R}^n$, а любое n-мерное комплексное пространство – арифметическому пространству $\mathbb{C}^n$.

§64. Линейные подпространства

Напомним определение линейного подпространства, данное в §18 (оно остается неизменным и в случае линейного пространства над произвольным полем). Подмножество L линейного пространства V над полем P называется *линейным подпространством пространства V*, если оно само является линейным пространством относительно законов композиции в V. Другое определение линейного подпространства, эквивалентное этому, устанавливается известной теоремой 18.1, которая имеет место и в общем случае.

который не является линейной комбинацией $e_1, \ldots, e_n$. Система векторов $e_1, \ldots, e_n, f$ линейно независима, и применение к ней процесса ортогонализации приводит к вектору e_{n+1}, который ортогонален $e_1, \ldots, e_n$ и, значит, $e_{n+1} \in L^\perp$. С другой стороны, $e_{n+1} \perp L^\perp$, так как e_{n+1} ортогонален $e_{k+1}, \ldots, e_n$. Следовательно, $e_{n+1} = \theta$. Отсюда согласно (70.8) вытекает линейная зависимость $e_1, \ldots, e_n, f$, что противоречит допущению. Таким образом, система векторов $e_1, \ldots, e_n$ является базисом $E\ (U)$ и $\dim L + \dim L^\perp = \dim E\ (\dim U)$. Так как $L \cap L^\perp = \{\theta\}$, то согласно теореме 66.2 получаем (72.1). ■

Следствие. Если L – линейное подпространство $E\ (U)$, то для любого вектора $f \in E\ (U)$ существует, и притом единственное, разложение

$$f = g + h\,, \tag{72.2}$$

где $g \in L$, $h \perp L$.

Вектор g в разложении (72.2) называется *ортогональной проекцией вектора f на подпространство L*, а вектор h – *ортогональной составляющей вектора f.*

Задачу разложения (72.2) вектора на ортогональную проекцию и ортогональную составляющую называют *задачей о перпендикуляре.* Этот термин заимствован из геометрии. С разложением (72.2) геометрического вектора мы уже встречались (рис. 1): чтобы получить это разложение, достаточно опустить перпендикуляр из конца вектора f на плоскость L. Имея в виду эту аналогию, ортогональную составляющую h в разложении (72.2) называют *перпендикуляром, опущенным из вектора f на подпространство L*, а сам вектор f – *наклонной к подпространству L.*

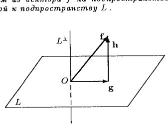


Рис. 1

Аналогия с геометрическими векторами состоит не только в названии. Отметим несколько тех свойств составляющих g и h в разложении (72.2), которые имеют место и в геометрии: так как $|f|^2 = (g + h, g + h) = (g, g) + (h, h)$, то

$$|f|^2 = |g|^2 + |h|^2\,, \tag{72.3}$$

$$|h| \le |f|\,.$$

Последнее неравенство свидетельствует о том, что *длина перпендикуляра не превосходит длины наклонной.* Равенство (72.3) называется *теоремой Пифагора в евклидовом (унитарном) пространстве.*

Решение задачи о перпендикуляре. Из теоремы 72.2 следует, что задача о перпендикуляре имеет, и притом единственное, решение. Укажем один из способов построения этого решения.

Пусть $L = \mathcal{L}(a_1, \ldots, a_k)$. Разложение (72.2) равносильно задаче построения векторов g и h таких, что

$$\begin{cases} f = g + h, \\ g \in L, \\ h \perp L \end{cases} \Leftrightarrow \begin{cases} g = \sum_{j=1}^{k} x_j a_j, \\ f - g \perp L, \\ h = f - g \end{cases} \Leftrightarrow \begin{cases} (f - g, a_i) = 0\,,\ i = \overline{1, k}, \\ g = \sum_{j=1}^{k} x_j a_j, \\ h = f - g \end{cases} \Leftrightarrow$$

$$\Leftrightarrow \begin{cases} (g, a_i) = (f, a_i)\,,\ i = \overline{1, k}, \\ g = \sum_{j=1}^{k} x_j a_j, \\ h = f - g \end{cases} \Leftrightarrow \begin{cases} (\sum_{j=1}^{k} x_j a_j, a_i) = (f, a_i)\,,\ i = \overline{1, k}, \\ g = \sum_{j=1}^{k} x_j a_j, \\ h = f - g. \end{cases}$$

Итак, задача о перпендикуляре сводится к поиску чисел $x_1, \ldots, x_k$, для которых

$$(\sum_{j=1}^{k} x_j a_j, a_i) = (f, a_i)\,,\ i = \overline{1, k}\,. \tag{72.4}$$

Сравнение этих соотношений с (71.4) и (71.3) говорит о том, что числа $x_1, \ldots, x_k$ являются решением системы уравнений с матрицей Грама $G(a_1, \ldots, a_k)$ и правой частью (f, a_i), $i = \overline{1, k}$. Эта система всегда имеет решение, так как она равносильна задаче о перпендикуляре.

Отметим возможные при этом варианты.

1°. Если $a_1, \ldots, a_k$ – ортонормированный базис L, то $G(a_1, \ldots, a_k) = I$ и система (72.4) имеет единственное решение $x_i = (f, a_i)$, $i = \overline{1, k}$, при этом $g = \sum_{i=1}^{k} x_i a_i$, $h = f - g$.

2°. Если $a_1, \ldots, a_k$ линейно независимы, то $\det G(a_1, \ldots, a_k) \ne 0$ и система (72.4) имеет единственное решение; решив систему (72.4), найдем $g = \sum_{i=1}^{k} x_i a_i$, $h = f - g$.

3°. Если $a_1, \ldots, a_k$ линейно зависимы, то $\det G(a_1, \ldots, a_k) = 0$ и система (72.4) имеет бесконечно много решений, которые соответствуют бесконечному числу разложений вектора g по линейно зависимой системе векторов $a_1, \ldots, a_k$.

§73. Линейное многообразие в евклидовом (унитарном) пространстве

Пусть $H = x_0 + L$ – линейное многообразие в евклидовом (унитарном) пространстве. Вектор $n \in H$, ортогональный L, называется *нормальным вектором линейного многообразия H* (рис. 1).

Следствие. Размерность суммы линейных подпространств равна рангу совокупности слагаемых подпространств:

$$\dim \sum_{i=1}^{k} L_i = \mathrm{rg}(e_1, \ldots, e_m, f_1, \ldots, f_s, \ldots, g_1, \ldots, g_t).$$

Это утверждение вытекает из теоремы 65.2 (с учетом 64.3).

Теорема 65.3. *Для любых двух линейных подпространств L_1 и L_2 линейного пространства V имеет место соотношение*

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim L_1 \cap L_2\,. \tag{65.2}$$

Доказательство. Рассмотрим сначала случай, когда $L_1 \cap L_2 \ne \{\theta\}$. Выполним следующее построение. Пусть $f_1, \ldots, f_s$ – базис $L_1 \cap L_2$. Так как $L_1 \cap L_2 \subset L_1$, $L_1 \cap L_2 \subset L_2$, то согласно теореме о неполном базисе векторы $f_1, \ldots, f_s$ можно дополнить как до базиса L_1, так и до базиса L_2. Пусть $e_1, \ldots, e_m, f_1, \ldots, f_s$ – базис L_1, а $f_1, \ldots, f_s, g_1, \ldots, g_t$ – базис L_2. Покажем, что система векторов

$$e_1, \ldots, e_m, f_1, \ldots, f_s, g_1, \ldots, g_t \tag{65.3}$$

образует базис $L_1 + L_2$. Действительно, она порождает $L_1 + L_2$ в силу теоремы 65.2 и, кроме того, она линейно независима, так как если

$$\sum_{i=1}^{m} \alpha_i e_i + \sum_{i=1}^{s} \beta_i f_i + \sum_{i=1}^{t} \gamma_i g_i = \theta\,, \tag{65.4}$$

то

$$g = \sum_{i=1}^{t} \gamma_i g_i = -\sum_{i=1}^{m} \alpha_i e_i - \sum_{i=1}^{s} \beta_i f_i. \tag{65.5}$$

Отсюда следует, что вектор g является вектором как L_1, так и L_2, т.е. $g \in L_1 \cap L_2$ и, следовательно, вектор g является линейной комбинацией векторов $f_1, \ldots, f_s$. Из единственности разложения вектора g по линейно независимой системе векторов $e_1, \ldots, e_m, f_1, \ldots, f_s$ следует, что в (65.5) $\alpha_1 = \ldots = \alpha_m = 0$. Тогда в (65.4) в силу линейной независимости $f_1, \ldots, f_s, g_1, \ldots, g_t$ получим, что $\beta_1 = \ldots = \beta_s = 0$, $\gamma_1 = \ldots = \gamma_t = 0$. Таким образом, только тривиальная линейная комбинация векторов (65.3) равна нулевому вектору. Итак, векторы (65.3) образуют базис $L_1 + L_2$ и, следовательно,

$$\dim(L_1 + L_2) = m + s + t\,. \tag{65.6}$$

С другой стороны, согласно схеме построения векторов (65.3), $\dim L_1 + \dim L_2 - \dim L_1 \cap L_2 = (m + s) + (s + t) - s = m + s + t$. Сравнение этого соотношения с (65.6) дает (65.2). В случае когда $L_1 \cap L_2 = \{\theta\}$, доказательство проводится аналогично, только в нем не участвуют векторы $f_1, \ldots, f_s$. ■

§66. Прямая сумма подпространств

Сумма подпространств линейного пространства называется *прямой суммой*, если разложение каждого вектора в ней по слагаемым подпространствам единственно. Обозначение: $L_1 \oplus \ldots \oplus L_k$.

Теорема 66.1 (критерии прямой суммы). *Для подпространств $L_1, \ldots, L_k$ конечномерного линейного пространства V следующие утверждения равносильны:*

1) сумма подпространств $L_1, \ldots, L_k$ – прямая;

2) совокупность базисов подпространств $L_1, \ldots, L_k$ линейно независима;

3) совокупность базисов подпространств $L_1, \ldots, L_k$ образует базис суммы $\sum_{i=1}^{k} L_i$;

4) $\dim \sum_{i=1}^{k} L_i = \sum_{i=1}^{k} \dim L_i$;

5) существует вектор $a \in \sum_{i=1}^{k} L_i$, для которого разложение по подпространствам $L_1, \ldots, L_k$ единственно;

6) произвольная система ненулевых векторов $a_1, \ldots, a_k$, взятых по одному из каждого подпространства L_i, $i = \overline{1, k}$, линейно независима;

7) $L_1 \cap L_2 = \{\theta\}$ (для $k = 2$).

Доказательство. 1 ⇒ 2. Пусть совокупность $e_1, \ldots, e_m$, $f_1, \ldots, f_s, \ldots, g_1, \ldots, g_t$ базисов подпространств $L_1, \ldots, L_k$ линейно зависима и

$$\sum_{i=1}^{m} \alpha_i e_i + \sum_{i=1}^{s} \beta_i f_i + \ldots + \sum_{i=1}^{t} \gamma_i g_i = \theta\,, \tag{66.1}$$

где

$$\sum_{i=1}^{m} \alpha_i^2 + \sum_{i=1}^{s} \beta_i^2 + \ldots + \sum_{i=1}^{t} \gamma_i^2 \ne 0\,. \tag{66.2}$$

Положим

$$x_1 = \sum_{i=1}^{m} \alpha_i e_i\,,\quad x_2 = \sum_{i=1}^{s} \beta_i f_i\,,\quad \ldots,\quad x_k = \sum_{i=1}^{t} \gamma_i g_i\,.$$

Отметим, что $x_i \in L_i$, $i = \overline{1, k}$, причем среди $x_1, \ldots, x_k$ в силу (66.2) и линейной независимости векторов каждого базиса существует вектор $x_i \ne \theta$. Тогда соотношение (66.1) может быть записано в виде

$$\theta = x_1 + \ldots + x_i + \ldots + x_k\,,\ x_i \ne \theta\,. \tag{66.3}$$

Это дает второе разложение нулевого вектора θ (после известного: $\theta = \theta + \ldots + \theta + \ldots + \theta$) по подпространствам $L_1, \ldots, L_i, \ldots, L_k$.

2 ⇒ 1. Пусть $L_1, \ldots, L_k$ не прямая сумма. Тогда существует вектор b из этой суммы, который имеет два разложения по подпространствам $L_1, \ldots, L_k$:

$$b = b_1 + \ldots + b_i + \ldots + b_k\,, \tag{66.4}$$

$|h + (g - y)| = \{$ в силу (72.3) $\} = |h|^2 + |g - y|^2$. Отсюда следует, что $\rho(f, y) \ge |h|$, $\forall y \in L$ и $\rho(f, y) = |h|$, если $y = g$. Это означает, что $|h| = \inf_{y\in L} \rho(f, y) = \rho(f, L)$. ■

Эта теорема может быть сформулирована и в других терминах, а именно:

1) *расстояние между вектором f и подпространством L равно расстоянию между вектором f и его ортогональной проекцией на L;*

2) *среди всех векторов подпространства L ближе всего к вектору f расположена его ортогональная проекция на L.*

Теорема 74.3. *Пусть $H = x_0 + L$ – линейное аффинное многообразие в евклидовом (унитарном) пространстве. Тогда*

$$\rho(f, H) = \rho(f - x_0, L).$$

Это утверждение следует из теоремы 74.2 и того, что для любого вектора $z = x_0 + y \in H$: $\rho(f, z) = |f - z| = |(f - x_0) - y| = \rho(f - x_0, y)$, где $y \in L$. ■

Теорема 74.4. *Пусть $H_1 = x_1 + L_1$ и $H_2 = x_2 + L_2$ – линейные аффинные многообразия в евклидовом (унитарном) пространстве. Тогда*

$$\rho(H_1, H_2) = \rho(x_1 - x_2, L_1 + L_2).$$

Это утверждение следует из того, что для любых векторов $z_1 = x_1 + y_1 \in H_1$ и $z_2 = x_2 + y_2 \in H_2$: $\rho(z_1, z_2) = |z_1 - z_2| = |(x_1 - x_2) - (y_1 - y_2)| = \rho(x_1 - x_2, y)$, где $y = y_1 - y_2 \in (L_1 + L_2)$. ■

§75. Изометрия

Евклидовы пространства E_1 и E_2 называются *изоморфными*, если существует биективное отображение $\varphi : E_1 \to E_2$, которое сохраняет законы композиции и скалярное произведение, т.е. если:

1) $\varphi(x + y) = \varphi(x) + \varphi(y)$, $\forall x, y \in E_1$;

2) $\varphi(\alpha x) = \alpha\varphi(x)$, $\forall x \in E_1$, $\forall \alpha \in \mathbb{R}$;

3) $(\varphi(x), \varphi(y)) = (x, y)$, $\forall x, y \in E_1$.

Само отображение φ при этом называется *изоморфизмом евклидовых пространств* или *изометрией.*

Точно так же определяется изоморфизм унитарных пространств U_1 и U_2 (с очевидным отличием в свойстве 2: $\alpha \in \mathbb{C}$). Из определения следует, что изоморфные евклидовы (унитарные) пространства изоморфны как линейные пространства.

Теорема 75.1. *Два евклидовых (унитарных) пространства изоморфны тогда и только тогда, когда равны их размерности.*

Доказательство. Необходимость вытекает из изоморфизма евклидовых (унитарных) пространств как линейных пространств.

Достаточность. Пусть V_1 и V_2 – оба евклидовы (оба унитарные) пространства и пусть $\dim V_1 = \dim V_2 = n$. Выберем в V_1 и V_2 ортонормированные базисы $e_1, \ldots, e_n$ и $e'_1, \ldots, e'_n$ и построим отображение $\varphi : V_1 \to V_2$, положив для каждого вектора $x = \sum_{i=1}^{n} x_i e_i \in V_1$ вектор $\varphi(x) = \sum_{i=1}^{n} x_i e'_i$. Из доказательства теоремы 63.1 следует, что отображение φ – изоморфизм линейных пространств V_1 и V_2. Оно сохраняет скалярное произведение, так как если $x = \sum_{i=1}^{n} x_i e_i$, $y = \sum_{i=1}^{n} y_i e'_i$, то согласно (70.7)

$$(x, y) = \sum_{i=1}^{n} x_i \overline{y_i}\,,\quad (\varphi(x), \varphi(y)) = \sum_{i=1}^{n} x_i \overline{y_i}. \ ■$$

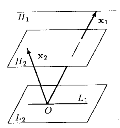


Рис. 1

Следствие 1. *Если линейные многообразия параллельны, то либо они не пересекаются, либо одно из них содержится в другом.*

Теорема 67.2. *Непустое пересечение линейных многообразий $H_1 = x_1 + L_1$ и $H_2 = x_2 + L_2$ является линейным многообразием с направляющим подпространством $L_1 \cap L_2$.*

Доказательство. По условию существует вектор $x_0 \in H_1 \cap H_2$, следовательно, $H_1 = x_0 + L_1$, $H_2 = x_0 + L_2$. Тогда $H_1 \cap H_2 = x_0 + L_1 \cap L_2$, так как для этих множеств имеет место двустороннее вложение. ■

Теорема 67.3. *Всякое k-мерное линейное многообразие в n-мерном пространстве можно задать в виде пересечения $n - k$ гиперплоскостей.*

Доказательство. Пусть $H = x_0 + L$ – k-мерное линейное многообразие и $e_1, \ldots, e_k$ – базис L. Дополним его до базиса $e_1, \ldots, e_k$, $e_{k+1}, \ldots, e_n$ пространства V. В качестве искомых гиперплоскостей можно взять линейные многообразия $H_i = x_0 + L_i$, $i = \overline{1, n-k}$, где

$$L_1 = \mathcal{L}(e_1, \ldots, e_k, e_{k+2}, e_{k+3}, \ldots, e_n),$$
$$L_2 = \mathcal{L}(e_1, \ldots, e_k, e_{k+1}, e_{k+3}, \ldots, e_n),$$
$$\ldots\ \ldots\ \ldots\ \ldots$$
$$L_{n-k} = \mathcal{L}(e_1, \ldots, e_k, e_{k+1}, e_{k+2}, \ldots, e_{n-1}),$$

т.е. L_i – линейная оболочка всех векторов базиса V, кроме e_{k+i}.

Очевидно, что $L_1 \cap L_2 \cap \ldots \cap L_{n-k} = \mathcal{L}(e_1, \ldots, e_k) = L$. В силу теоремы 67.2 отсюда следует, что $H = H_1 \cap H_2 \cap \ldots \cap H_{n-k}$. Так как $\dim L_i = n - 1$, $i = \overline{1, n-k}$, то все H_i – гиперплоскости. ■

Теорема 67.4. *Пересечение линейного многообразия $H = x_0 + L$ с любым подпространством, дополнительным к L, состоит ровно из одного вектора.*

Доказательство. Пусть L^δ – дополнительное подпространство к L. Так как $L \oplus L^\delta = V$, то для вектора x_0 имеет место разложение $x_0 = y + z$, где $y \in L$, $z \in L^\delta$. Тогда $z = x_0 - y \in H$, ибо $-y \in L$. Следовательно, z – общий вектор L^δ и H, т.е. $z \in L^\delta \cap H$.

Покажем, что z – единственный вектор пересечения $L^\delta \cap H$. Пусть $z' \in L^\delta \cap H$, тогда так как $z \in H$, $z' \in H$, то (§18) $z - z' \in L$, а так как $z \in L^\delta$, $z' \in L^\delta$, то $z - z' \in L^\delta$. Из того, что $L \cap L^\delta = \{\theta\}$, следует, что $z - z' = \theta$, т.е. $z = z'$. ■

Фактор-пространство. Пусть V – линейное пространство над полем P, L – некоторое его линейное подпространство. Рассмотрим множество $V|L$ всех линейных многообразий с направляющим подпространством L: $V|L = \{H = x + L \,|\, x \in V\}$. Введем на этом множестве операции сложения и умножения на число из поля P.

Определим сумму многообразий $H_1 = x_1 + L$, $H_2 = x_2 + L$ по следующему правилу:

$$H_1 + H_2 = (x_1 + x_2) + L\,. \tag{67.2}$$

Это определение корректно, так как результат не зависит от выбора векторов сдвига многообразий H_1, H_2. Фактически эта операция совпадает с известной операцией сложения смежных классов по нормальному делителю, т.е. с алгебраической операцией в фактор-группе (§40). Таким образом, правило (67.3) определяет внутренний закон композиции на $V|L$ (рис. 2).

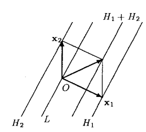


Рис. 2

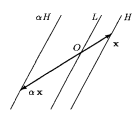


Рис. 3

Определим умножение многообразия $H = x_0 + L$ на число $\alpha \in P$ по следующему правилу:

$$\alpha H = \alpha x_0 + L\,. \tag{67.3}$$

Проверим корректность этого определения, т.е. что результат не зависит от выбора вектора сдвига x_0. Пусть $H = x'_0 + L$. Покажем, что

$$\alpha x_0 + L = \alpha x'_0 + L\,. \tag{67.4}$$

Для любого элемента $\alpha x_0 + y$, $y \in L$, из левого множества имеем $\alpha x_0 + y = \alpha(x'_0 + x_0 - x'_0) + y = \alpha x'_0 + \alpha(x_0 - x'_0) + y = \alpha x'_0 + y_1$, где

Глава XIV. **Линейные операторы**

В этой главе рассматриваются отображения (§8), которые действуют в линейных пространствах. Отображения в отличие от матриц будем обозначать рукописными латинскими буквами (например, $\mathcal{A}$), образ вектора x – символом $\mathcal{A}x$.

§76. Определение и простейшие свойства

Пусть V и W – линейные пространства над общим полем P. Отображение

$$\mathcal{A}: V \to W \qquad (76.1)$$

называется *линейным отображением пространства V в пространство W*, если для любых $x, y \in V$, $\alpha \in P$
1) $\mathcal{A}(x+y) = \mathcal{A}x + \mathcal{A}y$;
2) $\mathcal{A}(\alpha x) = \alpha\mathcal{A}x$.

Линейное отображение (76.1) называют также *линейным преобразованием пространства V в пространство W* или *линейным оператором, действующим из пространства V в пространство W*.

Если $V = W$, то линейное отображение $\mathcal{A}: V \to V$ называют *линейным отображением (преобразованием) пространства V в себя*, а чаще – *линейным оператором, действующим в V*.

Если $W = P$, то линейное отображение (76.1) называют *линейной формой* или *линейным функционалом в пространстве V*. Линейный функционал обозначают строчными латинскими буквами (например, f), при этом образ вектора x – символом $f(x)$.

Множество всех линейных операторов, действующих из пространства V в пространство W, будем обозначать символом $\mathcal{L}(V, W)$.

В соответствии с определением равенства отображений (§8) операторы $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V, W)$ равны, если $\mathcal{A}x = \mathcal{B}x$, $\forall x \in V$.

Примеры. 1. Пусть M_n – пространство вещественных многочленов степени не выше n. Отображение $\mathcal{D}: M_n \to M_n$, определенное правилом

$$\mathcal{D}p(t) = p'(t), \qquad (76.2)$$

является линейным и называется *оператором дифференцирования*.

2. Пусть $V = L_1 \oplus L_2$. Отображение $\mathcal{P}: V \to V$, определенное правилом

$$\mathcal{P}x = x_1 \qquad (76.3)$$

для вектора $x \in V$ с разложением $x = x_1 + x_2$, где $x_1 \in L_1$, $x_2 \in L_2$, является линейным и называется *оператором проектирования пространства V на L_1 параллельно L_2*.

Отображение $\mathcal{R}: V \to V$, определенное правилом

$$\mathcal{R}x = x_1 - x_2, \qquad (76.4)$$

также является линейным и называется *оператором отражения пространства V относительно L_1 параллельно L_2*.

3. Отображение $\mathcal{O}: V \to W$, которое каждый вектор $x \in V$ переводит в нулевой вектор $\theta \in W$, является линейным и называется *нулевым оператором*.

4. Отображение $\mathcal{I}: V \to V$, которое каждый вектор $x \in V$ переводит в x, является линейным и называется *тождественным (единичным) оператором*.

5. Изоморфизм φ линейных пространств V и W, очевидно, является линейным оператором, действующим из V в W.

Простейшие свойства. Из определения вытекают следующие свойства линейных операторов.

1°. *Линейный оператор переводит нулевой вектор в нулевой вектор*, так как $\mathcal{A}\theta_1 = \mathcal{A}(0x) = 0\mathcal{A}x = \theta_2$ (здесь θ_1, θ_2 – нулевые векторы V и W соответственно).

2°. *Линейный оператор сохраняет линейные комбинации*, т.е. переводит линейную комбинацию векторов в линейную комбинацию образов с теми же коэффициентами: $\mathcal{A}\left(\sum_{i=1}^k \alpha_i x_i\right) = \sum_{i=1}^k \alpha_i \mathcal{A}x_i$.

3°. *Линейный оператор сохраняет линейную зависимость*, т.е. переводит линейно зависимую систему векторов в линейно зависимую.

Задание линейного оператора. Свойство 2° говорит о том, что для задания линейного оператора $\mathcal{A} \in \mathcal{L}(V, W)$ достаточно определить его только на векторах $e_1, \ldots, e_n$ некоторого базиса пространства V. Зная векторы $\mathcal{A}e_1, \ldots, \mathcal{A}e_n$, можно однозначно найти образ любого вектора $x = \sum_{i=1}^n x_i e_i \in V$: $\mathcal{A}x = \sum_{i=1}^n x_i \mathcal{A}e_i \in W$. Формализуем это рассуждение.

Теорема 76.1. *Пусть $e_1, \ldots, e_n$ – базис пространства V, а $g_1, \ldots, g_n$ – произвольные векторы пространства W. Тогда существует, и притом единственный, линейный оператор $\mathcal{A} \in \mathcal{L}(V, W)$, который переводит векторы $e_1, \ldots, e_n$ в векторы $g_1, \ldots, g_n$ соответственно.*

Доказательство. Построим искомый оператор, положив для каждого вектора $x = \sum_{i=1}^n x_i e_i \in V$

$$\mathcal{A}x = \sum_{i=1}^n x_i g_i. \qquad (76.5)$$

Из единственности разложения вектора x по базису следует, что правило (76.5) однозначно определяет образ вектора x, при этом, как

векторов. Докажем его для k векторов $x_1, \ldots, x_k$. Приравняем нулевому вектору линейную комбинацию этих векторов:

$$\alpha_1 x_1 + \ldots + \alpha_k x_k = \theta. \qquad (85.2)$$

Под действием оператора $\mathcal{A}$ это равенство перейдет в равенство

$$\alpha_1\lambda_1 x_1 + \ldots + \alpha_k\lambda_k x_k = \theta. \qquad (85.3)$$

Умножая обе части (85.2) на λ_k и вычитая полученное равенство из (85.3), получаем, что $\alpha_1(\lambda_1 - \lambda_k)x_1 + \ldots + \alpha_{k-1}(\lambda_{k-1} - \lambda_k)x_{k-1} = \theta$. В силу индуктивного предположения отсюда следует, что $\alpha_1 = \ldots = \alpha_{k-1} = 0$. Тогда (85.2) перейдет в равенство $\alpha_k x_k = \theta$. Так как $x_k \neq \theta$, то $\alpha_k = 0$. Итак, в (85.2) все коэффициенты равны нулю. Следовательно, $x_1, \ldots, x_k$ линейно независимы. ■

Следствие. Линейный оператор, действующий в n-мерном пространстве, не может иметь более чем n различных собственных значений.

Не всякий линейный оператор обладает собственными векторами. Например, в геометрической плоскости V_2 оператор поворота на угол φ, не кратный π, не имеет ни одного собственного вектора, так как ни один ненулевой вектор после такого поворота не останется коллинеарным самому себе. Выясним вопрос о существовании собственных векторов.

§86. Характеристический многочлен

Это понятие уже встречалось в §58 в связи с матрицами второго и третьего порядка. Остановимся теперь на нем более подробно.

Характеристическим многочленом матрицы $A \in P^{n\times n}$ называется функция

$$f(\lambda) = \det(A - \lambda I), \quad \lambda \in P. \qquad (86.1)$$

Теорема 86.1. *Характеристический многочлен (86.1) матрицы $A \in P^{n\times n}$ является многочленом n-й степени от переменной λ над полем P.*

Доказательство. Пусть $A = (a_{ij}) \in P^{n\times n}$. Тогда

$$f(\lambda) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn}-\lambda \end{vmatrix}.$$

Каждый элемент матрицы $A - \lambda I$ представляет собой многочлен от λ степени не выше 1, значит, каждый член $\det(A - \lambda I)$ является многочленом от λ степени не выше n. Отсюда следует, что $f(\lambda)$ – многочлен от λ, степень которого не превосходит n. Осталось показать, что степень этого многочлена в точности равна n. Действительно, все члены $\det(A - \lambda I)$, отличные от $(a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda)$, имеют степень, не превосходящую $n - 2$, следовательно, в многочлене $f(\lambda)$ слагаемые, содержащие λ^{n-1} и λ^n, определяются только членом

$(a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda)$, который согласно формулам Виета имеет вид

$$(-\lambda)^n + (a_{11} + a_{22} + \ldots + a_{nn})(-\lambda)^{n-1} + g_{n-2}(\lambda),$$

где $g_{n-2}(\lambda)$ – многочлен от λ степени не выше $n - 2$. Таким образом,

$$f(\lambda) = a_0 + a_1(-\lambda) + a_2(-\lambda)^2 + \ldots + a_{n-1}(-\lambda)^{n-1} + (-\lambda)^n \qquad (86.2)$$

является многочленом n-й степени от λ над полем P. ■

Замечание 1. Попутно найден коэффициент a_{n-1} характеристического многочлена (86.2): $a_{n-1} = a_{11} + \ldots + a_{nn}$. Кроме того, $a_0 = f(0) = \det A$. Итак, в (86.2)

$$a_0 = \det A, \quad a_{n-1} = \operatorname{tr} A. \qquad (86.3)$$

Теорема 86.2. *Характеристические многочлены подобных матриц совпадают.*

Доказательство теоремы и сама теорема приводятся в §58 (теорема 58.1). ■

Следствие. Все матрицы одного и того же линейного оператора имеют одинаковые характеристические многочлены.

Таким образом, характеристический многочлен матрицы линейного оператора не зависит от базиса, а определяется самим оператором. *Характеристическим многочленом оператора* называется функция

$$f(\lambda) = \det(\mathcal{A} - \lambda\mathcal{I}), \quad \lambda \in P.$$

Из (82.3) следует, что характеристический многочлен оператора совпадает с характеристическим многочленом матрицы этого оператора в произвольном базисе.

Из следствия и второго равенства (86.3) вытекает, что след матрицы A_e линейного оператора $\mathcal{A}$ не зависит от выбора базиса e. Поэтому (как и в §82, п.9) можно ввести понятие *следа оператора* равенством $\operatorname{tr}\mathcal{A} = \operatorname{tr} A_e$.

Теорема 86.3. *Характеристический многочлен индуцированного оператора является делителем характеристического многочлена порождающего оператора.*

Это утверждение следует из (84.2) и замечания 3 из §84. ■

Теорема 86.4. *Если $V = L_1 \oplus \ldots \oplus L_k$ – прямая сумма подпространств $L_1, \ldots, L_k$, инвариантных относительно оператора $\mathcal{A} \in \mathcal{L}(V, V)$, то характеристический многочлен $f(\lambda)$ оператора $\mathcal{A}$ равен произведению характеристических многочленов $f_1(\lambda), \ldots, f_k(\lambda)$ индуцированных операторов $\mathcal{A}|L_1, \ldots, \mathcal{A}|L_k$:*

$$f(\lambda) = f_1(\lambda)\ldots f_k(\lambda). \qquad (86.4)$$

Утверждение является следствием теоремы 84.2. ■

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

Имеет место более общее утверждение.

Теорема 77.4. *Две матрицы A и B над полем P одинакового размера $m \times n$ эквивалентны тогда и только тогда, когда они являются матрицами одного и того же линейного оператора $\mathcal{A} \in \mathcal{L}(V, W)$, где V и W – линейные пространства над полем P размерностей n и m соответственно.*

Доказательство. Необходимость. Пусть $A, B \in P^{m\times n}$ и $B = D^{-1}AC$. Рассмотрим любые линейные пространства V и W над полем P такие, что $\dim V = n$, $\dim W = m$. Возьмем в пространстве V произвольный базис e, а в пространстве W – базис f. В силу взаимно однозначного соответствия между $P^{m\times n}$ и $\mathcal{L}(V, W)$ (теорема 77.1) существует единственный оператор $\mathcal{A} \in \mathcal{L}(V, W)$, который в паре базисов e и f имеет матрицу A. Тогда согласно (77.5) матрица B будет матрицей этого же оператора в паре базисов $t = eC$ и $s = fD$.

Достаточность рассмотрена в теореме 77.3. ■

§78. Линейное пространство операторов

На множестве $\mathcal{L}(V, W)$ введем операции сложения операторов и умножения оператора на число $\alpha \in P$.

Суммой линейных операторов $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V, W)$ называется отображение $\mathcal{C}: V \to W$, выполняемое по правилу $\mathcal{C}x = \mathcal{A}x + \mathcal{B}x$, $\forall x \in V$. Обозначение: $\mathcal{A} + \mathcal{B}$. Итак,

$$(\mathcal{A} + \mathcal{B})x = \mathcal{A}x + \mathcal{B}x, \quad \forall x \in V. \qquad (78.1)$$

Произведением линейного оператора $\mathcal{A} \in \mathcal{L}(V, W)$ *на число* $\alpha \in P$ называется отображение $\mathcal{C}: V \to W$, выполняемое по правилу $\mathcal{C}x = \alpha\mathcal{A}x$, $\forall x \in V$. Обозначение: $\alpha\mathcal{A}$. Итак,

$$(\alpha\mathcal{A})x = \alpha\mathcal{A}x, \quad \forall x \in V.$$

Теорема 78.1. *Для любых операторов $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V, W)$ и числа $\alpha \in P$*

$$\mathcal{A} + \mathcal{B} \in \mathcal{L}(V, W), \quad \alpha\mathcal{A} \in \mathcal{L}(V, W).$$

Доказательство. Действительно, для любых $x, y \in V$ согласно (78.1) имеем $(\mathcal{A} + \mathcal{B})(x + y) = \mathcal{A}(x + y) + \mathcal{B}(x + y)$. В силу линейности $\mathcal{A}$, $\mathcal{B}$ и аксиом линейного пространства $(\mathcal{A} + \mathcal{B})(x + y) = (\mathcal{A}x + \mathcal{A}y) + (\mathcal{B}x + \mathcal{B}y) = (\mathcal{A}x + \mathcal{B}x) + (\mathcal{A}y + \mathcal{B}y) = (\mathcal{A} + \mathcal{B})x + (\mathcal{A} + \mathcal{B})y$.

Аналогично показывается, что $(\mathcal{A} + \mathcal{B})(\lambda x) = \lambda((\mathcal{A} + \mathcal{B})x)$ для любых $x \in V$, $\lambda \in P$. Следовательно, $\mathcal{A} + \mathcal{B} \in \mathcal{L}(V, W)$.

Точно так же доказывается, что $\alpha\mathcal{A} \in \mathcal{L}(V, W)$. ■

Следствие 1. *Сложение операторов и умножение оператора на число являются внутренним и внешним законами композиции на множестве $\mathcal{L}(V, W)$.*

Теорема 78.2. *Множество $\mathcal{L}(V, W)$ – линейное пространство над полем P относительно введенных выше операций.*

Доказательство. Достаточно проверить аксиомы линейного пространства, взяв в качестве нулевого элемента нулевое отображение $\mathcal{O} \in \mathcal{L}(V, W)$, а в качестве противоположного к оператору $\mathcal{A}$ отображение $-\mathcal{A} \in \mathcal{L}(V, W)$, выполняемое по правилу $(-\mathcal{A})x = -\mathcal{A}x$, $\forall x \in V$. Все аксиомы вытекают из соответствующих аксиом линейного пространства, примененных к V и W, и проверяются по единой схеме. Проверим, например, ассоциативность. Для любых $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathcal{L}(V, W)$ и любого $x \in V$

$$((\mathcal{A} + \mathcal{B}) + \mathcal{C})x = (\mathcal{A} + \mathcal{B})x + \mathcal{C}x = (\mathcal{A}x + \mathcal{B}x) + \mathcal{C}x = \mathcal{A}x + (\mathcal{B}x + \mathcal{C}x),$$
$$(\mathcal{A} + (\mathcal{B} + \mathcal{C}))x = \mathcal{A}x + (\mathcal{B} + \mathcal{C})x = \mathcal{A}x + (\mathcal{B}x + \mathcal{C}x).$$

Таким образом, $(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C})$. ■

Теорема 78.3. *Если $\dim V = n$, $\dim W = m$, то линейное пространство $\mathcal{L}(V, W)$ изоморфно пространству матриц $P^{m\times n}$.*

Доказательство. Зафиксируем базисы e и f пространств V и W. Построим отображение $\varphi: \mathcal{L}(V, W) \to P^{m\times n}$, положив $\varphi(\mathcal{A}) = A_{fe}$. Это отображение биективно в силу теоремы 77.1. Покажем, что оно сохраняет законы композиции, т.е. что

$$(\mathcal{A} + \mathcal{B})_{fe} = A_{fe} + B_{fe}, \quad (\alpha\mathcal{A})_{fe} = \alpha A_{fe}. \qquad (78.2)$$

Пусть $A_{fe} = (a_{ij})$, $B_{fe} = (b_{ij})$. Тогда, согласно (77.1), $\mathcal{A}e_j = \sum_{i=1}^m a_{ij}f_i$, $\mathcal{B}e_j = \sum_{i=1}^m b_{ij}f_i$, поэтому $(\mathcal{A} + \mathcal{B})e_j = \mathcal{A}e_j + \mathcal{B}e_j = \sum_{i=1}^m (a_{ij} + b_{ij})f_i$. По определению (77.2) матрицы линейного оператора отсюда следует первое из соотношений (78.2).

Аналогично проверяется второе соотношение. ■

Следствие 2. $\dim\mathcal{L}(V, W) = \dim V \cdot \dim W$.

Замечание. Соотношения (78.2) означают, что при сложении операторов их матрицы складываются, при умножении оператора на число его матрица умножается на это же число.

§79. Умножение линейных операторов

Тот факт, что линейные операторы образуют линейное пространство, используется редко. Дело в том, что о линейных операторах можно сказать гораздо больше, используя еще одну операцию – умножение операторов, знакомую по §8 как суперпозиция или композиция отображений. Напомним определение.

Пусть V, W, Z – линейные пространства над полем P. *Произведением линейных операторов* $\mathcal{A} \in \mathcal{L}(V, W)$, $\mathcal{B} \in \mathcal{L}(W, Z)$ называется отображение $\mathcal{C}: V \to Z$, выполняемое по правилу $\mathcal{C}x = \mathcal{B}(\mathcal{A}x)$, $\forall x \in V$. Обозначение: $\mathcal{B}\mathcal{A}$. Итак, $(\mathcal{B}\mathcal{A})x = \mathcal{B}(\mathcal{A}x)$, $\forall x \in V$.

Теорема 79.1. *Если $\mathcal{A} \in \mathcal{L}(V, W)$, $\mathcal{B} \in \mathcal{L}(W, Z)$, то $\mathcal{B}\mathcal{A} \in \mathcal{L}(V, Z)$.*

Доказательство. Линейность оператора $\mathcal{B}\mathcal{A}$ проверяется непосредственно: для любых $x, y \in V$ и $\alpha \in P$

Из (87.1) следует, что собственное подпространство W_{λ_0} состоит из нулевого вектора θ и всех собственных векторов, отвечающих λ_0. Легко проверить, что собственное подпространство инвариантно относительно оператора $\mathcal{A}$. Размерность собственного подпространства W_{λ_0} называется *геометрической кратностью собственного значения* λ_0, а кратность λ_0 как корня характеристического многочлена называется его *алгебраической кратностью*.

Теорема 87.1. *Геометрическая кратность собственного значения не превосходит его алгебраической кратности.*

Доказательство. Пусть m и s – алгебраическая и геометрическая кратности собственного значения λ_0 оператора $\mathcal{A} \in \mathcal{L}(V, V)$.

Собственное подпространство W_{λ_0} инвариантно относительно оператора $\mathcal{A}$, следовательно, можно рассматривать индуцированный оператор $\mathcal{A}|W_{\lambda_0}$. Найдем его характеристический многочлен $f_1(\lambda)$. Пусть $e_1, \ldots, e_s$ – базис W_{λ_0}. Тогда согласно (87.1) матрицей оператора $\mathcal{A}|W_{\lambda_0}$ в этом базисе будет диагональная матрица s-го порядка с элементами λ_0 на главной диагонали. Следовательно, $f_1(\lambda) = (\lambda_0 - \lambda)^s$. Согласно теореме 86.3, $(\lambda_0 - \lambda)^s$ является делителем характеристического многочлена $f(\lambda)$ оператора $\mathcal{A}$, но $(\lambda_0 - \lambda)$ входит в характеристический многочлен $f(\lambda)$ ровно m раз. Значит, $s \leq m$. ■

Теорема 87.2. *Сумма собственных подпространств оператора, отвечающих различным собственным значениям, является прямой суммой.*

Доказательство. Пусть $\lambda_1, \ldots, \lambda_p$ – попарно различные собственные значения оператора $\mathcal{A}$. Тогда для собственных подпространств $W_1, \ldots, W_{\lambda_p}$ выполнено условие прямой суммы (теорема 66.1): любая система ненулевых векторов, взятых по одному из каждого W_{λ_k}, линейно независима как система собственных векторов, отвечающих различным собственным значениям (теорема 85.1). ■

§88. Операторы простой структуры

Критерии простой структуры. Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ называется *оператором простой структуры*, если в пространстве V существует базис из собственных векторов оператора $\mathcal{A}$.

Теорема 88.1. *Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ имеет простую структуру тогда и только тогда, когда в пространстве V существует базис, в котором он имеет диагональную матрицу.*

Доказательство. Пусть $\dim V = n$. Согласно определению оператор $\mathcal{A}$ имеет простую структуру тогда и только тогда, когда он имеет n линейно независимых собственных векторов $e_1, \ldots, e_n$. Это равносильно существованию базиса $e_1, \ldots, e_n$, в котором матрица A_e

оператора $\mathcal{A}$ имеет вид

$$A_e = \begin{bmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{bmatrix}, \qquad (88.1)$$

где $\lambda_1, \ldots, \lambda_n$ – собственные значения, соответствующие собственным векторам $e_1, \ldots, e_n$. ■

Следствие. В n-мерном пространстве линейный оператор, имеющий n различных собственных значений, является оператором простой структуры.

Обратное не верно. Тождественный оператор имеет простую структуру, однако все его собственные значения равны 1.

В соответствии с (88.1) оператор простой структуры называют также *диагонализуемым оператором*.

Теорема 88.2. *Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ имеет простую структуру тогда и только тогда, когда все его собственные подпространства в прямой сумме дают все пространство V:*

$$W_{\lambda_1} \oplus \ldots \oplus W_{\lambda_p} = V. \qquad (88.2)$$

Доказательство. Необходимость. Пусть $\mathcal{A}$ имеет простую структуру. Тогда в пространстве V существует базис $e_1, \ldots, e_n$, состоящий из собственных векторов оператора $\mathcal{A}$. Каждый вектор этого базиса принадлежит некоторому собственному пространству W_{λ_k}. Следовательно, $W_{\lambda_1} + \ldots + W_{\lambda_p} = V$. Но эта сумма является прямой в силу теоремы 87.2.

Достаточность вытекает из критерия прямой суммы (теорема 66.1): совокупность базисов собственных подпространств W_{λ_k}, $k = \overline{1, p}$, образует базис V, тем самым пространство V имеет базис из собственных векторов оператора $\mathcal{A}$. ■

Замечание 1. На основании теоремы 66.1 условие (88.2) может быть заменено равносильным условием

$$\dim W_{\lambda_1} + \ldots + \dim W_{\lambda_p} = \dim V. \qquad (88.3)$$

Теорема 88.3. *Линейный оператор, действующий в комплексном пространстве, имеет простую структуру тогда и только тогда, когда для каждого собственного значения этого оператора геометрическая кратность совпадает с алгебраической.*

Доказательство. Пусть $\lambda_1, \ldots, \lambda_p$, где $\lambda_i \neq \lambda_j$ при $i \neq j$, – все различные собственные значения оператора $\mathcal{A} \in \mathcal{L}(V, V)$ и пусть m_k и s_k, $k = \overline{1, p}$, – алгебраические и геометрические кратности λ_k.

в которых оператор $\mathcal{A}$ имеет матрицу $I_r \in P^{m\times n}$ вида

$$I_r = \left[\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & O \\ 0 & & 1 & \\ \hline & O & & O \end{array}\right] \Big\} r,$$

в которой все элементы равны нулю, кроме первых r диагональных элементов, равных 1.

Доказательство. Возьмем произвольные базисы t и s пространств V и W. Пусть A_{st} – матрица оператора $\mathcal{A}$ в паре базисов t и s. Тогда, согласно теореме 80.3, $\operatorname{rg}\mathcal{A}_{st} = r$. В силу теоремы 16.10 отсюда следует, что матрицы A_{st} и I_r эквивалентны. Следовательно (теорема 77.4), они являются матрицами одного линейного оператора. Отсюда вытекает утверждение теоремы. ■

Базисы e и f, в которых оператор $\mathcal{A}$ имеет матрицу I_r, называют *канонической парой базисов*.

Теоремой 80.5 решается вопрос о наиболее простой форме матрицы линейного оператора, действующего в различных пространствах V и W. Очевидно, этот вопрос далеко не праздный, так как одному и тому же линейному оператору соответствует целый класс эквивалентных друг другу матриц и выбор самой простой из них упрощает исследование свойств оператора.

Остановимся более подробно на одном из частных случаев линейных операторов – случае, когда пространство W совпадает с основным полем P.

§81. Линейные формы

Как уже отмечалось в §76, линейное отображение $f: V \to P$ линейного пространства V над полем P в это поле называется *линейной формой* (или *линейным функционалом*) *в пространстве V*.

Примеры. 1. Простейшей линейной формой в пространстве V является отображение $f: V \to P$, определенное равенством $f(x) = 0$, где $0 \in P$.

2. Линейной формой является отображение $f: \mathbb{R}^n \to \mathbb{R}$, определенное следующим правилом: если $(x_1, \ldots, x_n) \in \mathbb{R}^n$, то $f(x) = x_1$.

3. В пространстве M_n вещественных многочленов степени не выше n линейной формой является отображение $f: M_n \to \mathbb{R}$, определенное правилом: если $p(t) \in M_n$, то $f(p) = p(1)$.

Из общей теории линейных операторов следует, что если $e_1, \ldots, e_n$ – базис пространства V, то линейная форма $f \in \mathcal{L}(V, P)$ однозначно определяется числами

$$\alpha_1 = f(e_1), \ \ldots, \ \alpha_n = f(e_n).$$

При этом для произвольного вектора $x = \sum_{i=1}^n x_i e_i$ в силу линейности f имеем

$$f(x) = \alpha_1 x_1 + \ldots + \alpha_n x_n. \qquad (81.1)$$

Представление (81.1) называется *общим видом линейной формы f в базисе $e_1, \ldots, e_n$*. Числа $\alpha_1, \ldots, \alpha_n$ называются *коэффициентами линейной формы f в базисе $e_1, \ldots, e_n$*.

Сопряженное пространство. Известно (теорема 78.2), что множество $\mathcal{L}(V, P)$ всех линейных форм в линейном пространстве V образует линейное пространство относительно операций сложения и умножения на число, введенных в §78 правилами

$$(f_1 + f_2)(x) = f_1(x) + f_2(x), \quad (\alpha f)(x) = \alpha f(x).$$

Линейное пространство всех линейных форм на пространстве V называется *сопряженным пространством к пространству V*. Оно обозначается символом V^*.

Теорема 81.1. $\dim V^* = \dim V$.

Это равенство вытекает из теоремы 78.3 и ее следствия, так как $\dim P = 1$. ■

Следствие. Всякое конечномерное линейное пространство изоморфно своему сопряженному. Это вытекает из теоремы 63.1 об изоморфизме линейных пространств одинаковой размерности.

Специальное представление линейной формы в евклидовом (унитарном) пространстве.

Теорема 81.2. *Для любой линейной формы f в евклидовом (унитарном) пространстве V существует, и притом единственный, вектор $h \in V$ такой, что*

$$f(x) = (x, h), \quad \forall x \in V.$$

Доказательство. Пусть $e_1, \ldots, e_n$ – ортонормированный базис V и $\alpha_1 \ldots, \alpha_n$ – коэффициенты линейной формы f в этом базисе. Тогда вектор $h = \sum_{i=1}^n \overline{\alpha_i} e_i$ будет искомым в силу (81.1) и (70.7). Единственность вектора h следует из того, что если вектор h_1 удовлетворяет требованиям теоремы, то $(x, h_1) = (x, h)$, $\forall x \in V$, т.е. $(x, h_1 - h) = 0$, $\forall x \in V$. Так как $h_1 - h \in V$, то отсюда следует, что $h_1 - h = \theta$. ■

§82. Алгебра линейных операторов, действующих в одном пространстве

Обратимся теперь к линейным операторам, действующим в одном пространстве. Основной целью наших исследований будет поиск матрицы такого оператора, имеющей наиболее простую форму. Однако об этом речь пойдет позже. В ближайших параграфах будут рассматриваться общие вопросы, относящиеся к операторам, действующим в одном пространстве.

1) характеристический многочлен $f(\lambda) = (\lambda_0 - \lambda)^k$;
2) собственное значение $\lambda = \lambda_0$ алгебраической кратности k;
3) собственные векторы, которые являются нетривиальными решениями однородной системы уравнений с матрицей

$$B = J_k(\lambda_0) - \lambda_0 I = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix},$$

ранг которой, очевидно, равен $k - 1$. Таким образом, матрица $J_k(\lambda_0)$ (или оператор, задаваемый этой матрицей) имеет (теорема 30.2) один линейно независимый собственный вектор $(k - (k-1) = 1)$ и не может быть матрицей простой структуры.

§89. Треугольная форма матрицы линейного оператора

Теорема 89.1. *В n-мерном комплексном пространстве V для любого линейного оператора $\mathcal{A} \in \mathcal{L}(V, V)$ существует система n вложенных друг в друга инвариантных подпространств $L_1, \ldots, L_n$ всех размерностей от 1 до n, т.е. таких, что*

$$L_1 \subset L_2 \subset \ldots \subset L_n = V, \qquad (89.1)$$

где $\dim L_k = k$, $k = \overline{1, n}$.

Доказательство. Используем индукцию по n. Для $n = 1$ утверждение теоремы очевидно. Пусть теорема верна для всех линейных пространств размерности $n - 1$. Докажем ее для n-мерного пространства V, воспользовавшись следующей леммой.

Лемма. *Линейный оператор, действующий в n-мерном комплексном пространстве, обладает инвариантным подпространством размерности $n - 1$.*

Доказательство леммы. Линейный оператор $\mathcal{A}$, действующий в комплексном пространстве V, имеет собственное значение λ (теорема 86.6). Значит, $\det(\mathcal{A} - \lambda\mathcal{I}) = 0$ и $\operatorname{rg}(\mathcal{A} - \lambda\mathcal{I}) \leq n - 1$. Следовательно, $\dim\operatorname{im}(\mathcal{A} - \lambda\mathcal{I}) \leq n - 1$ и в пространстве V существует подпространство L размерности $n - 1$, которое содержит $\operatorname{im}(\mathcal{A} - \lambda\mathcal{I})$. Очевидно, L инвариантно относительно оператора $\mathcal{A} - \lambda\mathcal{I}$. Покажем, что оно инвариантно и относительно $\mathcal{A}$. Пусть $x \in L$, тогда $(\mathcal{A} - \lambda\mathcal{I})x = y \in L$. Отсюда следует, что $\mathcal{A}x = \lambda x + y \in L$. Лемма доказана.

Вернемся к доказательству теоремы. Согласно лемме оператор $\mathcal{A}$, действующий в n-мерном комплексном пространстве V, имеет инвариантное подпространство L_{n-1} размерности $n - 1$. Тогда индуцированный оператор $\mathcal{A}|L_{n-1}$ действует в $(n-1)$-мерном комплексном пространстве L_{n-1} и согласно индуктивному предположению для

него существует система вложенных инвариантных подпространств $L_1 \subset \ldots \subset L_{n-1}$ таких, что $\dim L_k = k$, $k = \overline{1, n-1}$. Так как действия операторов $\mathcal{A}$ и $\mathcal{A}|L_{n-1}$ совпадают, то подпространства $L_1, \ldots, L_{n-2} \subset L_{n-1}$ инвариантны относительно оператора $\mathcal{A}$. Остается добавить, что $L_{n-1} \subset L_n = V$. ■

Теорема 89.2. *Для любого линейного оператора $\mathcal{A}$, действующего в комплексном пространстве, существует базис, в котором матрица линейного оператора имеет треугольную форму.*

Доказательство. Согласно теореме 89.1 для оператора $\mathcal{A}$ найдется система инвариантных подпространств $L_1, \ldots, L_n$ таких, что $\dim L_k = k$ и $L_1 \subset \ldots \subset L_{n-1} \subset L_n = V$. Искомый базис $e_1, \ldots, e_n$ строим так: в качестве e_1 берем любой базис L_1, в качестве e_k, где $k > 1$, – вектор, дополняющий базис $\underline{L_{k-1}}$ до базиса L_k. В силу инвариантности подпространств L_k, $k = \overline{1, n}$, матрица A_e имеет верхнюю треугольную форму. ■

Замечание 1. Перенумерацией построенного базиса в обратном порядке можно получить нижнюю треугольную форму матрицы линейного оператора.

Замечание 2. На главной диагонали матрицы A_e расположены собственные значения оператора $\mathcal{A}$.

Теорема 89.3. *Любая квадратная комплексная матрица подобна матрице, имеющей треугольную форму.*

Доказательство теоремы повторяет доказательство теоремы 88.4 и ведется по схеме (88.6). ■

§90. Нильпотентный оператор

Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ называется *нильпотентным*, если существует число $q \in \mathbb{N}$ такое, что $\mathcal{A}^q = \mathcal{O}$. Наименьшее число q, обладающее этим свойством, называется *индексом нильпотентности (высотой)* оператора $\mathcal{A}$. Очевидно, что индекс нильпотентности ненулевого оператора $q \geq 2$. Аналогично определяется нильпотентная матрица $A \in P^{n\times n}$ и ее индекс.

Примеры. 1. В пространстве вещественных многочленов M_n оператор дифференцирования (76.3) является нильпотентным оператором индекса $n + 1$.

2. Жорданова клетка

$$J_k(0) = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & 0 \end{bmatrix}$$

является нильпотентной матрицей индекса k.

Теорема 90.1. *Если $\mathcal{A} \in \mathcal{L}(V, V)$ – нильпотентный оператор индекса q и $x_0 \in V$ – вектор, для которого $\mathcal{A}^{q-1}x_0 \neq \theta$, то*

(§5). Таким образом,

$$(A_e)^{-1} = (\mathcal{A}^{-1})_e. \quad ■ \qquad (83.2)$$

Оператор $\mathcal{A} \in \mathcal{L}(V, V)$ называется *невырожденным*, если его ядро состоит только из нулевого вектора, т.е.

$$\ker\mathcal{A} = \{\theta\}, \qquad (83.3)$$

и *вырожденным* в противном случае.

Теорема 83.5. *В конечномерном пространстве V следующие утверждения равносильны: для $\mathcal{A} \in \mathcal{L}(V, V)$*

1) $\mathcal{A}\mathcal{A}^{-1} = \mathcal{I}$; 4) $\operatorname{im}\mathcal{A} = V$; 6) $\mathcal{A}$ *обратим*;
2) $\mathcal{A}^{-1}\mathcal{A} = \mathcal{I}$; 5) $\det\mathcal{A} \neq 0$; 7) $\mathcal{A}$ *биективен*. (83.4)
3) $\mathcal{A}$ *не вырожден*;

Доказательство. $1 \iff 2 \iff 5 \iff 6 \iff 7$. Пусть $e = (e_1, \ldots, e_n)$ – произвольный базис пространства V. В силу свойства (79.2) утверждение 1 равносильно матричному равенству $A_e(\mathcal{A}^{-1})_e = I$ или, с учетом теоремы 5.3, соотношениям $A_e(A_e)^{-1} = (A_e)^{-1}A_e = I$, эквивалентным (в силу (79.2)) равенствам (83.1) и (в силу теоремы 5.2) условию

$$\det A_e \neq 0. \qquad (83.5)$$

Это доказывает импликации $1 \iff 5$, $1 \iff 2$, $1 \iff 6$ и, согласно теореме 83.1, $1 \iff 7$.

$1 \iff 3 \iff 4$. Утверждение 1 в силу (83.5) равносильно равенству

$$\operatorname{rg}\mathcal{A} = n, \qquad (83.6)$$

которое согласно (80.2) эквивалентно тому, что

$$\operatorname{def}\mathcal{A} = 0. \qquad (83.7)$$

Равенство (83.6) означает, что $\dim\operatorname{im}\mathcal{A} = \dim V$ или, в силу теоремы 64.3 о монотонности размерности, что $\operatorname{im}\mathcal{A} = V$. Это доказывает импликацию $1 \iff 4$. Равенство (83.7) означает, что $\dim\ker\mathcal{A} = 0$. Это доказывает импликацию $1 \iff 3$. ■

Теорема 83.6. *Произведение обратимых операторов обратимо, при этом*

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

Это утверждение уже было доказано в §8 применительно к отображениям общего вида. ■

Следствие 1. *Умножение линейных операторов является алгебраической операцией на множестве всех обратимых операторов, действующих в пространстве V.*

Следствие 2. *Множество всех обратимых линейных операторов из $\mathcal{L}(V, V)$ образует неабелеву мультипликативную группу.*

Глава XV. **Структура линейного оператора в комплексном пространстве**

Приступим к вопросу о поиске в классе подобных матриц, соответствующих линейному оператору $\mathcal{A} \in \mathcal{L}(V, V)$, матрицы наиболее простого вида. Ответ на этот вопрос совсем прост, когда линейный оператор действует в пространствах V и W, никак не связанных между собой (теорема 80.5). Гораздо интереснее (и сложнее), когда $V = W$. В первом случае базисы пространств V и W можно было выбирать независимо друг от друга, во втором – необходимо строить базис одного пространства $V = W$. Мы увидим, к какому разнообразию ответов приводит эта меньшая свобода выбора.

Наиболее полные результаты этой главы относятся к случаю, когда V – комплексное пространство.

§84. Инвариантные подпространства

Пусть V – линейное пространство над полем P и $\mathcal{A} \in \mathcal{L}(V, V)$. Линейное подпространство L пространства V называется *инвариантным подпространством относительно оператора $\mathcal{A}$*, если для любого вектора x из L его образ $\mathcal{A}x$ также лежит в L.

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

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

3. Для оператора дифференцирования (76.2) в пространстве M_n вещественных многочленов инвариантными подпространствами являются все подпространства $M_0, M_1, \ldots, M_{n-1}$.

Для линейного оператора наличие инвариантных подпространств означает возможность построить базис, в котором его матрица имеет несколько более простую форму, чем матрица общего вида.

Теорема 84.1. *Пусть $\mathcal{A} \in \mathcal{L}(V, V)$ и L – инвариантное подпространство относительно $\mathcal{A}$. Тогда существует базис пространства V, в котором матрица оператора $\mathcal{A}$ имеет квазитреугольную форму.*

Доказательство. Пусть $e_1, \ldots, e_k$ – базис L. Дополним его до базиса $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ пространства V. Построим матрицу оператора $\mathcal{A}$ в этом базисе. Из инвариантности L вытекает, что $\mathcal{A}e_1, \ldots, \mathcal{A}e_k \in L$ и, следовательно, векторы $\mathcal{A}e_1, \ldots, \mathcal{A}e_k$ линейно

1. Нильпотентность оператора $\mathcal{A}|N$ означает, что $\mathcal{A}^k x = \theta$, $\forall x \in N$, при некотором $k \in \mathbb{N}$. Следовательно, $N \subseteq N_k \subseteq N_q$ и

$$\dim N \leq \dim N_q. \qquad (90.3)$$

2. Обратимость оператора $\mathcal{A}|T$ означает, что $\operatorname{im}\mathcal{A}|T = T$. Следовательно, для любого вектора $y \in T$ имеет место представление $y = \mathcal{A}y_1$, где $y_1 \in T$. Используя такое же представление для вектора y_1 и всех последующих, получаем, что $y = \mathcal{A}y_1 = \mathcal{A}^2 y_2 = \ldots = \mathcal{A}^q y_q$. Таким образом, $T \subseteq T_q$ и

$$\dim T \leq \dim T_q. \qquad (90.4)$$

Так как $\dim N + \dim T = \dim V = \dim N_q + \dim T_q$, то из (90.3), (90.4) следует, что $N = N_q$, $T = T_q$. ■

Следствие 2. *Оператор $\mathcal{A}$ на подпространстве N_q имеет только нулевые собственные значения, а на подпространстве T_q не имеет нулевых собственных значений.* Это вытекает из доказанной теоремы с учетом теоремы 90.2 и того, что обратимый оператор не вырожден.

Следствие 3. *Для оператора $\mathcal{A}$, действующего в комплексном пространстве V, с характеристическим многочленом $f(\lambda) = \det(\mathcal{A} - \lambda\mathcal{I}) = (-\lambda)^{m_1}(\lambda_2 - \lambda)^{m_2}\ldots(\lambda_p - \lambda)^{m_p}$:*

а) характеристические многочлены $f_1(\lambda)$ и $f_2(\lambda)$ операторов $\mathcal{A}|N_q$ и $\mathcal{A}|T_q$ имеют вид

$$f_1(\lambda) = (-\lambda)^{m_1}, \ f_2(\lambda) = (\lambda_2 - \lambda)^{m_2}\ldots(\lambda_p - \lambda)^{m_p}; \qquad (90.5)$$

б) при этом

$$\dim N_q = m_1, \quad \dim T_q = m_2 + \ldots + m_p. \qquad (90.6)$$

Соотношение (90.5) вытекает из (86.4), если учесть, что $f_2(0) \neq 0$. Соотношение (90.6) следует из (90.5), так как размерность пространства совпадает со степенью характеристического многочлена.

§91. Корневые подпространства

Теоремой 90.3 осуществляется одновременное расщепление пространства $V = N_q \oplus T_q$, характеристического многочлена $f(\lambda) = f_1(\lambda)f_2(\lambda)$ и самого оператора $\mathcal{A}$ на индуцированные операторы $\mathcal{A}|N_q$ и $\mathcal{A}|T_q$. Такое расщепление направлено на выделение во всем пространстве V максимального подпространства N_q, на котором оператор $\mathcal{A}$ имеет только нулевые собственные значения. Этот же прием может быть использован и для других собственных значений.

Теорема 91.1 (о расщеплении линейного оператора). *Для любого линейного оператора $\mathcal{A}$, действующего в комплексном пространстве V, с характеристическим многочленом*

$$f(\lambda) = a_n(\lambda_1 - \lambda)^{m_1}\ldots(\lambda_p - \lambda)^{m_p}, \quad \text{где } \lambda_i \neq \lambda_j \text{ при } i \neq j, \qquad (91.1)$$

существуют инвариантные подпространства $K_{\lambda_1}, \ldots, K_{\lambda_p}$ такие, что

$$V = K_{\lambda_1} \oplus \ldots \oplus K_{\lambda_p}; \qquad (91.2)$$

$$\dim K_{\lambda_j} = m_j, \quad j = \overline{1, p}; \qquad (91.3)$$

$$f_j(\lambda) = \det(\mathcal{A}|K_{\lambda_j} - \lambda\mathcal{I}) = (\lambda_j - \lambda)^{m_j}, \quad j = \overline{1, p}. \qquad (91.4)$$

Доказательство. Заметим предварительно, что:

а) если $\lambda_{\mathcal{A}}$ и $\lambda_{\mathcal{B}}$ – собственные значения операторов $\mathcal{A}$ и $\mathcal{B} = \mathcal{A} - \lambda_0\mathcal{I}$, то, как нетрудно проверить,

$$\lambda_{\mathcal{B}} = \lambda_{\mathcal{A}} - \lambda_0; \qquad (91.5)$$

б) если подпространство L инвариантно относительно оператора $\mathcal{B} = \mathcal{A} - \lambda_0\mathcal{I}$, то L инвариантно относительно $\mathcal{A}$, так как если $x \in L$, то $\mathcal{B}x = y \in L$, т.е. $(\mathcal{A} - \lambda_0\mathcal{I})x = y \in L$, и, следовательно, $\mathcal{A}x = \lambda_0 x + y \in L$.

Доказательство теоремы состоит из $p - 1$ однотипных шагов.

Первый шаг. Рассмотрим оператор $\mathcal{B}_1 = \mathcal{A} - \lambda_1\mathcal{I}$. Применив к нему теорему 90.3, получим подпространства N_{q_1} и T_{q_1}, которые с учетом (91.5) обладают следующими свойствами:

$V = N_{q_1} \oplus T_{q_1}$;
N_{q_1}, T_{q_1} инвариантны относительно $\mathcal{B}_1$;
$\dim N_{q_1} = m_1$, $\dim T_{q_1} = m_2 + \ldots + m_p$;
$\det(\mathcal{B}_1|N_{q_1} - \lambda\mathcal{I}) = (-\lambda)^{m_1}$;
$\det(\mathcal{B}_1|T_{q_1} - \lambda\mathcal{I}) = (\lambda_2 - \lambda_1 - \lambda)^{m_2}\ldots(\lambda_p - \lambda_1 - \lambda)^{m_p}$.

Положив $K_{\lambda_1} = N_{q_1}$, $V_1 = T_{q_1}$, получим подпространства K_{λ_1} и V_1, инвариантные относительно оператора $\mathcal{A}$ (согласно п."б") и такие, что:

$$\left.\begin{array}{l} V = K_{\lambda_1} \oplus V_1; \\ \dim K_{\lambda_1} = m_1, \quad \dim V_1 = m_2 + \ldots + m_p; \\ \det(\mathcal{A}|K_{\lambda_1} - \lambda\mathcal{I}) = (\lambda_1 - \lambda)^{m_1}; \\ \det(\mathcal{A}|V_1 - \lambda\mathcal{I}) = (\lambda_2 - \lambda)^{m_2}\ldots(\lambda_p - \lambda)^{m_p}. \end{array}\right. \qquad (91.6)$$

Второй шаг состоит в применении первого шага к оператору $\mathcal{A}_1 = \mathcal{A}|V_1$. В результате второго шага получим подпространства K_{λ_2} и V_2, обладающие свойствами (91.6) с очевидными изменениями. Через $p - 1$ шагов получим все подпространства $K_{\lambda_1}, \ldots, K_{\lambda_p} = V_{p-1}$, удовлетворяющие требованиям теоремы. ■

Следствие. Для любого линейного оператора, действующего в комплексном пространстве, существует базис, в котором его матрица имеет квазидиагональную форму (согласно теореме 84.2), *у которой число диагональных клеток совпадает с числом различных собственных значений, а их размеры – с алгебраическими кратностями собственных значений, или, в матричной формулировке,* Условие существования собственных векторов линейного оператора определяется его характеристическим многочленом.

Теорема 86.5. *Пусть V – линейное пространство над полем P. Число $\lambda \in P$ является собственным значением оператора $\mathcal{A} \in \mathcal{L}(V, V)$ тогда и только тогда, когда λ – корень его характеристического многочлена.*

Доказательство. Согласно определению число λ является собственным значением оператора $\mathcal{A}$ тогда и только тогда, когда существует вектор x, удовлетворяющий условиям

$$\begin{cases} \mathcal{A}x = \lambda x, \\ x \neq \theta, \\ \lambda \in P, \end{cases} \iff \begin{cases} (\mathcal{A} - \lambda\mathcal{I})x = \theta, \\ x \neq \theta, \\ \lambda \in P. \end{cases}$$

Это равносильно вырожденности оператора $\mathcal{A} - \lambda\mathcal{I}$ при некотором $\lambda \in P$ или, в силу (83.4), условию

$$\det(\mathcal{A} - \lambda\mathcal{I}) = 0 \tag{86.5}$$

при некотором $\lambda \in P$. Таким образом, число λ является собственным значением оператора $\mathcal{A}$ тогда и только тогда, когда оно является корнем характеристического многочлена оператора $\mathcal{A}$, принадлежащим полю P. ■

Уравнение (86.5) называется *характеристическим уравнением* для оператора $\mathcal{A}$.

Собственные векторы линейного оператора в комплексном пространстве. Итак, вопрос о существовании собственных векторов сводится к вопросу о существовании корней характеристического многочлена, принадлежащих основному полю. Известно, что не во всяком поле многочлены имеют корни. Однако, как показано в §49, в алгебраически замкнутом поле $\mathbb{C}$ комплексных чисел любой многочлен степени $n \geq 1$ имеет n корней, если каждый корень считать столько раз, какова его кратность. Отсюда в соответствии с теоремой 86.5 вытекают следующие утверждения.

Теорема 86.6. *Произвольный линейный оператор, действующий в n-мерном комплексном пространстве, имеет:*

1) n собственных значений, если каждое собственное значение считать столько раз, какова его кратность как корня характеристического многочлена;

2) хотя бы один собственный вектор;

3) на любом своем инвариантном подпространстве хотя бы один собственный вектор.

Последнее утверждение следует из того, что индуцированный оператор, как и любой оператор, действующий в комплексном пространстве, имеет собственный вектор, который, очевидно, является собственным вектором основного оператора.

Замечание 2. Теорема 86.6 остается справедливой в вещественном пространстве для тех операторов, чьи характеристические многочлены имеют только вещественные корни.

Способ нахождения собственных векторов определен теоремой 86.5 и состоит в следующем.

1. Строится характеристический многочлен оператора и находятся его корни. Те корни, которые принадлежат полю P, будут собственными значениями оператора.

2. Для каждого найденного собственного значения λ_0 находятся ненулевые векторы ядра оператора $\mathcal{A} - \lambda_0\mathcal{I}$. Они и будут собственными векторами, отвечающими λ_0. Если $e_1, \ldots, e_n$ – базис пространства V, то поиск собственных векторов сводится к решению однородной системы уравнений $(A_e - \lambda_0 I)x_e = 0$, фундаментальная система решений которой дает полный набор линейно независимых собственных векторов, отвечающих λ_0, в координатах относительно базиса e.

Замечание 3. Прием, описанный здесь, представляет собой лишь теоретическую основу решения проблемы собственных значений и собственных векторов. Для реальных вычислений он нуждается в специальных приемах, так как поиск корней многочлена в общем случае (исключая разве что многочлены второй степени) сопряжен с дополнительными трудностями. Уже для многочленов степени $n \geq 5$ задача нахождения корней неразрешима в радикалах (доказано Э.Галуа в 30-х гг. прошлого столетия). Этот результат оказал решающее влияние на развитие вычислительных методов решения проблемы собственных значений. Их изложение не входит в нашу задачу, однако отметим, что вычислительные алгоритмы поиска собственных значений, как правило, не связаны с многочленами.

В ручных вычислениях иногда удается найти собственные значения, минуя прямое построение характеристического многочлена, если вычислять определитель матрицы $A - \lambda I$ *методом выделения линейных множителей*. Этот метод основан на выполнении элементарных преобразований, формирующих в какой-либо строке (столбце) матрицы $A - \lambda I$ общий множитель вида $\lambda - a$, который затем выносится за знак определителя. Это дает корень характеристического многочлена $\lambda = a$.

§87. Собственное подпространство

Пусть λ_0 – собственное значение оператора $\mathcal{A}$. Множество

$$W_{\lambda_0} = \{\, x \in V \mid \mathcal{A}x = \lambda_0 x \,\} \tag{87.1}$$

называется *собственным подпространством* оператора $\mathcal{A}$, отвечающим собственному значению λ_0.

Очевидно, что $W_{\lambda_0} = \ker(\mathcal{A} - \lambda_0\mathcal{I})$, поэтому собственное подпространство является линейным подпространством пространства V.

легко проверить, $\mathcal{A}e_i = g_i$, $i = \overline{1, n}$. Линейность построенного оператора вытекает из линейности координат.

Оператор $\mathcal{A}$ единствен, так как если $\mathcal{B}$ – любой другой линейный оператор, переводящий векторы $e_1, \ldots, e_n$ в $g_1, \ldots, g_n$, то

$$\mathcal{B}x = \mathcal{B}\Big(\sum_{i=1}^n x_i e_i\Big) = \sum_{i=1}^n x_i \mathcal{B}e_i = \sum_{i=1}^n x_i g_i = \mathcal{A}x, \quad \forall x \in V.$$

Следовательно, $\mathcal{B} = \mathcal{A}$. ■

Следствие. Линейные операторы $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V, W)$ равны тогда и только тогда, когда они совпадают на векторах базиса V.

§77. Матрица линейного оператора

Пусть $e = (e_1, \ldots, e_n)$ и $f = (f_1, \ldots, f_m)$ – базисы пространств V и W. Как следует из теоремы 76.1, линейный оператор $\mathcal{A} \in \mathcal{L}(V, W)$ однозначно определяется заданием векторов $\mathcal{A}e_1, \ldots, \mathcal{A}e_n$. В свою очередь, векторы $\mathcal{A}e_i$, $i = \overline{1, n}$, однозначно определяются своими координатами в базисе f, т.е. коэффициентами разложений

$$\begin{cases} \mathcal{A}e_1 = a_{11}f_1 + a_{21}f_2 + \ldots + a_{m1}f_m, \\ \mathcal{A}e_2 = a_{12}f_1 + a_{22}f_2 + \ldots + a_{m2}f_m, \\ \ldots \\ \mathcal{A}e_n = a_{1n}f_1 + a_{2n}f_2 + \ldots + a_{mn}f_m. \end{cases} \tag{77.1}$$

Матрица

$$A_{fe} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \tag{77.2}$$

называется *матрицей оператора $\mathcal{A}$ в паре базисов e и f*. Для обозначения этой матрицы используется также символ $(\mathcal{A})_{fe}$.

Из единственности разложения вектора по базису следует, что при фиксированных базисах e и f матрица линейного оператора определена однозначно.

Теорема 77.1. *Пусть* $\dim V = n$, $\dim W = m$. *Тогда существует взаимно однозначное соответствие между линейными операторами из $\mathcal{L}(V, W)$ и матрицами из $P^{m\times n}$.*

Доказательство. Построим это соответствие. Зафиксируем базисы $e = (e_1, \ldots, e_n)$ и $f = (f_1, \ldots, f_m)$ пространств V и W. Поставим в соответствие каждому линейному оператору $\mathcal{A} \in \mathcal{L}(V, W)$ его матрицу A_{fe} в паре базисов e и f. Очевидно, что матрица $A_{fe} \in P^{m\times n}$ определена однозначно. Докажем биективность построенного таким образом отображения. Действительно, оно:

1) сюръективно, так как любая матрица $B = (b_{ij}) \in P^{m\times n}$ является матрицей линейного оператора из $\mathcal{L}(V, W)$, переводящего векторы

e_j в векторы $\sum_{i=1}^m b_{ij} f_i$, $j = \overline{1, n}$ (в силу теоремы 76.1 такой оператор существует);

2) инъективно, ибо различные операторы из $\mathcal{L}(V, W)$ не совпадают на базисных векторах и, значит, имеют разные матрицы. ■

Доказанная теорема играет важную роль в теории линейных операторов. Она позволяет описывать свойства линейных операторов через известный нам аппарат теории матриц. В дальнейшем мы увидим, что связь между операторами и матрицами более тесная, чем взаимно однозначное соответствие.

Координаты вектора и его образа. Пусть $\mathcal{A} \in \mathcal{L}(V, W)$, e и f – базисы пространств V и W.

Теорема 77.2. *Если $y = \mathcal{A}x$, то*

$$y_f = A_{fe}x_e. \tag{77.3}$$

Доказательство. Пусть $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^m y_i f_i$ и $A_{fe} = (a_{ij})$. Утверждение (77.3) равносильно соотношениям

$$y_i = \sum_{j=1}^n a_{ij}x_j, \quad i = \overline{1, m}. \tag{77.4}$$

Докажем их. Имеем $y = \mathcal{A}x = \mathcal{A}\big(\sum_{j=1}^n x_j e_j\big) = \sum_{j=1}^n x_j \mathcal{A}e_j = \{$в силу $(77.1)\} = \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} f_i = \sum_{i=1}^m \big(\sum_{j=1}^n a_{ij}x_j\big) f_i$. Из единственности разложения вектора y по базису f следует (77.4). ■

Матрицы оператора в различных базисах. Пусть e и $t = eC$ – два базиса пространства V с матрицей перехода C, а f и $s = fD$ – два базиса пространства W с матрицей перехода D. Одному и тому же линейному оператору $\mathcal{A} \in \mathcal{L}(V, W)$ в паре базисов e и f соответствует матрица A_{fe}, а в паре базисов t и s – матрица A_{st}.

Теорема 77.3. *Матрицы A_{fe} и A_{st} линейного оператора $\mathcal{A} \in \mathcal{L}(V, W)$ в различных парах базисов связаны соотношением*

$$A_{st} = D^{-1}A_{fe}C. \tag{77.5}$$

Доказательство. Для произвольного вектора $x \in V$ и его образа $y = \mathcal{A}x$ в силу (77.3) имеем

$$y_f = A_{fe}x_e, \quad y_s = A_{st}x_t. \tag{77.6}$$

В свою очередь, согласно (17.8), $x_e = Cx_t$, $y_f = Dy_s$. Подставив эти соотношения в (77.6), получим, что $Dy_s = A_{fe}Cx_t$ или $DA_{st}x_t = A_{fe}Cx_t$. Так как это соотношение имеет место для любых x_t, то $DA_{st} = A_{fe}C$. В силу невырожденности матрицы перехода отсюда следует (77.5). ■

Следствие 1. *Матрицы линейного оператора в различных парах базисов эквивалентны* (§16).

Так как V – комплексное пространство, то, согласно теореме 86.6, $\dim V = m_1 + \ldots + m_p$. С другой стороны, $0 < \dim W_{\lambda_k} = s_k \leq m_k$, $k = \overline{1, p}$, поэтому равенство (88.3) возможно тогда и только тогда, когда

$$s_k = m_k, \quad k = \overline{1, p}. \quad ■ \tag{88.4}$$

Замечание 2. Теорема 88.3 в вещественном пространстве верна для тех операторов, чьи характеристические многочлены имеют только вещественные корни.

Матричная формулировка операторных свойств. В точности так же, как и для линейных операторов $\mathcal{A} \in \mathcal{L}(V, V)$, определяются собственные значения и собственные векторы квадратной матрицы $A \in P^{n\times n}$. Ненулевой вектор-столбец $x \in P^n$ называется *собственным вектором матрицы* $A \in P^{n\times n}$, если существует число $\lambda \in P$ такое, что

$$Ax = \lambda x.$$

При этом число λ называется *собственным значением матрицы A*, соответствующим собственному вектору x.

Если $e = (e_1, \ldots, e_n)$ – произвольный базис пространства V, то для оператора $\mathcal{A} \in \mathcal{L}(V, V)$ имеет место импликация

$$\mathcal{A}x = \lambda x \iff A_e x_e = \lambda x_e. \tag{88.5}$$

Это означает, что собственные значения оператора $\mathcal{A}$ и его матрицы в любом базисе $e = (e_1, \ldots, e_n)$ совпадают, а собственные векторы матрицы A_e являются координатными столбцами собственных векторов оператора $\mathcal{A}$ в этом базисе.

Характеристические многочлены оператора и его матрицы совпадают по определению, поэтому для квадратных матриц $A \in P^{n\times n}$ имеют место теоремы 86.5, 86.6.

Эта идентичность свойств оператора и его матрицы, замеченная уже в связи с первыми свойствами оператора (§77), продолжает наблюдаться.

Квадратная матрица $A \in P^{n\times n}$ называется *матрицей простой структуры*, если она имеет n линейно независимых собственных векторов.

Из (88.5) следует, что линейный оператор является оператором простой структуры тогда и только тогда, когда его матрица в любом базисе имеет простую структуру.

Теорема 88.4 (матричная формулировка теоремы 88.1). *Квадратная матрица является матрицей простой структуры тогда и только тогда, когда она подобна диагональной.*

Доказательство. Пусть $A \in P^{n\times n}$ – заданная матрица. Рассмотрим произвольное линейное пространство V над полем P размерности n. Зафиксируем в пространстве V произвольный базис f. Пусть $\mathcal{A} \in \mathcal{L}(V, V)$ – линейный оператор, матрица которого в базисе

f совпадает с матрицей A, так что $A = A_f$ (согласно теореме 77.1 такой оператор существует). Матрица $A = A_f$ имеет простую структуру тогда и только тогда, когда $\mathcal{A}$ – оператор простой структуры. В силу теоремы 88.1 это равносильно существованию базиса e, в котором оператор $\mathcal{A}$ имеет диагональную матрицу (88.1). Таким образом, матрица A имеет простую структуру тогда и только тогда, когда она и диагональная матрица (88.1) являются матрицами одного оператора. Это равносильно их подобию (§82). ■

Замечание 3. Доказательство теоремы можно схематически изобразить в виде лаконичной диаграммы, которой мы неоднократно будем пользоваться:

$$\begin{tikzcd} & \mathcal{A} \in \mathcal{L}(V,V) \arrow[dr, "e"] & \\ A = A_f \arrow[ur, "f"] \arrow[rr, phantom, "\cong"] & & A_e = [\,\ddots\,]. \end{tikzcd} \tag{88.6}$$

Замечание 4. Теорема может быть доказана и на матричном языке. Действительно, матрица A является матрицей простой структуры тогда и только тогда, когда она имеет n линейно независимых собственных векторов $s_1, \ldots, s_n \in P^n$. Это равносильно следующей цепочке равенств: если $S = [s_1 \ldots s_n]$ – матрица, столбцами которой являются собственные векторы $s_1, \ldots, s_n$, Λ – диагональная матрица с собственными значениями $\lambda_1, \ldots, \lambda_n$ на главной диагонали, то

$$\begin{cases} As_1 = \lambda_1 s_1 \\ \ldots \\ As_n = \lambda_n s_n \end{cases} \iff A[s_1 \ldots s_n] = [s_1 \ldots s_n]\Lambda \iff AS = S\Lambda \iff$$
$$\iff A = S\Lambda S^{-1}.$$

Жорданова клетка. Итак, самой 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 собственных векторов. Рассмотрим важный пример.

Матрица

$$J_k(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda_0 & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_0 & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda_0 \end{bmatrix} \tag{88.7}$$

называется *жордановой клеткой k-го порядка*. Эта матрица имеет:

$(\mathcal{B}\mathcal{A})(x+y) = \mathcal{B}(\mathcal{A}(x+y)) = \mathcal{B}(\mathcal{A}x + \mathcal{A}y) = \mathcal{B}(\mathcal{A}x) + \mathcal{B}(\mathcal{A}y) = \mathcal{B}\mathcal{A}x + \mathcal{B}\mathcal{A}y$,
$(\mathcal{B}\mathcal{A})(\alpha x) = \mathcal{B}(\mathcal{A}(\alpha x)) = \mathcal{B}(\alpha(\mathcal{A}x)) = \alpha\mathcal{B}(\mathcal{A}x) = \alpha(\mathcal{B}\mathcal{A}x) = (\alpha\mathcal{B}\mathcal{A})x$. ■

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

$$\begin{aligned} &1)\ (\mathcal{A}\mathcal{B})\mathcal{C} = \mathcal{A}(\mathcal{B}\mathcal{C}) \quad \text{(ассоциативность)}; \\ &2)\ \alpha(\mathcal{A}\mathcal{B}) = (\alpha\mathcal{A})\mathcal{B} = \mathcal{A}(\alpha\mathcal{B}); \\ &3)\ (\mathcal{A} + \mathcal{B})\mathcal{C} = \mathcal{A}\mathcal{C} + \mathcal{B}\mathcal{C}, \\ &\quad\ \mathcal{A}(\mathcal{B} + \mathcal{C}) = \mathcal{A}\mathcal{B} + \mathcal{A}\mathcal{C} \quad \text{(дистрибутивность)}. \end{aligned} \tag{79.1}$$

Эти свойства легко проверяются непосредственно.

Умножение линейных операторов не обладает свойством коммутативности. В самом деле, о коммутативности можно говорить лишь для операторов, действующих в одном пространстве, т.е. для операторов $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V, V)$. Но и в этом случае умножение не коммутативно. В этом можно убедиться на простом примере операторов ортогонального проектирования плоскости V_2 на ось Ox и поворота на угол $\pi/2$.

Теорема 79.2. *При умножении линейных операторов их матрицы умножаются, т.е. если e, f, g – базисы пространств V, W, Z, то*

$$(\mathcal{B}\mathcal{A})_{ge} = B_{gf}A_{fe}. \tag{79.2}$$

Доказательство. Пусть $A_{fe} = (a_{ij})$, $B_{gf} = (b_{ij})$, $(\mathcal{B}\mathcal{A})_{ge} = (c_{ij})$, $\dim V = n$, $\dim W = m$, $\dim Z = k$. Тогда в силу (77.1)

$$\mathcal{B}\mathcal{A}e_j = \sum_{i=1}^k c_{ij}g_i. \tag{79.3}$$

В то же время $\mathcal{B}\mathcal{A}e_j = \mathcal{B}(\mathcal{A}e_j) = \mathcal{B}\big(\sum_{s=1}^m a_{sj}f_s\big) = \sum_{s=1}^m a_{sj}(\mathcal{B}f_s) = \sum_{s=1}^m a_{sj}\sum_{i=1}^k b_{is}g_i = \sum_{s=1}^m\sum_{i=1}^k a_{sj}b_{is}g_i = \sum_{i=1}^k\big(\sum_{s=1}^m b_{is}a_{sj}\big)g_i$.

Сравнение этого разложения с (79.3) приводит к равенству $c_{ij} = \sum_{s=1}^m b_{is}a_{sj}$, которое означает (79.2). ■

§80. Образ и ядро линейного оператора

Образом линейного оператора $\mathcal{A} \in \mathcal{L}(V, W)$ называется множество

$$\operatorname{im}\mathcal{A} = \{y \in W \mid \exists x \in V : \mathcal{A}x = y\},$$

ядром оператора $\mathcal{A}$ – множество

$$\ker\mathcal{A} = \{x \in V \mid \mathcal{A}x = \theta\}.$$

Примеры. 1. В пространстве многочленов M_n для оператора дифференцирования (76.2): $\operatorname{im}\mathcal{D} = M_{n-1}$, $\ker\mathcal{D} = M_0$.

2. Для оператора проектирования (76.3): $\operatorname{im}\mathcal{P} = L_1$, $\ker\mathcal{P} = L_2$.

3. Для оператора отражения (76.4): $\operatorname{im}\mathcal{R} = V$, $\ker\mathcal{R} = \{\theta\}$.

Теорема 80.1. *Если $\mathcal{A} \in \mathcal{L}(V, W)$, то $\ker\mathcal{A}$ – линейное подпространство пространства V, $\operatorname{im}\mathcal{A}$ – линейное подпространство пространства W.*

Эти утверждения вытекают из теоремы 18.1, условия которой легко проверяются. ■

Рангом линейного оператора называется размерность его образа, а *дефектом* – размерность ядра. Обозначения: $\operatorname{rg}\mathcal{A}$, $\operatorname{def}\mathcal{A}$. Итак, $\operatorname{rg}\mathcal{A} = \dim\operatorname{im}\mathcal{A}$, $\operatorname{def}\mathcal{A} = \dim\ker\mathcal{A}$.

Теорема 80.2. *Если $e_1, \ldots, e_n$ – базис пространства V, то*

$$\operatorname{im}\mathcal{A} = \mathcal{L}(\mathcal{A}e_1, \ldots, \mathcal{A}e_n). \tag{80.1}$$

Доказательство. Достаточно показать, что для множеств (80.1) имеет место двустороннее вложение: с одной стороны, если $y \in \operatorname{im}\mathcal{A}$, то $y = \mathcal{A}x$ для некоторого вектора $x \in V$, т.е. $y = \mathcal{A}\big(\sum_{i=1}^n x_i e_i\big) = \sum_{i=1}^n x_i\mathcal{A}e_i \in \mathcal{L}(\mathcal{A}e_1, \ldots, \mathcal{A}e_n)$; с другой стороны, если $y \in \mathcal{L}(\mathcal{A}e_1, \ldots, \mathcal{A}e_n)$, то $y = \sum_{i=1}^n x_i\mathcal{A}e_i = \mathcal{A}\big(\sum_{i=1}^n x_i e_i\big) = \mathcal{A}x$, где $x = \sum_{i=1}^n x_i e_i$, т.е. $y \in \operatorname{im}\mathcal{A}$. ■

Теорема 80.3. *Ранг линейного оператора равен рангу его матрицы в произвольной паре базисов.*

Доказательство. Из теоремы 80.2 и (64.3) следует, что $\operatorname{rg}\mathcal{A} = \dim\operatorname{im}\mathcal{A} = \dim\mathcal{L}(\mathcal{A}e_1, \ldots, \mathcal{A}e_n) = \operatorname{rg}(\mathcal{A}e_1, \ldots, \mathcal{A}e_n)$. Ранг системы векторов $\mathcal{A}e_1, \ldots, \mathcal{A}e_n$ совпадает с рангом системы арифметических векторов, составленных из координат этих векторов в базисе f пространства W, т.е. с рангом системы столбцов матрицы A_{fe}. ■

Теорема 80.4 (о ранге и дефекте). *Если $\mathcal{A} \in \mathcal{L}(V, W)$, то*

$$\operatorname{rg}\mathcal{A} + \operatorname{def}\mathcal{A} = \dim V. \tag{80.2}$$

Доказательство. Пусть $e_1, \ldots, e_k$ – базис $\ker\mathcal{A}$. Дополним его до базиса $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$ пространства V. Согласно теореме 80.2, $\operatorname{im}\mathcal{A} = \mathcal{L}(\mathcal{A}e_1, \ldots, \mathcal{A}e_k, \mathcal{A}e_{k+1}, \ldots, \mathcal{A}e_n) = \mathcal{L}(\mathcal{A}e_{k+1}, \ldots, \mathcal{A}e_n)$. Докажем, что векторы $\mathcal{A}e_{k+1}, \ldots, \mathcal{A}e_n$ линейно независимы. Пусть это не так. Тогда для нетривиальной линейной комбинации этих векторов имеют место соотношения $\alpha_{k+1}\mathcal{A}e_{k+1} + \ldots + \alpha_n\mathcal{A}e_n = \theta$ или $\mathcal{A}(\alpha_{k+1}e_{k+1} + \ldots + \alpha_n e_n) = \theta$. Следовательно, $\alpha_{k+1}e_{k+1} + \ldots + \alpha_n e_n \in \ker\mathcal{A}$. Это означает, что вектор $\alpha_{k+1}e_{k+1} + \ldots + \alpha_n e_n$ линейно выражается через $e_1, \ldots, e_k$, что невозможно в силу линейной независимости $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$. Таким образом, $\dim\operatorname{im}\mathcal{A} = n - k$, $\dim\ker\mathcal{A} = k$. Отсюда следует (80.2). ■

Теорема 80.5. *Пусть $\mathcal{A} \in \mathcal{L}(V, W)$, $\operatorname{rg}\mathcal{A} = r$, $\dim V = n$, $\dim W = m$. Тогда существуют базисы e и f пространств V и W,*

векторы

$$x_0, \mathcal{A}x_0, \ldots, \mathcal{A}^{q-1}x_0 \tag{90.1}$$

линейно независимы.

Доказательство. Рассмотрим равенство $\alpha_0 x_0 + \alpha_1\mathcal{A}x_0 + \ldots + \alpha_{q-1}\mathcal{A}^{q-1}x_0 = \theta$. Применяя последовательно операторы $\mathcal{A}^{q-1}$, $\mathcal{A}^{q-2}$, $\ldots$, $\mathcal{A}$ к обеим частям этого равенства, получим, что $\alpha_0 = \alpha_1 = \ldots = \alpha_{q-1} = 0$. Это доказывает линейную независимость системы векторов (90.1). ■

Следствие 1. *Индекс нильпотентности не превосходит размерности пространства.*

Теорема 90.2. *В комплексном пространстве линейный оператор нильпотентен тогда и только тогда, когда все его собственные значения равны нулю.*

Доказательство. Необходимость. Если λ – собственное значение нильпотентного оператора $\mathcal{A} \in \mathcal{L}(V, V)$ индекса q и x – соответствующий собственный вектор, то, как нетрудно проверить, имеет место цепочка импликаций

$$\mathcal{A}x = \lambda x \Rightarrow \mathcal{A}^2x = \lambda^2x \Rightarrow \ldots \Rightarrow \mathcal{A}^qx = \lambda^qx.$$

Отсюда следует, что $\lambda^q x = \theta$. Так как $x \neq \theta$, то $\lambda = 0$.

Достаточность. Рассмотрим базис e комплексного пространства V, в котором оператор $\mathcal{A}$ имеет верхнюю треугольную матрицу (§89), главная диагональ которой состоит целиком из нулей (§89, замечание 2). Итак,

$$A_e = \begin{bmatrix} 0 & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & 0 & a_{23} & \ldots & a_{2n} \\ & & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{n-1,n} \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Как нетрудно проверить, при последовательном возведении этой матрицы в степени $q = 2, 3, \ldots, n$ нетривиальный треугольник, расположенный над главной диагональю, перемещается каждый раз на одну диагональ выше, так что $(A_e)^n = O$. Значит, $\mathcal{A}^n = \mathcal{O}$. ■

Замечание. Необходимость этого утверждения, как видно из самого доказательства, имеет место и в вещественном пространстве.

Прямая сумма операторов. Если $V = L_1 \oplus L_2 \oplus \ldots \oplus L_p$ – прямая сумма подпространств $L_1, L_2, \ldots, L_p$, инвариантных относительно линейного оператора $\mathcal{A} \in \mathcal{L}(V, V)$, то оператор $\mathcal{A}$ называется *прямой суммой индуцированных операторов* $\mathcal{A}|L_1, \ldots, \mathcal{A}|L_p$. Этот термин мотивирован тем обстоятельством, что для любого вектора $x \in V$ с разложением $x = x_1 + \ldots + x_p$, где $x_k \in L_k$, $k = \overline{1, p}$, имеет место равенство $\mathcal{A}x = \mathcal{A}x_1 + \ldots + \mathcal{A}x_p = (\mathcal{A}|L_1)x_1 + \ldots + (\mathcal{A}|L_p)x_p$.

Теорема 90.3. *Произвольный линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ является прямой суммой нильпотентного и обратимого операторов, причем это разложение единственно.*

Доказательство. Для доказательства теоремы необходимо показать, что существует единственная пара подпространств L_1, L_2, ин-

вариантных относительно линейного оператора $\mathcal{A}$ и таких, что $V = L_1 \oplus L_2$, $\mathcal{A}|L_1$ нильпотентен, $\mathcal{A}|L_2$ обратим. Будем считать, что оператор $\mathcal{A}$ вырожден (для невырожденного $\mathcal{A}$: $L_1 = \{\theta\}$, $L_2 = V$) и не нильпотентен (для нильпотентного $\mathcal{A}$: $L_1 = V$, $L_2 = \{\theta\}$).

Существование. Обозначим для $k \in \mathbb{N}$

$$N_k = \ker\mathcal{A}^k, \quad T_k = \operatorname{im}\mathcal{A}^k.$$

1. Покажем, что *подпространства N_k строго вложены друг в друга до некоторого момента q, начиная с которого все N_k совпадают*, т.е.

$$N_1 \subset N_2 \subset \ldots \subset N_q = N_{q+1} = \ldots \tag{90.2}$$

а) Вложение $N_k \subseteq N_{k+1}$ очевидно, так как если $\mathcal{A}^kx = \theta$, то $\mathcal{A}^{k+1}x = \mathcal{A}(\mathcal{A}^k) = \mathcal{A}\theta = \theta$.

б) Пусть $N_k = N_{k+1}$, тогда $N_{k+1} = N_{k+2}$, так как $N_{k+1} \subseteq N_{k+2}$, $N_{k+2} \subseteq N_{k+1}$. Второе из этих вложений следует из того, что если $x \in N_{k+2}$, то $\mathcal{A}^{k+2}x = \theta$, т.е. $\mathcal{A}^{k+1}(\mathcal{A}x) = \theta$. Значит, $\mathcal{A}x \in N_{k+1} = N_k$, следовательно, $\mathcal{A}^k(\mathcal{A}x) = \theta$, т.е. $\mathcal{A}^{k+1}x = \theta$.

Из "а" и "б" следует, что подпространство N_k либо строго вложено в N_{k+1}, либо совпадает со всеми последующими ядрами. Так как в конечномерном пространстве размерности подпространств N_k не могут бесконечно возрастать, то наступит момент q, начиная с которого все ядра N_k будут совпадать с N_q.

2. Зафиксируем этот момент q и покажем, что

$$V = N_q \oplus T_q.$$

Действительно, $\dim V = \dim N_q + \dim T_q$ в силу теоремы 80.4 о ранге и дефекте, при этом $N_q \cap T_q = \{\theta\}$, так как если $y \in N_q$, $y \in T_q$, то $\mathcal{A}^qy = \theta$, $y = \mathcal{A}^qx$, т.е. $\mathcal{A}^{2q}x = \theta$. Значит, $x \in N_{2q} = N_q$ и $\mathcal{A}^qx = \theta$, откуда следует, что $y = \theta$.

3. Подпространства N_q, T_q *инвариантны относительно $\mathcal{A}$*, ибо:

а) если $x \in N_q$, то $x \in N_{q+1} = N_q$, поэтому $\mathcal{A}^{q+1}x = \theta$, т.е. $\mathcal{A}^q(\mathcal{A}x) = \theta$, следовательно, $\mathcal{A}x \in N_q$;

б) если $y \in T_q$, то $y = \mathcal{A}^qx$ и $\mathcal{A}y = \mathcal{A}^{q+1}x = \mathcal{A}^q(\mathcal{A}x) = \mathcal{A}^qx_1$, где $x_1 = \mathcal{A}x$, следовательно, $\mathcal{A}y \in T_q$.

4. Оператор $\mathcal{A}|N_q$ – *нильпотентный оператор индекса q*, так как:

а) $\mathcal{A}^qx = \theta$ для любого вектора $x \in N_q$;

б) существует вектор $x_0 \in N_q$ такой, что $\mathcal{A}^{q-1}x_0 \neq \theta$, ибо $N_{q-1} \neq N_q$.

5. Оператор $\mathcal{A}|T_q$ *обратим*, так как его ядро состоит только из нулевого вектора. Действительно, если $y \in \ker\mathcal{A}|T_q$, то $y \in T_q$, $\mathcal{A}y = \theta$, т.е. $y = \mathcal{A}^qx$ и $\mathcal{A}^{q+1}x = \theta$. Отсюда следует, что $x \in N_{q+1} = N_q$, т.е. $\mathcal{A}^qx = \theta$ и $y = \theta$.

Утверждения пунктов 2 – 5 доказывают существование искомого разложения: $L_1 = N_q$, $L_2 = T_q$.

Единственность. Пусть существует другое разложение $V = N \oplus T$, обладающее всеми свойствами первого.

Пусть V – линейное пространство над полем P. Рассмотрим множество $\mathcal{L}(V, V)$ всех линейных операторов, действующих в пространстве V. В этом множестве для любых операторов выполнимы не только сложение и умножение на число, но и умножение операторов друг на друга. Из общей теории линейных операторов вытекают следующие факты.

1°. *$\mathcal{L}(V, V)$ – линейное пространство над полем P* (теорема 78.2).

2°. *$\mathcal{L}(V, V)$ – некоммутативное кольцо с единицей* (см. (79.1)).

3°. Линейное пространство над полем P, которое является кольцом и удовлетворяет условию 2 из (79.1), называется *алгеброй* (или *линейной алгеброй*) *над полем P*. Размерность линейного пространства при этом называется также *размерностью алгебры*. Из свойства 1° и (79.1) следует, что *$\mathcal{L}(V, V)$ – алгебра над полем P*.

Отметим, что множество $P^{n\times n}$ квадратных матриц n-го порядка над полем P также является алгеброй. Примером бесконечномерной алгебры служит кольцо многочленов $P[x]$.

Может быть построена другая алгебра, в которой нарушена как коммутативность, так и ассоциативность произведения. Примером такой алгебры является множество всех векторов геометрического пространства V_3 относительно линейных операций над векторами и векторного произведения (§23).

4°. При построении матрицы линейного оператора $\mathcal{A} \in \mathcal{L}(V, V)$ используется один базис e пространства V: столбцами этой матрицы являются коэффициенты разложений векторов $\mathcal{A}e_1, \ldots, \mathcal{A}e_n$ по базису $e_1, \ldots, e_n$ (§77); матрица обозначается символом A_e или $(\mathcal{A})_e$ и называется *матрицей оператора $\mathcal{A}$ в базисе e*. Очевидно, $A_e \in P^{n\times n}$.

5°. *Линейное пространство $\mathcal{L}(V, V)$ изоморфно пространству $P^{n\times n}$ и $\dim\mathcal{L}(V, V) = n^2$* (теорема 78.3 и ее следствие).

6°. Как и в любом кольце, оператор $\mathcal{A} \in \mathcal{L}(V, V)$ можно возводить в степень $n \in \mathbb{N}$, и если $p(t) = a_0 + a_1t + \ldots + a_nt^n$ – произвольный многочлен над полем P от переменной t, то однозначно определен оператор

$$p(\mathcal{A}) = a_0I + a_1\mathcal{A} + \ldots + a_n\mathcal{A}^n, \tag{82.1}$$

называемый *многочленом от оператора $\mathcal{A}$*. Из свойств матрицы линейного оператора следует, что матрицей многочлена (82.1) от оператора $\mathcal{A}$ является тот же многочлен от матрицы A_e: $(p(\mathcal{A}))_e = p(A_e)$.

7°. При переходе от базиса e к базису $f = eQ$ матрица оператора изменяется согласно (77.5) по следующему закону:

$$A_f = Q^{-1}A_eQ. \tag{82.2}$$

Из (82.2) и (58.6) следует, что *одному и тому же линейному оператору $\mathcal{A} \in \mathcal{L}(V, V)$ соответствует целый класс матриц, подобных друг другу.*

8°. Очевидно, что *две матрицы $A, B \in P^{n\times n}$ подобны тогда и только тогда, когда они являются матрицами одного и того же линейного оператора, действующего в n-мерном линейном пространстве над полем P* (теорема 77.4).

9°. Из (82.2) следует, что *все матрицы одного и того же линейного оператора имеют одинаковый определитель*. Это означает, что определитель матрицы линейного оператора не зависит от матрицы и определяется самим оператором. *Определителем линейного оператора* $\mathcal{A} \in \mathcal{L}(V, V)$ называется определитель матрицы этого оператора в произвольном базисе. Обозначение: $\det\mathcal{A}$. Итак,

$$\det\mathcal{A} = \det A_e. \tag{82.3}$$

Заметим, что кольцо $\mathcal{L}(V, V)$ не является полем хотя бы потому, что оно некоммутативно. К тому же вопрос об обратимости линейных операторов пока остается открытым.

§83. Обратный оператор

Напомним определение обратного отображения и некоторые факты (§8) применительно к линейному оператору, действующему в одном пространстве.

Пусть $\mathcal{A} \in \mathcal{L}(V, V)$. Отображение $\mathcal{A}^{-1} : V \to V$ называется *обратным оператором* к оператору $\mathcal{A}$, если

$$\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} = \mathcal{I}. \tag{83.1}$$

Теорема 83.1. *Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ обратим тогда и только тогда, когда он биективен.*

Теорема 83.2. *Обратный оператор единствен.*

Исследуем теперь свойства обратимости, относящиеся непосредственно к линейным операторам.

Теорема 83.3. *Обратный оператор линеен.*

Доказательство. Пусть $\mathcal{A} \in \mathcal{L}(V, V)$. Покажем, что обратный оператор $\mathcal{A}^{-1}$, если он существует, является линейным оператором, действующим в пространстве V. Действительно, если $\mathcal{A}$ обратим, то он биективен и, значит, сюръективен. Это означает, что для любых векторов $y_1, y_2 \in V$ существуют $x_1, x_2 \in V$ такие, что $y_1 = \mathcal{A}x_1$, $y_2 = \mathcal{A}x_2$. При этом $x_1 = \mathcal{A}^{-1}y_1$, $x_2 = \mathcal{A}^{-1}y_2$. Отсюда получим, что $\mathcal{A}^{-1}(y_1 + y_2) = \mathcal{A}^{-1}(\mathcal{A}x_1 + \mathcal{A}x_2) = \mathcal{A}^{-1}\mathcal{A}(x_1 + x_2) = x_1 + x_2 = \mathcal{A}^{-1}y_1 + \mathcal{A}^{-1}y_2$. Аналогично, $\mathcal{A}^{-1}(\alpha y_1) = \mathcal{A}^{-1}(\alpha\mathcal{A}x_1) = \mathcal{A}^{-1}\mathcal{A}(\alpha x_1) = \alpha x_1 = \alpha\mathcal{A}^{-1}y_1$, $\forall\alpha \in P$. ■

Теорема 83.4. *Матрица обратного оператора $\mathcal{A}^{-1}$ в произвольном базисе является обратной к матрице оператора $\mathcal{A}$ в этом же базисе.*

Доказательство. Пусть e – произвольный базис пространства V и для оператора $\mathcal{A} \in \mathcal{L}(V, V)$ существует обратный оператор $\mathcal{A}^{-1}$.

Перейдем в равенствах (83.1) к матрицам операторов в базисе e. Согласно (79.2) получим, что $A_e(\mathcal{A}^{-1})_e = (\mathcal{A}^{-1})_eA_e = I$. Эти равенства совпадают с определением обратной матрицы для матрицы A_e

любая квадратная комплексная матрица подобна квазидиагональной матрице, обладающей указанным выше свойством.

Замечание. Разложение (91.2), обладающее всеми свойствами из теоремы 91.1, для каждого оператора единственно в силу теоремы 90.3.

Корневые векторы. Пусть λ_j – собственное значение оператора $\mathcal{A}$. Вектор $x \in V$ называется *корневым вектором* оператора $\mathcal{A}$, *отвечающим собственному значению* λ_j, если $(\mathcal{A} - \lambda_j\mathcal{I})^kx = \theta$ при некотором $k \in \mathbb{Z}$, $k \geq 0$. *Высотой* корневого вектора называется наименьшее k, обладающее указанным свойством. Очевидно, высота нулевого корневого вектора (и только нулевого) равна нулю.

Отметим простейшие свойства корневых векторов, вытекающие из определения.

1°. *Корневые векторы высоты* 1 *являются собственными векторами*, так как удовлетворяют условиям $(\mathcal{A} - \lambda_j\mathcal{I})x = \theta$, $x \neq \theta$.

2°. *Если x – корневой вектор высоты $k > 0$, то вектор $(\mathcal{A} - \lambda_j\mathcal{I})x$ является корневым вектором высоты $k - 1$.*

3°. *Если x – корневой вектор высоты $k > 0$, то векторы*

$$x, (\mathcal{A} - \lambda_j\mathcal{I})x, \ldots, (\mathcal{A} - \lambda_j\mathcal{I})^{k-1}x$$

линейно независимы. Доказательство этого утверждения аналогично доказательству теоремы 90.1.

4°. *Корневые векторы различных высот линейно независимы.*

Корневые векторы высоты $k > 1$ называются *присоединенными векторами* $(k-1)$*-го порядка.* Итак, если x – присоединенный вектор $(k-1)$-го порядка, то $(\mathcal{A} - \lambda_j\mathcal{I})^kx = \theta$, $(\mathcal{A} - \lambda_j\mathcal{I})^{k-1}x \neq \theta$ или, другими словами, $(\mathcal{A} - \lambda_j\mathcal{I})^{k-1}x$ – собственный вектор оператора $\mathcal{A}$, отвечающий собственному значению λ_j. Таким образом, корневой вектор – это либо нулевой вектор, либо собственный вектор, либо присоединенный вектор.

Корневые подпространства. Множество всех корневых векторов оператора $\mathcal{A}$, отвечающих собственному значению λ_j, называется *корневым подпространством* оператора $\mathcal{A}$, *отвечающим собственному значению* λ_j. Обозначение: K_{λ_j}. Итак, $K_{\lambda_j} = \{x \in V \mid \exists k \in \mathbb{N} : (\mathcal{A} - \lambda_j\mathcal{I})^kx = \theta\}$. Опишем структуру корневого подпространства.

Пусть $N_k = \ker(\mathcal{A} - \lambda_j\mathcal{I})^k$. Тогда согласно (90.2)

$$N_1 \subset N_2 \subset \ldots \subset N_q = N_{q+1} = \ldots \tag{91.7}$$

1. Очевидно, что N_1 состоит из корневых векторов высоты, не превосходящей 1, т.е. из нулевого вектора θ и всех собственных векторов, отвечающих собственному значению λ_j, т.е. совпадает с собственным подпространством W_{λ_j}. Таким образом,

$$N_1 = W_{\lambda_j} \tag{91.8}$$

и, следовательно,

$$\dim N_1 = s_j, \tag{91.9}$$

где s_j – геометрическая кратность собственного значения λ_j.

2. Подпространство N_2 в (91.7) состоит из корневых векторов высоты, не превосходящей 2, а подпространство N_q состоит из корневых векторов всех высот, т.е. совпадает с корневым подпространством K_{λ_j}. Таким образом, q – максимальная высота корневого вектора, отвечающего собственному значению λ_j, $K_{\lambda_j} = \ker(\mathcal{A} - \lambda_j\mathcal{I})^q$ и

$$W_{\lambda_j} = N_1 \subset N_2 \subset \ldots \subset N_q = K_{\lambda_j}. \tag{91.10}$$

Это означает, что подпространства $K_{\lambda_1}, \ldots, K_{\lambda_p}$, построенные при доказательстве теоремы 91.1, совпадают с корневыми подпространствами, отвечающими собственным значениям $\lambda_1, \ldots, \lambda_p$ соответственно. Следовательно, корневые подпространства обладают свойствами (91.2)–(91.4).

3. Из (91.10) следует, что максимальная высота q корневых векторов, отвечающих собственному значению λ_j, совпадает с индексом нильпотентности оператора $\mathcal{A} - \lambda_j\mathcal{I}$ и согласно свойству 3° не превосходит размерности K_{λ_j}, т.е. алгебраической кратности собственного значения λ_j.

4. Из (91.10) следует также, что собственное подпространство W_{λ_j} является подпространством корневого подпространства K_{λ_j}, так что $s_j \leq m_j$. При этом

$$W_{\lambda_j} = K_{\lambda_j} \iff s_j = m_j. \tag{91.11}$$

Таким образом, цепочка (91.10) дает еще одно доказательство теорем 87.1 и 88.3.

§92. Жорданова форма

Перейдем к построению базиса, в котором матрица линейного оператора имеет наиболее простую форму. Этим базисом будет совокупность построенных специальным образом базисов корневых подпространств.

Канонический базис корневого подпространства. Пусть K_{λ_j} – корневое подпространство оператора $\mathcal{A}$, отвечающее собственному значению λ_j. Положим

$$\mathcal{B} = \mathcal{A} - \lambda_j\mathcal{I}, \ N_k = \ker\mathcal{B}^k, \ n_k = \dim N_k, \ r_k = \operatorname{rg}\mathcal{B}^k.$$

Построим сначала само корневое подпространство K_{λ_j}. Для этого согласно (91.7) необходимо найти момент q, начиная с которого все ядра N_k будут совпадать с $N_q = K_{\lambda_j}$; при этом в силу (91.9), (91.3) $n_1 = s_j < n_2 < \ldots < n_q = m_j$, где s_j и m_j – геометрическая и алгебраическая кратности λ_j.

выражаются только через $e_1, \ldots, e_k$. Таким образом,

$$\begin{cases} \mathcal{A}e_1 = a_{11}e_1 + \ldots + a_{k1}e_k, \\ \ldots \\ \mathcal{A}e_k = a_{1k}e_1 + \ldots + a_{kk}e_k, \\ \mathcal{A}e_{k+1} = a_{1,k+1}e_1 + \ldots + a_{k,k+1}e_k + \ldots + a_{n,k+1}e_n, \\ \ldots \\ \mathcal{A}e_n = a_{1n}e_1 + \ldots + a_{kn}e_k + \ldots + a_{nn}e_n. \end{cases} \tag{84.1}$$

Это означает, что матрица A_e имеет вид

$$A_e = \begin{bmatrix} a_{11} & \ldots & a_{1k} & a_{1,k+1} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} & a_{k,k+1} & \ldots & a_{kn} \\ 0 & \ldots & 0 & a_{k+1,k+1} & \ldots & a_{k+1,n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & a_{n,k+1} & \ldots & a_{nn} \end{bmatrix}$$

и, следовательно, имеет квазитреугольную форму:

$$A_e = \left[\begin{array}{c|c} A_1 & B \\ \hline O & C \end{array}\right]. \ ■ \tag{84.2}$$

Замечание 1. Верна и теорема, обратная доказанной: переход от (84.2) к (84.1) очевиден и приводит к тому, что $L = \mathcal{L}(e_1, \ldots, e_k)$ инвариантно относительно $\mathcal{A}$.

Теорема 84.2. *Если пространство V является прямой суммой подпространств $L_1, \ldots, L_k$, инвариантных относительно оператора $\mathcal{A} \in \mathcal{L}(V, V)$, то в пространстве V существует базис, в котором матрица оператора $\mathcal{A}$ имеет квазидиагональную форму.*

Доказательство аналогично доказательству теоремы 84.1. В качестве искомого базиса берется базис e, составленный из базисов слагаемых подпространств (теорема 66.1). Тогда в силу инвариантности подпространств $L_1, \ldots, L_k$ матрица A_e имеет вид

$$A_e = \begin{bmatrix} \boxed{A_1} & & & O \\ & \boxed{A_2} & & \\ & & \ddots & \\ O & & & \boxed{A_k} \end{bmatrix}. \ ■ \tag{84.3}$$

Замечание 2. Верна и теорема, обратная доказанной.

Индуцированный оператор. Рассматривая линейный оператор только на его инвариантном подпространстве, можно получить новый оператор. Пусть L – подпространство, инвариантное относительно оператора $\mathcal{A} \in \mathcal{L}(V, V)$. Отображение $\mathcal{A}|L : L \to L$, определенное

равенством

$$(\mathcal{A}|L)x = \mathcal{A}x, \quad \forall x \in L,$$

называется *индуцированным оператором, порожденным оператором* $\mathcal{A}$. В силу линейности оператора $\mathcal{A}$ индуцированный оператор также будет линейным. Он совпадает с оператором $\mathcal{A}$ на подпространстве L и не определен вне его. Итак, $\mathcal{A}|L \in \mathcal{L}(L, L)$.

Замечание 3. Из разложений (84.1) следует, что матрицы A_1, $\ldots$, A_k в (84.3) являются матрицами индуцированных операторов $\mathcal{A}|L_1, \ldots, \mathcal{A}|L_k$ в базисах инвариантных подпространств $L_1, \ldots, L_k$.

§85. Собственные значения и собственные векторы

Пусть V – линейное пространство над полем P. Ненулевой вектор $x \in X$ называется *собственным вектором оператора* $\mathcal{A} \in \mathcal{L}(V, V)$, если существует такое число $\lambda \in P$, что

$$\mathcal{A}x = \lambda x. \tag{85.1}$$

Число λ называется *собственным значением оператора $\mathcal{A}$, соответствующим собственному вектору x*. Множество всех собственных значений оператора $\mathcal{A}$ называется *спектром* этого оператора.

Примеры. 1. В пространстве вещественных многочленов M_n любой многочлен нулевой степени будет собственным вектором оператора дифференцирования (76.2), ему соответствует собственное значение $\lambda = 0$.

2. Для оператора проектирования (76.3) любой ненулевой вектор из L_1 будет собственным вектором, отвечающим собственному значению $\lambda = 1$, так как $\mathcal{P}x = x$, $\forall x \in L_1$, а любой ненулевой вектор из L_2 будет собственным вектором, отвечающим собственному значению $\lambda = 0$, так как $\mathcal{P}x = 0x$, $\forall x \in L_2$.

Из определения следует, что если x – собственный вектор оператора $\mathcal{A}$, отвечающий собственному значению λ, то любой вектор αx, где $\alpha \neq 0$, также будет собственным вектором оператора $\mathcal{A}$, отвечающим тому же собственному значению λ. Это означает, что любой собственный вектор порождает целое одномерное подпространство собственных векторов, из которого исключен вектор θ. Таким образом, в отношении собственных векторов понятие "различные" не имеет содержательного смысла, существеннее ставить вопрос об их линейной независимости.

Теорема 85.1. *Собственные векторы $x_1, \ldots, x_k$ оператора, отвечающие различным собственным значениям $\lambda_1, \ldots, \lambda_k$, линейно независимы.*

Доказательство. Применим индукцию по k. Для $k = 1$ утверждение заведомо верно, так как собственный вектор является ненулевым по определению. Пусть оно верно для любой системы из $k - 1$

Теперь будем строить базис K_{λ_j}, последовательно просматривая подпространства $N_q, N_{q-1}, \ldots, N_1$.

N_q) Пусть $f_1, \ldots, f_{t_q}$ – векторы, дополняющие произвольный базис N_{q-1} до базиса N_q. Ясно, что:

1) они будут корневыми векторами высоты q;
2) их количество равно $n_q - n_{q-1}$;
3) $t_q = n_q - n_{q-1} = (n_q - n_{q-1}) - (n_{q+1} - n_q) = -n_{q+1} + 2n_q - n_{q-1}$, так как $n_{q+1} = n_q$;
4) никакая нетривиальная линейная комбинация этих векторов не принадлежит N_{q-1} (такие векторы будем называть *линейно независимыми над* N_{q-1}).

N_{q-1}) Построим векторы $\mathcal{B}f_1, \ldots, \mathcal{B}f_{t_q}$. Эти векторы являются корневыми векторами высоты $q-1$, и они линейно независимы над N_{q-2}, так как в противном случае для нетривиального набора чисел $\alpha_1, \ldots, \alpha_{t_q}$ будем иметь $\mathcal{B}^{q-2}\sum_{k=1}^{t_q}\alpha_k\mathcal{B}f_k = \theta$, т.е. $\mathcal{B}^{q-1}\sum_{k=1}^{t_q}\alpha_k f_k = \theta$, и $\sum_{k=1}^{t_q}\alpha_k f_k \in N_{q-1}$, что противоречит линейной независимости $f_1, \ldots, f_{t_q}$ над N_{q-1}.

Дополним эти векторы векторами $g_1, \ldots, g_{t_{q-1}} \in N_{q-1}$ так, чтобы векторы $\mathcal{B}f_1, \ldots, \mathcal{B}f_{t_q}, g_1, \ldots, g_{t_{q-1}}$ дополняли произвольный базис N_{q-2} до базиса N_{q-1}. Ясно, что:

1) они будут корневыми векторами высоты $q-1$;
2) их количество равно $n_{q-1} - n_{q-2}$;
3) $t_{q-1} = (n_{q-1} - n_{q-2}) - (n_q - n_{q-1}) = -n_q + 2n_{q-1} - n_{q-2}$;
4) они линейно независимы над N_{q-2}.

N_{q-2}) Аналогично строятся векторы
$$\mathcal{B}^2 f_1, \ldots, \mathcal{B}^2 f_{t_q}, \mathcal{B}g_1, \ldots, \mathcal{B}g_{t_{q-1}}, h_1, \ldots, h_{t_{q-2}},$$
дополняющие произвольный базис N_{q-3} до базиса N_{q-2}. Для этой системы векторов справедливы те же факты 1–4, что и для векторов, построенных на предыдущих шагах .

Выполняя далее такие же построения в подпространствах N_{q-3}, N_{q-4}, …, придем к подпространству N_1.

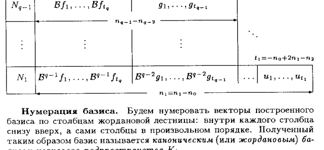
N_1) Здесь строятся векторы
$$\mathcal{B}^{q-1}f_1, \ldots, \mathcal{B}^{q-1}f_{t_q}, \mathcal{B}^{q-2}g_1, \ldots, \mathcal{B}^{q-2}g_{t_{q-1}}, \ldots, \mathcal{B}v_1, \ldots, \mathcal{B}v_{t_2},$$
которые дополняются векторами $u_1, \ldots, u_{t_1}$ до базиса N_1. Таким образом, векторы
$$\mathcal{B}^{q-1}f_1, \ldots, \mathcal{B}^{q-1}f_{t_q}, \mathcal{B}^{q-2}g_1, \ldots, \mathcal{B}^{q-2}g_{t_{q-1}}, \ldots, \mathcal{B}v_1, \ldots, \mathcal{B}v_{t_2}, u_1, \ldots, u_{t_1}$$

1) являются собственными векторами;
2) их количество равно $n_1 = n_1 - n_0$ (очевидно, $n_0 = \operatorname{def}\mathcal{B}^0 = 0$);
3) $t_1 = (n_1 - n_0) - (n_2 - n_1) = -n_2 + 2n_1 - n_0$;
4) они линейно независимы.

Полученную за q шагов систему векторов удобно объединить в приводимую ниже таблицу, которую будем называть *жордановой лестницей*.

Теорема 92.1. *Построенная система векторов образует базис корневого подпространства* K_{λ_j}.

Доказательство. Количество векторов в построенной системе равно размерности пространства K_{λ_j}, так как $n_1 + (n_2 - n_1) + (n_3 - n_2) + \ldots + (n_q - n_{q-1}) = n_q = \dim K_{\lambda_j}$. Эти векторы линейно независимы, так как если приравнять их линейную комбинацию нулевому вектору и последовательно применить операторы $\mathcal{B}^{q-1}, \mathcal{B}^{q-2}, \ldots, \mathcal{B}$ к обеим частям полученного равенства, то все коэффициенты линейной комбинации окажутся равными нулю. ■



Нумерация базиса. Будем нумеровать векторы построенного базиса по столбцам жордановой лестницы: внутри каждого столбца снизу вверх, а сами столбцы в произвольном порядке. Полученный таким образом базис называется *каноническим* (или *жордановым*) *базисом корневого подпространства* K_{λ_j}.

Матрица оператора $\mathcal{A}|K_{\lambda_j}$ в каноническом базисе. 1. Пусть $e_1, \ldots, e_q$ – векторы первого столбца жордановой лестницы. Тогда

$$\begin{cases} e_1 = \mathcal{B}^{q-1}f_1, \\ e_2 = \mathcal{B}^{q-2}f_1, \\ \ldots \\ e_q = f_1 \end{cases} \Rightarrow \begin{cases} \mathcal{B}e_1 = \theta, \\ \mathcal{B}e_2 = e_1, \\ \ldots \\ \mathcal{B}e_q = e_{q-1} \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} (\mathcal{A} - \lambda_j\mathcal{I})e_1 = \theta, \\ (\mathcal{A} - \lambda_j\mathcal{I})e_2 = e_1, \\ \ldots \\ (\mathcal{A} - \lambda_j\mathcal{I})e_q = e_{q-1} \end{cases} \Rightarrow \begin{cases} \mathcal{A}e_1 = \lambda_j e_1, \\ \mathcal{A}e_2 = \lambda_j e_2 + e_1, \\ \ldots \\ \mathcal{A}e_q = \lambda_j e_q + e_{q-1}. \end{cases} \qquad (92.1)$$

Этой группе векторов канонического базиса соответствуют первые q столбцов матрицы $\mathcal{A}|K_{\lambda_j}$ в каноническом базисе, которые согласно

вида
$$\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & & \boxed{\begin{matrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{matrix}} & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}. \qquad (97.9)$$

Простым отражением называется оператор в евклидовом пространстве, который в некотором ортонормированном базисе имеет матрицу вида
$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & O & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & 1 & \\ & O & & & & \ddots \\ & & & & & & 1 \end{bmatrix}. \qquad (97.10)$$

Из определения следует, что простое вращение и простое отражение – ортогональные операторы, так как их матрицы (97.9) и (97.10) в ортонормированном базисе ортогональны. Простое вращение представляет собой поворот в некоторой двумерной плоскости и оставляет неизменным $(n-2)$-мерное подпространство, ортогональное этой плоскости. Простое отражение меняет направление всех векторов некоторого одномерного подпространства и оставляет неизменным его $(n-1)$-мерное ортогональное дополнение.

Теорема 97.5. *Всякий ортогональный оператор может быть представлен как произведение некоторого числа простых вращений и простых отражений.*

Это утверждение следует из того, что матрица вида (97.8) может быть представлена в виде произведения некоторого числа матриц вида (97.9), (97.10). ■

§98. Самосопряженный оператор

Линейный оператор $\mathcal{A}$, действующий в унитарном (евклидовом) пространстве, называется *самосопряженным*, если $\mathcal{A} = \mathcal{A}^*$. Самосопряженный оператор в унитарном пространстве называют *эрмитовым*, а в евклидовом пространстве – *симметрическим*.

Квадратная матрица A (комплексная или вещественная) называется *самосопряженной*, если $A = A^H$. Комплексную самосопряженную

матрицу называют *эрмитовой*, а вещественную – *симметрической* или *вещественно-эрмитовой* (очевидно, для симметрической матрицы: $A = A^T$).

Из определения вытекают следующие факты.

1°. Самосопряженный оператор нормален.

2°. Оператор самосопряжен тогда и только тогда, когда в любом ортонормированном базисе он имеет самосопряженную матрицу.

3°. Определитель самосопряженного оператора веществен.

4°. Если подпространство L инвариантно относительно самосопряженного оператора $\mathcal{A}$, то $L^\perp$ также инвариантно относительно $\mathcal{A}$ (теорема 95.4).

5°. Самосопряженный оператор на любом инвариантном подпространстве индуцирует самосопряженный оператор.

Теорема 98.1 (спектральная характеристика самосопряженного оператора). *Нормальный оператор в унитарном (евклидовом) пространстве самосопряжен тогда и только тогда, когда все корни его характеристического многочлена вещественны.*

Доказательство. Необходимость. В унитарном пространстве это утверждение означает, что все собственные значения эрмитова оператора вещественны, и вытекает из равенств $\mathcal{A}x = \lambda x$ и, с учетом теоремы 96.1, $\mathcal{A}x = \bar{\lambda}x$. Докажем утверждение для евклидова пространства E. Пусть e – ортонормированный базис E, тогда A_e – самосопряженная (вещественная) матрица. Рассмотрим произвольное унитарное пространство U той же размерности, что и пространство E, и в нем произвольный ортонормированный базис f. Тогда матрице A_e отвечает самосопряженный оператор $\mathcal{B} \in \mathcal{L}(U, U)$, для которого матрица A_e является матрицей в базисе f: $A_e = B_f$. Следовательно, характеристические многочлены операторов $\mathcal{A}$ и $\mathcal{B}$ совпадают и по доказанному выше (применительно к оператору $\mathcal{B}$) все корни характеристического многочлена оператора $\mathcal{A}$ вещественны.

Достаточность. Пусть $\mathcal{A}$ – нормальный оператор и все корни его характеристического многочлена вещественны. Тогда как в евклидовом, так и в унитарном пространстве существует ортонормированный базис $e_1, \ldots, e_n$ из собственных векторов оператора $\mathcal{A}$. Если $x = \sum_{i=1}^n x_i e_i$ – любой вектор пространства, то $\mathcal{A}x = \sum_{i=1}^n x_i\lambda_i e_i$ и $\mathcal{A}^*x = \sum_{i=1}^n x_i\overline{\lambda_i}e_i = \sum_{i=1}^n x_i\lambda_i e_i$, так как $\lambda_i \in \mathbb{R}$. Следовательно, $\mathcal{A}x = \mathcal{A}^*x$, $\forall x \in V$, откуда следует, что $\mathcal{A} = \mathcal{A}^*$. ■

Итак, для самосопряженных операторов основное поле ($\mathbb{R}$ или $\mathbb{C}$) не играет такой "роковой" роли, как это было с другими операторами; все факты, касающиеся комплексных пространств, справедливы и в вещественном пространстве.

В частности, как следует из теоремы 98.1, *оператор, действующий в унитарном (или евклидовом) пространстве, самосопряжен тогда и только тогда, когда существует ортонормированный базис, в котором его матрица имеет вещественную диагональную*

Это утверждение следует из того, что квадратные матрицы одинакового порядка над общим полем подобны тогда и только тогда, когда они являются матрицами одного и того же линейного оператора (§82, п. 8°). ■

Теорема 92.4 (теорема Гамильтона–Кэли). *Линейный оператор, действующий в комплексном (или в вещественном) пространстве, является корнем своего характеристического многочлена.*

Доказательство. 1. Докажем сначала это утверждение для комплексного пространства V. Пусть $\mathcal{A} \in \mathcal{L}(V, V)$ и его характеристический многочлен имеет вид (91.1). Согласно теореме 91.1, $V = K_{\lambda_1} \oplus \ldots \oplus K_{\lambda_p}$ и, следовательно, для любого вектора $x \in V$ имеет место разложение $x = x_1 + \ldots + x_p$, где $x_j \in K_{\lambda_j}$, $j = \overline{1,p}$. Тогда
$$f(\mathcal{A})x = f(\mathcal{A})x_1 + \ldots + f(\mathcal{A})x_j + \ldots + f(\mathcal{A})x_p. \qquad (92.6)$$

Каждое слагаемое в этом разложении равно нулевому вектору, так как $f(\mathcal{A})x_j = (\lambda_1\mathcal{I} - \mathcal{A})^{m_1}\ldots(\lambda_j\mathcal{I} - \mathcal{A})^{m_j}\ldots(\lambda_p\mathcal{I} - \mathcal{A})^{m_p}x_j = \theta$, ибо операторы в этом произведении перестановочны, а $(\mathcal{A} - \lambda\mathcal{I})^{m_j}x_j = \theta$ в силу (91.10). Следовательно, $f(\mathcal{A})x = \theta$, $\forall x \in V$, т.е. $f(\mathcal{A}) = \mathcal{O}$.

2. Пусть V – вещественное линейное пространство. Возьмем какой-либо базис e пространства V, и пусть A_e – матрица оператора $\mathcal{A}$ в этом базисе. Для доказательства теоремы достаточно показать, что $f(A_e) = O$. Рассмотрим любое комплексное пространство V_1 той же размерности, что и V. Пусть f – произвольный базис V_1, тогда матрица A_e является матрицей оператора $\mathcal{B} \in \mathcal{L}(V_1, V_1)$ в базисе f, т.е. $A_e = B_f$. Значит, характеристические многочлены операторов $\mathcal{A}$ и $\mathcal{B}$ совпадают и, согласно п. 1 доказательства, $f(A_e) = O$. ■

§93. Инвариантные подпространства минимальной размерности

Теорема 93.1. *У всякого линейного оператора в комплексном пространстве существует одномерное инвариантное подпространство.*

Доказательство. Утверждение следует из существования собственного вектора для любого оператора, действующего в комплексном пространстве (§86): если e – собственный вектор оператора $\mathcal{A}$, то $\mathcal{L}(e)$ – одномерное подпространство, инвариантное относительно $\mathcal{A}$. ■

Теорема 93.2. *У всякого линейного оператора в вещественном пространстве существует одномерное или двумерное инвариантное подпространство.*

Доказательство. Пусть V – вещественное пространство, $\mathcal{A} \in \mathcal{L}(V, V)$, $e = (e_1, \ldots, e_n)$ – базис V и A – матрица оператора $\mathcal{A}$ в базисе e. Характеристический многочлен $f(\lambda)$ оператора $\mathcal{A}$ является многочленом с вещественными коэффициентами, так как $f(\lambda) = \det(A - \lambda I)$,

где $A \in \mathbb{R}^{n\times n}$. Пусть α_0 – корень характеристического многочлена $f(\lambda)$.

1. Если $\lambda_0 \in \mathbb{R}$, то λ_0 – собственное значение оператора $\mathcal{A}$. Тогда линейная оболочка $\mathcal{L}(e)$, натянутая на соответствующий собственный вектор e, образует одномерное подпространство, инвариантное относительно оператора $\mathcal{A}$.

2. Если $\lambda_0 = \alpha + i\beta$, $\beta \neq 0$, то $|A - \lambda_0 I| = 0$ и система уравнений
$$Az = \lambda_0 z \qquad (93.1)$$
над полем $\mathbb{C}$ имеет нетривиальное решение $z_0 = (x_1 + iy_1, \ldots, x_n + iy_n)^T$. В обозначениях $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$, $y = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$ система (93.1) может быть записана в виде $A(x + iy) = (\alpha + i\beta)(x + iy)$ или
$$\begin{cases} Ax = \alpha x - \beta y, \\ Ay = \beta x + \alpha y. \end{cases} \qquad (93.2)$$

Тогда, если $u = \sum_{i=1}^n x_i e_i$, $v = \sum_{i=1}^n y_i e_i$, то системе (93.2) соответствует система векторных уравнений
$$\begin{cases} \mathcal{A}u = \alpha u - \beta v, \\ \mathcal{A}v = \beta u + \alpha v, \end{cases} \qquad (93.3)$$
где u, v – векторы пространства V, одновременно не равные нулю. Из (93.3) следует, что $\mathcal{L}(u, v)$ – инвариантное подпространство. Нетрудно проверить, что $\dim\mathcal{L}(u, v) = 2$.

Итак, каждый вещественный корень характеристического многочлена порождает одномерное инвариантное пространство, а каждый комплексный (не вещественный) корень – двумерное инвариантное подпространство. ■

тем самым в силу (99.3) будет доказано, что $\mathcal{C} = \mathcal{B}$.

Разложим вектор e_i по базису f:
$$e_i = \sum_{k=1}^n \alpha_k f_k. \qquad (99.5)$$

Отметим, что в этом равенстве участвуют собственные векторы e_i, $f_1, \ldots, f_n$ оператора $\mathcal{A}$, отвечающие собственным значениям λ_i, $\mu_1^2, \ldots, \mu_n^2$. Из линейной независимости собственных векторов, отвечающих различным собственным значениям, следует, что в разложении (99.5) отличными от нуля будут коэффициенты α_k лишь при тех f_k, которые отвечают собственному значению $\mu_k^2 = \lambda_i$. Поэтому $\mathcal{C}e_i = \sum_{k=1}^n \alpha_k\mathcal{C}f_k = \sum_{k=1}^n \alpha_k\sqrt{\lambda_i}f_k = \sqrt{\lambda_i}\sum_{k=1}^n \alpha_k f_k = \sqrt{\lambda_i}e_i$, т.е. (99.4). ■

Оператор $\mathcal{B}$ называется *квадратным корнем из оператора* $\mathcal{A}$.

§100. Разложения линейного оператора

Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ в унитарном (евклидовом) пространстве V называется *косоэрмитовым* (соответственно *кососимметрическим*), если $\mathcal{A}^* = -\mathcal{A}$.

Квадратная комплексная матрица называется *косоэрмитовой*, если $A^H = -A$. Квадратная вещественная матрица называется *кососимметрической*, если $A^T = -A$.

Из определения следует, что оператор $\mathcal{A}$ косоэрмитов (кососимметричен) тогда и только тогда, когда его матрица в любом ортонормированном базисе пространства косоэрмитова (соответственно кососимметрична).

Теорема 100.1. *Линейный оператор $\mathcal{A}$ в унитарном пространстве эрмитов тогда и только тогда, когда оператор $i\mathcal{A}$ косоэрмитов.*

Доказательство. Заметим, что $(i\mathcal{A})^* = -i\mathcal{A}^*$ (теорема 94.4). Следовательно, $(i\mathcal{A})^* = -i\mathcal{A}$ тогда и только тогда, когда $\mathcal{A} = \mathcal{A}^*$. ■

Из теоремы следует, что все свойства самосопряженных операторов переносятся на косоэрмитовы (кососимметрические) операторы с той лишь разницей, что в этом случае все корни характеристического многочлена чисто мнимые.

Теорема 100.2. *Линейный оператор $\mathcal{A}$ в унитарном (евклидовом) пространстве может быть представлен, и притом единственным образом, в виде суммы*
$$\mathcal{A} = \mathcal{B} + \mathcal{C} \qquad (100.1)$$
эрмитова (симметрического) оператора $\mathcal{B}$ и косоэрмитова (кососимметрического) оператора $\mathcal{C}$.

Доказательство. Положим
$$\mathcal{B} = \frac{1}{2}(\mathcal{A} + \mathcal{A}^*), \quad \mathcal{C} = \frac{1}{2}(\mathcal{A} - \mathcal{A}^*). \qquad (100.2)$$

Тогда, как легко проверить, $\mathcal{B}^* = \mathcal{B}$, $\mathcal{C}^* = -\mathcal{C}$ и $\mathcal{A} = \mathcal{B} + \mathcal{C}$. Единственность такой пары операторов следует из того, что для любой другой пары операторов $\mathcal{B}_1$ и $\mathcal{C}_1$ таких, что $\mathcal{B}_1^* = \mathcal{B}_1$, $\mathcal{C}_1^* = -\mathcal{C}_1$ и $\mathcal{A} = \mathcal{B}_1 + \mathcal{C}_1$, имеем $\mathcal{A}^* = \mathcal{B}_1 - \mathcal{C}_1$ или $\frac{1}{2}(\mathcal{A} + \mathcal{A}^*) = \mathcal{B}_1$, $\frac{1}{2}(\mathcal{A} - \mathcal{A}^*) = \mathcal{C}_1$, т.е. в силу (100.2) $\mathcal{B}_1 = \mathcal{B}$, $\mathcal{C}_1 = \mathcal{C}$. ■

Разложение (100.1) называется *эрмитовым разложением* оператора $\mathcal{A}$.

Теорема 100.3. *Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ в унитарном (евклидовом) пространстве нормален тогда и только тогда, когда операторы $\mathcal{B}$ и $\mathcal{C}$ в эрмитовом разложении (100.1) этого оператора перестановочны.*

Доказательство. Если $\mathcal{A} = \mathcal{B} + \mathcal{C}$, то $\mathcal{A}^* = \mathcal{B} - \mathcal{C}$ и $\mathcal{A}\mathcal{A}^* = \mathcal{B}^2 - \mathcal{B}\mathcal{C} + \mathcal{C}\mathcal{B} - \mathcal{C}^2$, $\mathcal{A}^*\mathcal{A} = \mathcal{B}^2 - \mathcal{C}\mathcal{B} + \mathcal{B}\mathcal{C} - \mathcal{C}^2$, т.е. $\mathcal{A}\mathcal{A}^* - \mathcal{A}^*\mathcal{A} = \frac{1}{2}(\mathcal{C}\mathcal{B} - \mathcal{B}\mathcal{C})$. Значит, $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$ тогда и только тогда, когда $\mathcal{C}\mathcal{B} = \mathcal{B}\mathcal{C}$. ■

Теорема 100.4. *Любой оператор $\mathcal{A}$ в унитарном (евклидовом) пространстве может быть представлен в виде произведения*
$$\mathcal{A} = \mathcal{B}\mathcal{U} \qquad (100.3)$$
неотрицательного оператора $\mathcal{B}$ и унитарного оператора $\mathcal{U}$. При этом оператор $\mathcal{B}$ определен однозначно, а если $\mathcal{A}$ обратим, то и оператор $\mathcal{U}$ определен однозначно.

Доказательство. 1. Рассмотрим оператор $\mathcal{A}^*\mathcal{A}$.

а) $\mathcal{A}^*\mathcal{A}$ – самосопряженный оператор, так как $(\mathcal{A}^*\mathcal{A})^* = \mathcal{A}^*\mathcal{A}$.

б) $\mathcal{A}^*\mathcal{A} \geq \mathcal{O}$, так как $(\mathcal{A}^*\mathcal{A}x, x) = (\mathcal{A}x, \mathcal{A}x) \geq 0$, $\forall x \neq \theta$.

в) Существует ортонормированный базис пространства из собственных векторов $e_1, \ldots, e_n$ оператора $\mathcal{A}^*\mathcal{A}$. При этом можно считать, что векторы $e_1, \ldots, e_r$ отвечают ненулевым собственным значениям $\rho_1^2, \ldots, \rho_r^2$, а остальные – нулевым, так что
$$\rho_k > 0 \text{ при } k \leq r \text{ и } \rho_k = 0 \text{ при } k > r. \qquad (100.4)$$

Отметим, что
$$r = \operatorname{rg}\mathcal{A}^*\mathcal{A}. \qquad (100.5)$$

2. Согласно (80.1), $\operatorname{im}\mathcal{A} = \mathcal{L}(\mathcal{A}e_1, \ldots, \mathcal{A}e_n)$. Для векторов $\mathcal{A}e_1, \ldots, \mathcal{A}e_n$ имеем
$$(\mathcal{A}e_i, \mathcal{A}e_j) = (\mathcal{A}^*\mathcal{A}e_i, e_j) = (\rho_i^2 e_i, e_j) = \begin{cases} \rho_i^2, & i = j, \\ 0, & i \neq j. \end{cases}$$

Следовательно, $\mathcal{A}e_{r+1} = \ldots = \mathcal{A}e_n = \theta$ (в силу (100.5)), а $\mathcal{A}e_1, \ldots, \mathcal{A}e_r$ – ненулевые попарно ортогональные векторы, причем $|\mathcal{A}e_k| = \rho_k$, $k = \overline{1,r}$. Следовательно, векторы
$$g_k = \frac{1}{\rho_k}\mathcal{A}e_k, \quad k = \overline{1,r}, \qquad (100.6)$$

Итак,
$$(e_i, f_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \qquad (95.1)$$

Примером биортогональной пары базисов могут служить тройки некомпланарных векторов **a**, **b**, **c** и [**b**, **c**], [**c**, **a**], [**a**, **b**] геометрического пространства V_3, для которых (**a**, **b**, **c**) = 1.

Очевидно, что ортонормированный базис биортогонален самому себе.

Теорема 95.1. *Для любого базиса $e_1, \ldots, e_n$ унитарного (евклидова) пространства существует, и притом единственный, биортогональный базис $f_1, \ldots, f_n$.*

Доказательство. Согласно (95.1) вектор f_j, $j = \overline{1,n}$, ортогонален всем векторам $e_1, \ldots, e_n$, кроме e_j. Следовательно, $f_j \in L_j$, где $L_j = \mathcal{L}^\perp(e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n)$. Очевидно, $\dim L_j = 1$. Если g – базис L_j, то $f_j = \alpha g$. Из (95.1) следует, что $(f_j, e_j) = 1$, откуда получаем, что $\alpha = 1/(g, e_j)$. Тем самым доказаны существование и единственность векторов f_j, $j = \overline{1,n}$, биортогонального базиса. ■

Теорема 95.2. *В паре биортогональных базисов e и f унитарного (евклидова) пространства V матрицы операторов $\mathcal{A}$ и $\mathcal{A}^*$ связаны соотношением*
$$(\mathcal{A}^*)_f = (A_e)^H. \qquad (95.2)$$

Доказательство. Пусть $A_e = (a_{ij})$, $(\mathcal{A}^*)_f = (b_{ij})$. Тогда $\mathcal{A}e_j = \sum_{k=1}^n a_{kj}e_k$, $\mathcal{A}^*f_i = \sum_{k=1}^n b_{ki}f_k$. Умножив первое из этих равенств скалярно на f_i, получим, что $(\mathcal{A}e_j, f_i) = \sum_{k=1}^n a_{kj}(e_k, f_i) = a_{ij}$. С другой стороны, $(\mathcal{A}e_j, f_i) = (e_j, \mathcal{A}^*f_i) = (e_j, \sum_{k=1}^n b_{ki}f_k) = \sum_{k=1}^n \overline{b_{ki}}(e_j, f_k) = \overline{b_{ji}}$. Следовательно, $a_{ij} = \overline{b_{ji}}$, $i = \overline{1,n}$, $j = \overline{1,n}$. Это равносильно (95.2). ■

Следствие 1. *В ортонормированном базисе e*
$$(\mathcal{A}^*)_e = (A_e)^H.$$

Следствие 2. *Для любого оператора $\mathcal{A} \in \mathcal{L}(V, V)$*
$$\det\mathcal{A}^* = \overline{\det\mathcal{A}}, \quad \operatorname{rg}\mathcal{A}^* = \operatorname{rg}\mathcal{A}. \qquad (95.3)$$

Теорема 95.3. *Для любого оператора $\mathcal{A} \in \mathcal{L}(V, V)$*
$$\ker\mathcal{A} = \operatorname{im}^\perp\mathcal{A}^*, \quad \ker\mathcal{A}^* = \operatorname{im}^\perp\mathcal{A}. \qquad (95.4)$$

Доказательство. Для любых векторов $x \in \ker\mathcal{A}$, $y \in \operatorname{im}\mathcal{A}^*$ имеем $\mathcal{A}x = \theta$, $y = \mathcal{A}^*y_1$. Значит, $(x, y) = (x, \mathcal{A}^*y_1) = (\mathcal{A}x, y_1) = 0$. Следовательно, $\ker\mathcal{A} \subset \operatorname{im}^\perp\mathcal{A}^*$. С другой стороны (см. (80.2) и (95.3)), $\dim\ker\mathcal{A} = \dim V - \dim\operatorname{im}\mathcal{A} = \dim V - \dim\operatorname{im}\mathcal{A}^* = \dim\operatorname{im}^\perp\mathcal{A}^*$. Теорема 64.3 о монотонности размерности доказывает первое из равенств (95.4). Второе доказывается аналогично. ■

Теорема 95.4. *Если подпространство L инвариантно относительно оператора $\mathcal{A}$, то его ортогональное дополнение $L^\perp$ инвариантно относительно сопряженного оператора $\mathcal{A}^*$.*

Доказательство. Если $x \in L$, $y \in L^\perp$, то $(\mathcal{A}x, y) = 0$, так как по условию теоремы $\mathcal{A}x \in L$. С другой стороны, $(\mathcal{A}x, y) = (x, \mathcal{A}^*y)$. Следовательно, $(x, \mathcal{A}^*y) = 0$, $\forall x \in L$ и, значит, $\mathcal{A}^*y \in L^\perp$. ■

§96. Нормальный оператор

Пусть V – унитарное или евклидово пространство. Линейный оператор $\mathcal{A} \in \mathcal{L}(V, V)$ называется *нормальным оператором*, если
$$\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}.$$

Квадратная матрица A (комплексная или вещественная) называется *нормальной матрицей*, если
$$AA^H = A^HA.$$

Замечание 1. Из определения и теоремы 95.2 следует, что оператор нормален тогда и только тогда, когда в любом ортонормированном базисе его матрица нормальна.

Теорема 96.1. *Собственный вектор нормального оператора, отвечающий собственному значению λ, является собственным вектором сопряженного оператора, отвечающим собственному значению $\bar{\lambda}$.*

Доказательство. Легко проверить, что если $\mathcal{A}$ – нормальный оператор, то $\mathcal{A} - \lambda\mathcal{I}$ также нормален. Пусть теперь x – собственный вектор оператора $\mathcal{A}$, отвечающий собственному значению λ, тогда $(\mathcal{A} - \lambda\mathcal{I})x = \theta$ и $((\mathcal{A} - \lambda\mathcal{I})x, (\mathcal{A} - \lambda\mathcal{I})x) = 0$. Согласно (94.1) отсюда следует, что $(x, (\mathcal{A} - \lambda\mathcal{I})^*(\mathcal{A} - \lambda\mathcal{I})x) = 0$ или, с учетом нормальности оператора $\mathcal{A} - \lambda\mathcal{I}$, что $(x, (\mathcal{A} - \lambda\mathcal{I})(\mathcal{A} - \lambda\mathcal{I})^*x) = 0$, т.е. $((\mathcal{A} - \lambda\mathcal{I})^*x, (\mathcal{A} - \lambda\mathcal{I})^*x) = 0$ и $(\mathcal{A} - \lambda\mathcal{I})^*x = \theta$. Отсюда в силу теоремы 94.4 следует, что $(\mathcal{A}^* - \bar{\lambda}\mathcal{I})x = \theta$, т.е. $\mathcal{A}^*x = \bar{\lambda}x$. ■

Следствие 1. *Если $\mathcal{A}$ – нормальный оператор, то*
$$\ker\mathcal{A} = \ker\mathcal{A}^*, \qquad (96.1)$$
так как нетривиальные векторы ядра являются собственными векторами, отвечающими нулевому собственному значению.

Следствие 2. *Если $\mathcal{A}$ – нормальный оператор, то*
$$\ker\mathcal{A} = \operatorname{im}^\perp\mathcal{A}, \quad \ker\mathcal{A}^* = \operatorname{im}^\perp\mathcal{A}^*.$$

Это следует из (95.4) и (96.1).

Глава XVII. Квадратичные формы

В §58 рассматривалась одна из основных задач аналитической геометрии на плоскости – приведение общего уравнения линии второго порядка к каноническому виду. Теория квадратичных форм, излагаемая здесь, основной своей целью ставит решение этой задачи и задач, связанных с нею, в пространстве любой размерности.

§101. Квадратичные формы в линейном пространстве

Билинейные формы. Пусть V – линейное пространство над полем P[1]. Отображение $\mathcal{A} : V \times V \to P$ называется *билинейной формой* в пространстве V, если для любых $x, y, z \in V$, $\alpha \in P$:
$$\begin{aligned} &1)\ \mathcal{A}(x + y, z) = \mathcal{A}(x, z) + \mathcal{A}(y, z); \\ &2)\ \mathcal{A}(\alpha x, y) = \alpha\mathcal{A}(x, y); \\ &3)\ \mathcal{A}(x, y + z) = \mathcal{A}(x, y) + \mathcal{A}(x, z); \\ &4)\ \mathcal{A}(x, \alpha y) = \alpha\mathcal{A}(x, y). \end{aligned} \qquad (101.1)$$

Билинейная форма называется *симметричной*, если $\mathcal{A}(y, x) = \mathcal{A}(x, y)$, $\forall x, y \in V$.

Примеры. 1. Скалярное произведение (x, y) в евклидовом пространстве является симметричной билинейной формой.

2. В любом линейном пространстве V, если $f(x)$, $g(x)$ – линейные формы, то функция $\mathcal{A}(x, y) = f(x)g(y)$ является симметричной билинейной формой.

3. В n-мерном пространстве V с базисом $e_1, \ldots, e_n$ отображение $\mathcal{A} : V \times V \to P$, определенное правилом
$$\mathcal{A}(x, y) = \sum_{i,j=1}^n a_{ij}x_iy_j, \qquad (101.2)$$
$\forall x = \sum_{i=1}^n x_ie_i$, $y = \sum_{i=1}^n y_ie_i$, где a_{ij} $(i, j = \overline{1,n})$ – фиксированные числа, является билинейной формой (в силу линейности координат).

Теорема 101.1. *Пусть V – линейное пространство над полем P и $e_1, \ldots, e_n$ – базис V. Для любых чисел $a_{ij} \in P$, $i, j = \overline{1,n}$, существует, и притом единственная, билинейная форма $\mathcal{A}(x, y)$ в пространстве V, для которой $\mathcal{A}(e_i, e_j) = a_{ij}$, $i, j = \overline{1,n}$.*

Доказательство. Пусть $e_1, \ldots, e_n$ – базис пространства V и a_{ij}, $i, j = \overline{1,n}$, – заданные числа. Отображение (101.2) является билинейной формой в пространстве V (пример 3), причем $\mathcal{A}(e_i, e_j) = a_{ij}$,

[1] В этой главе предполагается, что характеристика основного поля равна 0.

$i, j = \overline{1,n}$ (так как векторы e_i и e_j имеют очевидные наборы координат в базисе $e_1, \ldots, e_n$). Покажем, что любая билинейная форма $\mathcal{B}(x, y)$ в пространстве V, для которой $\mathcal{B}(e_i, e_j) = a_{ij}$, $i, j = \overline{1,n}$, совпадает с $\mathcal{A}(x, y)$. В самом деле, для любых векторов $x = \sum_{i=1}^n x_ie_i$, $y = \sum_{i=1}^n y_ie_i$ в силу (101.1) имеем $\mathcal{B}(x, y) = \mathcal{B}(\sum_{i=1}^n x_ie_i, \sum_{j=1}^n y_je_j) = \sum_{i=1}^n\sum_{j=1}^n x_iy_j\mathcal{B}(e_i, e_j) = \sum_{i,j=1}^n a_{ij}x_iy_j$, т.е. $\mathcal{B}(x, y) = \mathcal{A}(x, y)$, $\forall x, y \in V$. ■

Таким образом, в примере 3 приведен самый общий вид билинейной формы в n-мерном пространстве.

Представление билинейной формы в виде (101.2) называется *общим видом билинейной формы в базисе e*. Матрица $A_e = (a_{ij}) \in P^{n\times n}$, элементы которой определены равенством $a_{ij} = \mathcal{A}(e_i, e_j)$, $i, j = \overline{1,n}$, называется *матрицей билинейной формы $\mathcal{A}(x, y)$ в базисе e*.

Как следует из теоремы 101.1, билинейная форма однозначно определяется своим заданием на базисных векторах, при этом матрица билинейной формы представляет собой таблицу ее значений на парах базисных векторов и в силу (101.2) однозначно определяет значения этой формы для любой пары векторов.

Общий вид (101.2) билинейной формы $\mathcal{A}(x, y)$ может быть записан в компактной форме: если x_e и y_e – координатные столбцы векторов x и y в базисе e, то
$$\mathcal{A}(x, y) = x_e^TA_ey_e, \qquad \mathcal{A}(x, y) = y_e^TA_e^Tx_e. \qquad (101.3)$$

Первое из равенств (101.3) проверяется непосредственно, второе равенство можно получить транспонированием обеих частей первого.

Выражение, стоящее в правой части (101.2), называется *билинейной формой от переменных $x_1, \ldots, x_n$ и $y_1, \ldots, y_n$*. Правые части равенств (101.3) представляют собой компактные формы записи этой билинейной формы.

Теорема 101.2. *Произвольная матрица $A = (a_{ij}) \in P^{n\times n}$ является матрицей единственной билинейной формы в заданном базисе пространства.*

Доказательство. Пусть $e_1, \ldots, e_n$ – заданный базис пространства V. Матрица $A = (a_{ij}) \in P^{n\times n}$ является матрицей билинейной формы (101.2). Единственность этой билинейной формы вытекает из теоремы 101.1. ■

Таким образом, соответствие между билинейными формами и их матрицами является биективным отображением множества всех билинейных форм в n-мерном пространстве V над полем P на множестве всех матриц $P^{n\times n}$.

Теорема 101.3. *Матрицы билинейной формы $\mathcal{A}(x, y)$ в базисах e и $f = eQ$ связаны соотношением $A_f = Q^TA_eQ$.*

Доказательство. Согласно (101.3), с одной стороны, $\mathcal{A}(x, y) = x_e^TA_ey_e = \{x_e = Qx_f, y_e = Qy_f\} = x_f^TQ^TA_eQy_f$. С другой стороны, $\mathcal{A}(x, y) = x_f^TA_fy_f$. Отсюда с учетом произвольности x, y следует утверждение теоремы. ■

Следствие. $\operatorname{rg}A_e = \operatorname{rg}A_f$.

§97. Унитарный (ортогональный) оператор

Линейный оператор $\mathcal{U}$, действующий в унитарном (евклидовом) пространстве, называется *унитарным* (соответственно *ортогональным*) оператором, если $\mathcal{U}^*\mathcal{U} = \mathcal{U}\mathcal{U}^* = \mathcal{I}$.

Из определения вытекают следующие факты.

1°. Оператор $\mathcal{U}$ унитарен (ортогонален) только тогда, когда в любом ортонормированном базисе он имеет унитарную (соответственно ортогональную) матрицу.

2°. Для унитарного (ортогонального) оператора $\mathcal{U}$
$$|\det\mathcal{U}| = 1. \qquad (97.1)$$

3°. Унитарный (ортогональный) оператор нормален.

Теорема 97.1 (критерии унитарности). *В унитарном (евклидовом) пространстве V следующие утверждения равносильны:*

1) оператор $\mathcal{U}$ унитарен (ортогонален);
2) $\mathcal{U}^*\mathcal{U} = \mathcal{I}$;
3) $\mathcal{U}\mathcal{U}^* = \mathcal{I}$;
4) оператор $\mathcal{U}$ сохраняет скалярное произведение, т.е.
$$(\mathcal{U}x, \mathcal{U}y) = (x, y), \quad \forall x, y \in V;$$
5) оператор $\mathcal{U}$ сохраняет длину, т.е.
$$|\mathcal{U}x| = |x|, \quad \forall x \in V;$$
6) оператор $\mathcal{U}$ переводит любой ортонормированный базис V в ортонормированный базис;
7) оператор $\mathcal{U}$ переводит хотя бы один ортонормированный базис V в ортонормированный базис.

Доказательство. $1 \Leftrightarrow 2 \Leftrightarrow 3$. Любое из равенств $\mathcal{U}^*\mathcal{U} = \mathcal{I}$ или $\mathcal{U}\mathcal{U}^* = \mathcal{I}$ означает невырожденность $\mathcal{U}$ и существование $\mathcal{U}^{-1}$. Умножение обеих частей этих равенств на $\mathcal{U}^{-1}$ (справа или слева) приводит к тому, что $\mathcal{U}^* = \mathcal{U}^{-1}$. Следовательно, имеют место импликации $2 \Rightarrow 1$, $3 \Rightarrow 1$. Переход $1 \Rightarrow 2$, $1 \Rightarrow 3$ очевиден.

$1 \Rightarrow 4$. Так как $\mathcal{U}^*\mathcal{U} = \mathcal{I}$, то
$$(\mathcal{U}x, \mathcal{U}y) = (x, \mathcal{U}^*\mathcal{U}y) = (x, y), \quad \forall x, y \in V.$$

$4 \Rightarrow 1$. Так как $(x, \mathcal{U}^*\mathcal{U}y) = (\mathcal{U}x, \mathcal{U}y) = (x, y)$, $\forall x, y \in V$, то на основании теоремы 94.1 $\mathcal{U}^*\mathcal{U} = \mathcal{I}$.

$4 \Rightarrow 5$. Очевидно, так как $|\mathcal{U}x| = \sqrt{(\mathcal{U}x, \mathcal{U}x)} = \sqrt{(x, x)} = |x|$, $\forall x \in V$.

$5 \Rightarrow 4$. Это следует из легко проверяемых соотношений:
$$(x, y) = (|x + y|^2 - |x|^2 - |y|^2)/2$$
в евклидовом пространстве и
$$(x, y) = (|x + y|^2 - |x - y|^2 + i|x + iy|^2 - i|x - iy|^2)/4$$
в унитарном пространстве.

$4 \Rightarrow 6$. Очевидно, так как $(\mathcal{U}e_i, \mathcal{U}e_j) = (e_i, e_j) = \delta_{ij}$.

$6 \Rightarrow 4$. Если $e_1, \ldots, e_n$ – ортонормированный базис V и $x = \sum_{i=1}^n x_ie_i$, $y = \sum_{i=1}^n y_ie_i$, то $\mathcal{U}x = \sum_{i=1}^n x_i\mathcal{U}e_i$, $\mathcal{U}y = \sum_{i=1}^n y_i\mathcal{U}e_i$, и

так как $\mathcal{U}e_1, \ldots, \mathcal{U}e_n$ – ортонормированный базис, то, согласно (70.7), $(\mathcal{U}x, \mathcal{U}y) = \sum_{i=1}^n x_i\overline{y_i} = (x, y)$, $\forall x, y \in V$.

$6 \Rightarrow 7$. Очевидно.

$7 \Rightarrow 6$. Этот переход фактически доказан в п. $6 \Rightarrow 4$. ■

Следствие. Унитарный (ортогональный) оператор на любом инвариантном подпространстве индуцирует унитарный (соответственно ортогональный) оператор, так как сохраняет скалярное произведение любой пары векторов этого подпространства.

Теорема 97.2 (спектральная характеристика унитарного оператора). *Нормальный оператор в унитарном пространстве унитарен тогда и только тогда, когда все его собственные значения по модулю равны единице.*

Доказательство. Необходимость справедлива в унитарном и евклидовом пространстве и означает, что все собственные значения унитарного (и ортогонального) оператора $\mathcal{U}$ по модулю равны единице. Докажем это. Пусть x – собственный вектор оператора $\mathcal{U}$ и λ – отвечающее ему собственное значение. Тогда $(\mathcal{U}x, \mathcal{U}x) = (\lambda x, \lambda x) = |\lambda|^2(x, x)$. С другой стороны, $(\mathcal{U}x, \mathcal{U}x) = (x, \mathcal{U}^*\mathcal{U}x) = (x, x)$. Из этих равенств следует, что $|\lambda| = 1$.

Достаточность. Если $\mathcal{U}$ – нормальный оператор, то согласно теореме 96.4 в пространстве V существует ортонормированный базис $e_1, \ldots, e_n$ из собственных векторов оператора $\mathcal{U}$. Тогда для любого вектора $x = \sum_{i=1}^n x_ie_i \in V$ имеем $\mathcal{U}x = \sum_{i=1}^n x_i\lambda_ie_i$, где $|\lambda_i| = 1$, $i = \overline{1,n}$. В силу ортонормированности базиса e согласно (70.7) получаем отсюда, что $(x, x) = \sum_{i=1}^n |x_i|^2$ и $(\mathcal{U}x, \mathcal{U}x) = \sum_{i=1}^n |\lambda_i|^2|x_i|^2 = \sum_{i=1}^n |x_i|^2$. Следовательно, $|\mathcal{U}x| = |x|$, $\forall x \in V$, и $\mathcal{U}$ – унитарный оператор (теорема 97.1). ■

Теорема 97.3. *Если подпространство L инвариантно относительно унитарного (ортогонального) оператора $\mathcal{U}$, то его ортогональное дополнение $L^\perp$ также инвариантно относительно $\mathcal{U}$.*

Доказательство. Пусть $y \in L^\perp$. Покажем, что $\mathcal{U}y \in L^\perp$, т.е. $(x, \mathcal{U}y) = 0$, $\forall x \in L$. Оператор $\mathcal{U}$ индуцирует на подпространстве L унитарный (ортогональный) оператор $\mathcal{U}|L$. Значит, оператор $\mathcal{U}|L$ обратим и его образ совпадает со всем подпространством L, т.е. $\operatorname{im}\mathcal{U}|L = L$ (теорема 83.5). Это означает, что для любого вектора $x \in L$ существует вектор $x_1 \in L$ такой, что $x = \mathcal{U}x_1$. Тогда $(x, \mathcal{U}y) = (\mathcal{U}x_1, \mathcal{U}y) = (x_1, y) = 0$, так как $x_1 \in L$, $y \in L^\perp$. ■

Каноническая форма матрицы унитарного оператора. Из того что унитарный оператор нормален и все его собственные значения по модулю равны единице, следует, что *в пространстве V существует ортонормированный базис e, в котором матрица унитарного оператора $\mathcal{U}$ имеет диагональную форму*
$$U_e = \begin{bmatrix} \lambda_1 & & \text{O} \\ & \ddots & \\ \text{O} & & \lambda_n \end{bmatrix}, \quad \text{где } |\lambda_i| = 1, \ i = \overline{1,n},$$

Доказательство. Пусть $e = (e_1, \ldots, e_n)$ – базис V. Квадратичная форма $\mathcal{A}(x, x)$ с матрицей $A_e = (a_{ij})$ имеет в этом базисе вид (101.7). Обозначим $g(x_1, \ldots, x_n) = \sum_{i,j=1}^n a_{ij}x_ix_j$.

Переход к другому базису $f = eQ$ равносилен преобразованию координат по известному закону $x_e = Qx_f$, $|Q| \neq 0$. В терминах координат утверждение теоремы означает, что *любая квадратичная форма $g(x_1, \ldots, x_n)$ от переменных $x_1, \ldots, x_n$ невырожденным преобразованием координат приводится к сумме квадратов.* Это мы и будем доказывать. Будем считать, что $A_e \neq O$ (если $A_e = O$, то e – искомый базис). Обозначим через Δ_k *угловые миноры* k-го порядка матрицы A_e, т.е. $\Delta_k = M^{1,\ldots,k}_{1,\ldots,k}$, $k = \overline{1,n}$, и положим $\Delta_0 = 1$.

I. Рассмотрим сначала случай, когда $\Delta_k \neq 0$, $k = \overline{1, n-1}$. В этом случае процесс приведения к каноническому виду состоит из $n-1$ однотипных шагов.

Первый шаг основан на том, что $\Delta_1 \neq 0$, т.е. $a_{11} \neq 0$. Соберем все члены квадратичной формы $g(x_1, \ldots, x_n)$, содержащие x_1, и выделим из них полный квадрат:
$$\mathcal{A}(x, x) = g(x_1, \ldots, x_n) = a_{11}x_1^2 + 2\sum_{k=2}^n a_{1k}x_1x_k + \sum_{i,k=2}^n a_{ik}x_ix_k =$$
$$= a_{11}\left(x_1 + \sum_{k=2}^n \frac{a_{1k}}{a_{11}}x_k\right)^2 - a_{11}\left(\sum_{k=2}^n \frac{a_{1k}}{a_{11}}x_k\right)^2 + \sum_{i,k=2}^n a_{ik}x_ix_k.$$

Перейдем к новым координатам:
$$x_1' = x_1 + \sum_{k=2}^n \frac{a_{1k}}{a_{11}}x_k \text{ и } x_j' = x_j \text{ при } j \neq 1,$$
очевидно выполнив при этом невырожденное преобразование координат с матрицей
$$Q_1 = \begin{bmatrix} 1 & -a_{12}/a_{11} & \ldots & -a_{1n}/a_{11} \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix}.$$

Тогда квадратичная форма $\mathcal{A}(x, x)$ в новых координатах примет вид $\mathcal{A}(x, x) = a_{11}x_1'^2 + h(x_2', \ldots, x_n')$, где $h(x_2', \ldots, x_n')$ – квадратичная форма от переменных $x_2', \ldots, x_n'$. При этом матрица $A_1 = Q_1^TA_eQ_1$ квадратичной формы $\mathcal{A}(x, x)$ в новом базисе будет иметь вид
$$A_1 = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22}' & \ldots & a_{2n}' \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{n2}' & \ldots & a_{nn}' \end{bmatrix}.$$

Каждая строка (столбец) матрицы A_1, начиная со второй, получена из соответствующей строки (соответственно столбца) матрицы A_e вычитанием из нее первой строки (соответственно столбца) A_e, умноженной на некоторое число, поэтому угловые миноры матрицы

A_1 совпадают с $\Delta_1, \ldots, \Delta_n$. Следовательно, $\Delta_1 = a_{11}$, $\Delta_2 = a_{11}a_{22}'$ и
$$a_{22}' = \Delta_2/\Delta_1 \neq 0. \qquad (101.10)$$

Второй шаг основан на том, что $\Delta_2 \neq 0$, т.е. $a_{22}' \neq 0$, и состоит в применении действий первого шага к квадратичной форме $h(x_2', \ldots, x_n')$: выделяется полный квадрат среди всех членов, содержащих x_2', выполняется невырожденное преобразование координат
$$x_2'' = x_2' + \sum_{k=3}^n \frac{a_{2k}'}{a_{22}'}x_k' \text{ и } x_j'' = x_j' \text{ при } j \neq 2$$
и квадратичная форма $\mathcal{A}(x, x)$ приводится к виду $\mathcal{A}(x, x) = a_{11}x_1''^2 + a_{22}'x_2''^2 + v(x_3'', \ldots, x_n'')$, где $v(x_3'', \ldots, x_n'')$ – квадратичная форма от переменных $x_3'', \ldots, x_n''$, а ее матрица – к виду
$$A_2 = \begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22}' & 0 & \ldots & 0 \\ 0 & 0 & a_{33}'' & \ldots & a_{3n}'' \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & a_{n3}'' & \ldots & a_{nn}'' \end{bmatrix}.$$

По-прежнему угловые миноры A_2 совпадают с $\Delta_1, \Delta_2, \ldots, \Delta_n$, поэтому
$$a_{33}'' = \Delta_3/\Delta_2 \neq 0. \qquad (101.11)$$

Повторяя этот процесс, через $(n-1)$ шагов мы придем к базису, в котором матрица квадратичной формы $\mathcal{A}(x, x)$ имеет диагональную форму: $A_{n-1} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, где с учетом (101.10), (101.11) и обозначения $\Delta_0 = 1$
$$\lambda_i = \Delta_i/\Delta_{i-1}, \quad i = \overline{1,n}. \qquad (101.12)$$

Отметим, что каждый шаг (для определенности i-й шаг) процесса начинается с проверки условия $\Delta_i \neq 0$, так как только при выполнении этого условия можно провести i-й шаг.

Метод приведения к каноническому виду, описанный здесь, называется *методом Лагранжа*.

II. Пусть теперь среди угловых миноров $\Delta_1, \ldots, \Delta_{n-1}$ могут встретиться нулевые. Модифицируем метод Лагранжа.

Опишем i-й шаг. Пусть после $(i-1)$-го шага матрица квадратичной формы $\mathcal{A}(x, x)$ имеет вид
$$A_{i-1} = \begin{bmatrix} a_{11} & & & O \\ & \ddots & & \\ & & a_{i-1,i-1} & \\ O & & & \boxed{C} \end{bmatrix}, \text{ где } C = \begin{bmatrix} a_{ii} & \ldots & a_{in} \\ \ldots & \ldots & \ldots \\ a_{ni} & \ldots & a_{nn} \end{bmatrix}$$
(здесь штрихи опущены для упрощения записи). Будем считать, что

форму, или, в матричной формулировке: *квадратная матрица (комплексная или вещественная) является самосопряженной тогда и только тогда, когда она унитарно подобна вещественной диагональной матрице.*

§99. Знакоопределенные операторы

Теорема 99.1. *Если в унитарном пространстве V для любого вектора x имеет место равенство $(x, \mathcal{B}x) = 0$, то $\mathcal{B} = \mathcal{O}$.*

Доказательство. Для любых векторов $y, z \in V$ имеем $(\mathcal{B}(y+z), y+z) = 0$ и $(\mathcal{B}(iy+z), iy+z) = 0$, т.е. $(\mathcal{B}y,y)+(\mathcal{B}z,y)+(\mathcal{B}y,z)+(\mathcal{B}z,z) = 0$, $(\mathcal{B}y,y) - i(\mathcal{B}z,y) + i(\mathcal{B}y,z) + (\mathcal{B}z,z) = 0$ или, с учетом условия леммы, $(\mathcal{B}z,y)+(\mathcal{B}y,z) = 0$, $-i(\mathcal{B}z,y)+i(\mathcal{B}y,z) = 0$. Прибавив к первому равенству второе, умноженное на i, получим, что $(\mathcal{B}y,z) = 0$, $\forall y,z \in V$. Следовательно, $\mathcal{B} = \mathcal{O}$ (теорема 94.1). ■

Теорема 99.2. *Линейный оператор $\mathcal{A}$ в унитарном пространстве V эрмитов тогда и только тогда, когда*

$$(\mathcal{A}x, x) \in \mathbb{R}, \quad \forall x \in V. \tag{99.1}$$

Доказательство. Необходимость. Если $\mathcal{A}$ – эрмитов оператор, то $(\mathcal{A}x,x) = (x, \mathcal{A}^*x) = \overline{(\mathcal{A}x, x)}$, $\forall x \in V$; следовательно,

$(\mathcal{A}x,x) = \overline{(\mathcal{A}x,x)}$ и $(\mathcal{A}x,x) \in \mathbb{R}$, $\forall x \in V$.

Достаточность. Пусть $(\mathcal{A}x,x) \in \mathbb{R}$. Тогда $(\mathcal{A}x,x) = (x,\mathcal{A}x)$ и $(\mathcal{A}x,x) = (x,\mathcal{A}^*x)$. Отсюда следует, что

$$(x, (\mathcal{A}-\mathcal{A}^*)x) = 0, \quad \forall x \in V. \tag{99.2}$$

Отсюда на основании теоремы 99.1 следует, что $\mathcal{A} = \mathcal{A}^*$. ■

Замечание. Условие (99.1) формально верно и в евклидовом пространстве, однако лишено смысла, так как справедливо для любых линейных операторов $\mathcal{A}$.

Теорема 99.2 позволяет для самосопряженных операторов $\mathcal{A}$ говорить о знаке числа $(\mathcal{A}x,x)$. Самосопряженный оператор в унитарном (евклидовом) пространстве называется *положительно определенным*, если

$$(\mathcal{A}x,x) > 0, \quad \forall x \neq \theta,$$

неотрицательно определенным (отрицательно определенным или *неположительно определенным)*, если $(\mathcal{A}x,x) \geq 0$, $\forall x \neq \theta$ (соответственно $(\mathcal{A}x,x) < 0$ или $(\mathcal{A}x,x) \leq 0$). Обозначение: $\mathcal{A} > \mathcal{O}$, $\mathcal{A} \geq \mathcal{O}$, $\mathcal{A} < \mathcal{O}$, $\mathcal{A} \leq \mathcal{O}$ соответственно.

Теорема 99.3. *Самосопряженный оператор $\mathcal{A}$ в унитарном (и евклидовом) пространстве положительно определен (соответствено $\mathcal{A} \geq \mathcal{O}$, $\mathcal{A} < \mathcal{O}$, $\mathcal{A} \leq \mathcal{O}$) тогда и только тогда, когда все его собственные значения $\lambda > 0$ ($\lambda \geq 0$, $\lambda < 0$, $\lambda \leq 0$).*

Доказательство. Докажем вариант теоремы, относящийся к $\mathcal{A} > \mathcal{O}$. Необходимость. Если $\mathcal{A} > \mathcal{O}$, то $(\mathcal{A}x,x) > 0$ для любого вектора $x \neq \theta$. В частности, это верно и для собственного вектора x, поэтому $(\mathcal{A}x,x) = (\lambda x, x) = \lambda(x,x) > 0$, где $(x,x) > 0$. Значит, $\lambda > 0$.

Достаточность. Если $\mathcal{A}$ – самосопряженный оператор, то существует ортонормированный базис из собственных векторов $e_1, \ldots, e_n$ оператора $\mathcal{A}$. При этом соответствующие собственные значения $\lambda_i > 0$, $i = \overline{1,n}$. Тогда для любого вектора $x = \sum_{i=1}^n x_i e_i \neq \theta$ имеем

$$(\mathcal{A}x,x) = \left(\sum_{i=1}^n x_i\lambda_i e_i, \sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n \lambda_i |x_i|^2 > 0. \ ■$$

Следствие. Если $\mathcal{A} > \mathcal{O}$ (или $\mathcal{A} < \mathcal{O}$), то $\mathcal{A}$ обратим, так как $\det \mathcal{A} = \lambda_1 \cdot \ldots \cdot \lambda_n$.

Теорема 99.4. *Оператор, обратный к положительно (отрицательно) определенному оператору, положительно (соответственно отрицательно) определен.*

Доказательство. Если $\mathcal{A}$ – самосопряженный оператор, то $\mathcal{A}^{-1}$ – тоже самосопряженный оператор, так как согласно теореме 94.4 $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1} = \mathcal{A}^{-1}$. Если $\mathcal{A} > \mathcal{O}$ ($\mathcal{A} < \mathcal{O}$), то все собственные значения оператора $\mathcal{A}^{-1}$ положительны (соответственно отрицательны), так как они обратны собственным значениям оператора $\mathcal{A}$. Из теоремы 99.3 следует, что $\mathcal{A}^{-1} > \mathcal{O}$ ($\mathcal{A}^{-1} < \mathcal{O}$). ■

Теорема 99.5. *Для любого неотрицательно (положительно) определенного оператора $\mathcal{A}$ существует, и притом единственный, неотрицательно (соответственно положительно) определенный оператор $\mathcal{B}$ такой, что $\mathcal{B}^2 = \mathcal{A}$.*

Доказательство. Существование. Построим оператор $\mathcal{B}$. Пусть $e_1, \ldots, e_n$ – ортонормированный базис пространства из собственных векторов оператора $\mathcal{A}$ и $\mathcal{A}e_i = \lambda_i e_i$, $i = \overline{1,n}$. По условию теоремы $\lambda_i \geq 0$ ($\lambda_i > 0$), $i = \overline{1,n}$. Положим

$$\mathcal{B}e_i = \sqrt{\lambda_i}\, e_i, \quad i = \overline{1,n}. \tag{99.3}$$

Из (99.3) следует, что оператор $\mathcal{B} \geq \mathcal{O}$ ($\mathcal{B} > \mathcal{O}$), так как $\mathcal{B}$ – нормальный оператор (ибо существует ортонормированный базис $e_1, \ldots, e_n$ из собственных векторов $\mathcal{B}$), при этом он самосопряжен (ибо его собственные значения $\sqrt{\lambda_i} \in \mathbb{R}$) и, кроме того, $\sqrt{\lambda_i} \geq 0$ ($\sqrt{\lambda_i} > 0$). Оператор $\mathcal{B}$ – искомый, так как в силу (99.3) $\mathcal{B}^2 e_i = \lambda_i e_i = \mathcal{A}e_i$, $i = \overline{1,n}$.

Единственность. Пусть существует другой оператор $\mathcal{C} \geq \mathcal{O}$ ($\mathcal{C} > \mathcal{O}$) такой, что $\mathcal{C}^2 = \mathcal{A}$. Тогда существует ортонормированный базис $f_1, \ldots, f_n$ пространства из собственных векторов оператора $\mathcal{C}$. Если $\mathcal{C}f_i = \mu_i f_i$, $i = \overline{1,n}$, то $\mathcal{A}f_i = \mathcal{C}^2 f_i = \mu_i^2 f_i$, $i = \overline{1,n}$. Значит, числа $\mu_1^2, \ldots, \mu_n^2$ являются собственными значениями оператора $\mathcal{A}$ и, следовательно, совпадают с числами $\lambda_1, \ldots, \lambda_n$. Покажем, что

$$\mathcal{C}e_i = \sqrt{\lambda_i}\, e_i, \quad i = \overline{1,n}, \tag{99.4}$$

(92.1) имеют вид

$$\left[\begin{array}{c} \boxed{J_q(\lambda_j)} \\ O \end{array}\right], \tag{92.2}$$

где $J_q(\lambda_j)$ – клетка Жордана (88.7) q-го порядка с λ_j на главной диагонали.

Точно так же устроены столбцы матрицы $\mathcal{A}|K_{\lambda_j}$, определяемые векторами второго столбца жордановой лестницы: диагональная клетка имеет тот же вид $J_q(\lambda_j)$, а все остальные элементы равны нулю. Таким образом, первая группа из t_q столбцов жордановой лестницы порождает клетки Жордана q-го порядка с λ_j на главной диагонали. Число этих клеток равно t_q.

2. Следующая группа из t_{q-1} столбцов жордановой лестницы определяет клетки $J_{q-1}(\lambda_j)$ на главной диагонали матрицы $\mathcal{A}|K_{\lambda_j}$. Число клеток $(q-1)$-го порядка равно t_{q-1}.

3. Рассмотрев все столбцы жордановой лестницы, получим матрицу A_j оператора $\mathcal{A}|K_{\lambda_j}$ в каноническом базисе. Из (92.2) и последующих рассуждений вытекает, что A_j – квазидиагональная матрица с клетками Жордана $J_k(\lambda_j)$ на главной диагонали. Всего этих клеток столько, сколько столбцов в жордановой лестнице, т.е. n_1 или, согласно (91.9), s_j (геометрическая кратность λ_j). Таким образом,

$$A_j = \begin{bmatrix} \boxed{J_{q_1}(\lambda_j)} & & & O \\ & \boxed{J_{q_2}(\lambda_j)} & & \\ & & \ddots & \\ O & & & \boxed{J_{q_{s_j}}(\lambda_j)} \end{bmatrix}, \tag{92.3}$$

где $q_1 + \ldots + q_{s_j} = \dim K_{\lambda_j} = m_j$, а число клеток k-го порядка равно числу $t_k = -n_{k-1} + 2n_k - n_{k+1}$, $k = \overline{1,q}$.

4. Несмотря на некоторый произвол в выборе базиса, матрица A_j определена однозначно с точностью до порядка клеток, так как количество всех клеток равно геометрической кратности s_j собственного значения λ_j, а количество клеток k-го порядка равно числу $t_k = -n_{k-1} + 2n_k - n_{k+1}$, $k = \overline{1,q}$, или, согласно (80.2), $t_k = r_{k-1} - 2r_k + r_{k+1}$, $k = \overline{1,q}$.

Таким образом, структура клеток Жордана в матрице A_j определяется только оператором $\mathcal{A}$. Порядок клеток определен порядком нумерации столбцов жордановой лестницы.

Жорданова форма матрицы линейного оператора в комплексном пространстве. Будем придерживаться прежних обозна-

чений: m_j и s_j – алгебраическая и геометрическая кратности λ_j, $r_k = \mathrm{rg}(\mathcal{A} - \lambda_j\mathcal{I})^k$.

Теорема 92.2. *Пусть $\mathcal{A} \in \mathcal{L}(V,V)$ – линейный оператор, действующий в комплексном пространстве V, и его характеристический многочлен имеет вид*

$f(\lambda) = a_n(\lambda_1 - \lambda)^{m_1}\ldots(\lambda_p - \lambda)^{m_p}$, где $\lambda_i \neq \lambda_j$ при $i \neq j$.

Тогда в пространстве V существует базис e, в котором матрица оператора $\mathcal{A}$ имеет квазидиагональную форму

$$A_e = \begin{bmatrix} \boxed{A_1} & & & O \\ & \boxed{A_2} & & \\ & & \ddots & \\ O & & & \boxed{A_p} \end{bmatrix}, \tag{92.4}$$

в которой матрицы A_j, $j = \overline{1,p}$, имеют вид (92.3).

Доказательство. В соответствии с теоремой 91.1, $V = K_{\lambda_1} \oplus \ldots \oplus K_{\lambda_p}$. В качестве искомого базиса e возьмем совокупность канонических базисов корневых подпространств $K_{\lambda_1}, \ldots, K_{\lambda_p}$. Согласно теореме 84.2 матрица A_e имеет вид (92.4), где A_j – матрица оператора $\mathcal{A}|K_{\lambda_j}$ в каноническом базисе K_{λ_j}. Следовательно, матрица A_j имеет вид (92.3). ■

Следствие. Для собственных значений оператора $\mathcal{A}$ имеют место соотношения $\lambda_1 + \ldots + \lambda_n = \mathrm{tr}\,\mathcal{A}$, $\lambda_1 \cdot \ldots \cdot \lambda_n = \det \mathcal{A}$.

Полученная форма матрицы линейного оператора называется *жордановой формой*, а построенный базис – *каноническим* (или *жордановым*) *базисом*.

Замечание 1. *Жорданова форма матрицы линейного оператора определена однозначно с точностью до порядка клеток Жордана.* Порядок клеток определен порядком нумерации корневых подпространств и тем порядком, в котором нумеруются столбцы жордановой лестницы в каноническом базисе K_{λ_j}.

Замечание 2. *Для операторов простой структуры, и только для них, жорданова форма совпадает с диагональной*, так как критерий (88.4) с учетом (91.11) равносилен условию

$$W_{\lambda_j} = K_{\lambda_j}, \quad j = \overline{1,p}. \tag{92.5}$$

Замечание 3. В матричной формулировке теорема 92.2 означает, что *любая квадратная комплексная матрица подобна матрице, имеющей жорданову форму.* Это может быть доказано по известной схеме (88.6).

Матрица A_e, имеющая жорданову форму и подобная матрице A, называется *жордановой формой матрицы A.*

Теорема 92.3. *Две матрицы $A, B \in \mathbb{C}^{n\times n}$ подобны тогда и только тогда, когда их жордановы формы совпадают.*

образуют ортонормированный базис $\mathrm{im}\,\mathcal{A}$. Отметим, что тем самым получено (в силу (64.3)) равенство

$$r = \mathrm{rg}\,\mathcal{A}. \tag{100.7}$$

Дополним векторы $g_1, \ldots, g_r$ до ортонормированного базиса $g_1, \ldots, g_r, g_{r+1}, \ldots, g_n$ всего пространства. Тогда из (100.6) с учетом (100.4) получим, что

$$\mathcal{A}e_k = \rho_k g_k, \quad k = \overline{1,n}. \tag{100.8}$$

3. Построим операторы $\mathcal{B}$ и $\mathcal{U}$ из (100.3). Положим

$$\mathcal{U}e_k = g_k, \quad \mathcal{B}g_k = \rho_k g_k, \ k = \overline{1,n}. \tag{100.9}$$

Тогда $\mathcal{U}$ – унитарный оператор, так как переводит ортонормированный базис e в ортонормированный базис g; $\mathcal{B} \geq \mathcal{O}$ как нормальный оператор (ибо ортонормированный базис g состоит из его собственных векторов), собственные значения которого $\rho_k \geq 0$, $k = \overline{1,n}$. При этом $\mathcal{A} = \mathcal{BU}$, так как, согласно (100.8) и (100.9), $(\mathcal{BU})e_k = \mathcal{B}(\mathcal{U}e_k) = \mathcal{B}g_k = \rho_k g_k = \mathcal{A}e_k$, $k = \overline{1,n}$.

4. Докажем единственность. Пусть $\mathcal{A} = \mathcal{BU}$ – разложение (100.3) для оператора $\mathcal{A}$. Тогда $\mathcal{A}^* = \mathcal{U}^*\mathcal{B}$ и $\mathcal{AA}^* = \mathcal{B}^2$. Таким образом, $\mathcal{B}$ – квадратный корень из оператора $\mathcal{AA}^*$, который на основании теоремы 99.4 существует и определен однозначно. Если оператор $\mathcal{A}$ обратим, то из (100.5) и (100.7) следует, что обратим и оператор $\mathcal{A}^*\mathcal{A}$, при этом $\rho_k \neq 0$, $k = \overline{1,n}$, поэтому согласно (100.9) обратим оператор $\mathcal{B}$. В силу (100.3) отсюда следует, что $\mathcal{U} = \mathcal{B}^{-1}\mathcal{A}$. ■

Отметим, что попутно доказаны следующие факты:

а) $\mathrm{rg}\,\mathcal{A} = \mathrm{rg}\,\mathcal{A}^*\mathcal{A} = \mathrm{rg}\,\mathcal{AA}^*$ (это следует из (100.5), (100.7) и равенства $\mathcal{AA}^* = \mathcal{B}^2$);

б) оператор $\mathcal{U}$ переводит ортонормированный базис из собственных векторов оператора $\mathcal{A}^*\mathcal{A}$ в ортонормированный базис из собственных векторов оператора $\mathcal{AA}^*$ (это следует из (100.9));

в) операторы $\mathcal{A}^*\mathcal{A}$ и $\mathcal{AA}^*$ имеют одинаковые собственные значения (это следует из того, что $\mathcal{AA}^* = \mathcal{B}^2$).

Арифметические значения квадратных корней из собственных значений оператора $\mathcal{A}^*\mathcal{A}$ называются *сингулярными числами* оператора $\mathcal{A}$.

Разложение (100.3) называется *полярным разложением оператора $\mathcal{A}$*.

Замечание. Для любого оператора $\mathcal{A}$ имеет место разложение вида $\mathcal{A} = \mathcal{VC}$, где $\mathcal{V}$ – унитарный оператор, $\mathcal{C} \geq \mathcal{O}$. Оно может быть получено из полярного разложения сопряженного оператора $\mathcal{A}^* = \mathcal{BU}$ переходом к оператору $\mathcal{A}$: $\mathcal{A} = \mathcal{U}^*\mathcal{B}$, где $\mathcal{U}^*$, очевидно, унитарный оператор.

Теорема 100.5. *Оператор $\mathcal{A}$ нормален тогда и только тогда, когда в любом его полярном разложении операторы $\mathcal{B}$ и $\mathcal{U}$ перестановочны.*

Доказательство. Пусть $\mathcal{A} = \mathcal{BU}$ – полярное разложение оператора $\mathcal{A}$. Тогда $\mathcal{AA}^* = \mathcal{B}^2$, $\mathcal{A}^*\mathcal{A} = \mathcal{U}^*\mathcal{B}^2\mathcal{U}$. Достаточность очевидна, так как если $\mathcal{BU} = \mathcal{UB}$, то $\mathcal{A}^*\mathcal{A} = \mathcal{U}^*\mathcal{B}^2\mathcal{U} = \mathcal{U}^*\mathcal{B}(\mathcal{BU}) = \mathcal{U}^*\mathcal{B}(\mathcal{UB}) = \mathcal{U}^*(\mathcal{BU})\mathcal{B} = \mathcal{U}^*(\mathcal{UB})\mathcal{B} = \mathcal{U}^*\mathcal{UB}^2 = \mathcal{B}^2 = \mathcal{AA}^*$.

Необходимость. Если $e_1, \ldots, e_n$ – ортонормированный базис из собственных векторов $\mathcal{A}^*\mathcal{A}$, то

$$\mathcal{A}^*\mathcal{A}e_k = \rho_k^2 e_k, \quad k = \overline{1,n}. \tag{100.10}$$

С другой стороны, $\mathcal{A}^*\mathcal{A} = \mathcal{U}^*\mathcal{B}^2\mathcal{U}$ и, значит, $\mathcal{U}^*\mathcal{B}^2\mathcal{U}e_k = \rho_k^2 e_k$, т.е.

$$\mathcal{B}^2\mathcal{U}e_k = \rho_k^2\mathcal{U}e_k, \quad k = \overline{1,n}. \tag{100.11}$$

Так как $\mathcal{U}$ унитарен, то векторы $\mathcal{U}e_1, \ldots, \mathcal{U}e_n$ образуют ортонормированный базис, поэтому в соответствии с (99.3) из (100.11) получим

$$\mathcal{BU}e_k = \rho_k\mathcal{U}e_k, \quad k = \overline{1,n}. \tag{100.12}$$

С другой стороны, $\mathcal{A}^*\mathcal{A} = \mathcal{AA}^* = \mathcal{B}^2$ и согласно (100.10) $\mathcal{B}^2 e_k = \rho_k^2 e_k$ или, в соответствии с (99.3), $\mathcal{B}e_k = \rho_k e_k$, откуда умножением на $\mathcal{U}$ слева получим, что

$$\mathcal{UB}e_k = \rho_k\mathcal{U}e_k, \quad k = \overline{1,n}. \tag{100.13}$$

Сравнение равенств (100.12) и (100.13) доказывает, что $\mathcal{BU} = \mathcal{UB}$. ■

Глава XVI. Линейные операторы в унитарных (евклидовых) пространствах

§94. Сопряженный оператор

В этом параграфе V и W – два пространства, оба унитарных или оба евклидовых.

Теорема 94.1. *Если $\mathcal{A}, \mathcal{B}$ – линейные операторы из $\mathcal{L}(V,W)$ и $(\mathcal{A}x, y) = (\mathcal{B}x, y)$, $\forall x \in V$, $y \in W$, то $\mathcal{A} = \mathcal{B}$.*

Доказательство. Утверждение вытекает из следующей цепочки простейших импликаций:

$(\mathcal{A}x,y) = (\mathcal{B}x,y)$, $\forall x \in V$, $y \in W$, $\Rightarrow$ $(\mathcal{A}x - \mathcal{B}x, y) = 0$, $\forall y \in W$, $\Rightarrow$
$\Rightarrow$ $\mathcal{A}x - \mathcal{B}x = \theta$ $\Rightarrow$ $\mathcal{A}x = \mathcal{B}x$, $\forall x \in V$, $\Rightarrow$ $\mathcal{A} = \mathcal{B}$. ■

Замечание. Очевидно, что из равенства $(x, \mathcal{A}y) = (x, \mathcal{B}y)$, $\forall x \in W$, $y \in V$, также следует, что $\mathcal{A} = \mathcal{B}$.

Пусть $\mathcal{A} \in \mathcal{L}(V,W)$. Отображение $\mathcal{A}^* : W \to V$ называется *сопряженным оператором к оператору $\mathcal{A}$*, если

$$(\mathcal{A}x, y) = (x, \mathcal{A}^*y), \quad \forall x \in V, y \in W. \tag{94.1}$$

Теорема 94.2. *Сопряженный оператор линеен.*

Доказательство. Пусть $y_1, y_2 \in W$, тогда согласно (94.1)

$$(\mathcal{A}x, y_1 + y_2) = (x, \mathcal{A}^*(y_1 + y_2)). \tag{94.2}$$

С другой стороны, $(\mathcal{A}x, y_1+y_2) = (\mathcal{A}x, y_1) + (\mathcal{A}x, y_2) = (x, \mathcal{A}^*y_1) + (x, \mathcal{A}^*y_2) = (x, \mathcal{A}^*y_1 + \mathcal{A}^*y_2)$. Отсюда с учетом (94.2) получим, что $(x, \mathcal{A}^*(y_1+y_2) - (\mathcal{A}^*y_1 + \mathcal{A}^*y_2)) = 0$, $\forall x \in V$, следовательно,

$$\mathcal{A}^*(y_1+y_2) = \mathcal{A}^*y_1 + \mathcal{A}^*y_2, \quad \forall y_1, y_2 \in W. \tag{94.3}$$

Аналогично показывается, что

$$\mathcal{A}^*(\alpha y) = \alpha\mathcal{A}^*y, \quad \forall y \in W, \forall\alpha \in \mathbb{C}(\mathbb{R}). \tag{94.4}$$

Свойства (94.3), (94.4) доказывают линейность оператора $\mathcal{A}^*$. ■

Теорема 94.3. *Для любого оператора $\mathcal{A} \in \mathcal{L}(V,W)$ существует, и притом единственный, сопряженный оператор.*

Доказательство. Существование. Пусть $e_1, \ldots, e_n$ – ортонормированный базис V. Тогда для любого вектора $x \in V$ согласно

(70.6) имеет место разложение $x = \sum_{k=1}^n (x, e_k)e_k$. Следовательно, $\mathcal{A}x = \sum_{k=1}^n (x,e_k)\mathcal{A}e_k$ и в силу (70.7)

$$(\mathcal{A}x, y) = \sum_{k=1}^n (x, e_k)(\mathcal{A}e_k, y), \quad \forall y \in W. \tag{94.5}$$

Покажем, что оператор $\mathcal{B} \in \mathcal{L}(W,V)$, определенный равенством

$$\mathcal{B}y = \sum_{k=1}^n (y, \mathcal{A}e_k)e_k, \quad \forall y \in W,$$

является сопряженным к $\mathcal{A}$. Действительно, из того, что $(x, \mathcal{B}y) = (x, \sum_{k=1}^n (y,\mathcal{A}e_k)e_k) = \sum_{k=1}^n (\mathcal{A}e_k, y)(x, e_k)$, и равенства (94.5) следует, что $(\mathcal{A}x, y) = (x, \mathcal{B}y)$, $\forall x \in V$, $y \in W$, т.е. $\mathcal{B} = \mathcal{A}^*$.

Единственность вытекает из теоремы 94.1, так как для любых двух операторов $\mathcal{B}$ и $\mathcal{C}$, сопряженных к $\mathcal{A}$, имеет место соотношение $(x, \mathcal{B}y) = (\mathcal{A}x, y) = (x, \mathcal{C}y)$, $\forall x \in V$, $y \in W$. ■

Теорема 94.4. *Операция сопряжения линейного оператора обладает следующими свойствами:*

1) $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$,
2) $(\alpha\mathcal{A})^* = \overline{\alpha}\mathcal{A}^*$,
3) $(\mathcal{AB})^* = \mathcal{BA}$,
4) $(\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^*$,
5) $(\mathcal{A}^*)^* = \mathcal{A}$,

выполненными для любых операторов, для которых определены указанные операции.

Доказательство. Все свойства доказываются однотипно. Докажем свойства 3 и 4. Свойство 3 вытекает из (94.1) и теоремы 94.1, так как $(x, (\mathcal{AB})^*y) = (\mathcal{AB}x, y) = (\mathcal{B}x, \mathcal{A}^*y) = (x, \mathcal{B}^*\mathcal{A}^*y)$, $\forall x, y$. Свойство 4 (очевидно, для невырожденных $\mathcal{A} \in \mathcal{L}(V,V)$) вытекает из свойства 3, так как если $\mathcal{AA}^{-1} = \mathcal{A}^{-1}\mathcal{A} = \mathcal{I}$, то $(\mathcal{A}^{-1})^*\mathcal{A}^* = \mathcal{A}^*(\mathcal{A}^{-1})^* = \mathcal{I}$. Это означает, что оператор $(\mathcal{A}^{-1})^*$ является обратным к $\mathcal{A}^*$. ■

§95. Биортогональные базисы

Здесь и далее будем рассматривать линейные операторы, действующие в одном пространстве V, унитарном или евклидовом.

Две системы векторов $x_1, \ldots, x_k$ и $y_1, \ldots, y_k$ в унитарном (евклидовом) пространстве называются *биортогональными системами*, если $(x_i, y_j) = \delta_{ij}$, где δ_{ij} – символ Кронекера (§70).

Каждая из двух биортогональных систем векторов линейно независима. В этом можно убедиться, если приравнять линейную комбинацию одной системы нулевому вектору и последовательно умножать обе части полученного равенства скалярно на векторы другой системы.

Биортогональные системы $e_1, \ldots, e_n$ и $f_1, \ldots, f_n$, образующие базисы пространства V, называют *биортогональной парой базисов.*

Теорема 101.4. *Билинейная форма симметрична тогда и только тогда, когда ее матрица в любом базисе симметрична.*

Доказательство. Необходимость проверяется непосредственно. Достаточность. Если $A_e^T = A_e$, то, согласно (101.3), $\mathcal{A}(x,y) = y_e^T A_e^T x_e = y_e^T A_e x_e = \mathcal{A}(y,x)$. ■

Все матрицы одной билинейной формы имеют одинаковый ранг (см. следствие из теоремы 101.3). *Рангом билинейной формы* называется ранг ее матрицы в произвольном базисе. Билинейная форма $\mathcal{A}(x,y)$ называется *вырожденной*, если $\mathrm{rg}\,\mathcal{A}(x,y) < \dim V$, и *невырожденной*, если $\mathrm{rg}\,\mathcal{A}(x,y) = \dim V$.

Теорема 101.5. *Билинейная форма $\mathcal{A}(x,y)$ вырождена тогда и только тогда, когда существует вектор $x \neq \theta$ такой, что*

$$\mathcal{A}(x,y) = 0, \quad \forall y \in V. \tag{101.4}$$

Доказательство. Пусть $e_1, \ldots, e_n$ – базис V и $A_e = (a_{ij})$ – матрица билинейной формы $\mathcal{A}(x,y)$ в этом базисе. Соотношение (101.4) равносильно системе равенств $\mathcal{A}(x, e_j) = 0$, $j = \overline{1,n}$, или, если учесть разложение $x = \sum_{i=1}^n x_i e_i$, системе равенств $\sum_{i=1}^n x_i\mathcal{A}(e_i, e_j) = 0$, $j = \overline{1,n}$, т.е.

$$\left\{\begin{array}{l} a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n = 0, \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{nn}x_n = 0. \end{array}\right. \tag{101.5}$$

Значит, билинейная форма $\mathcal{A}(x,y)$ вырождена тогда и только тогда, когда однородная система линейных уравнений (101.5) имеет нетривиальное решение, т.е. (теорема 28.5) когда $\mathrm{rg}\,A_e < n$. ■

Квадратичные формы. Пусть $\mathcal{A}(x,y)$ – симметричная билинейная форма в пространстве V над полем P. *Квадратичной формой* называется отображение $\mathcal{A}: V \to P$, которое каждому вектору $x \in V$ ставит в соответствие число $\mathcal{A}(x,x)$. Обозначение: $\mathcal{A}(x,x)$ или $\mathcal{A}(x)$. Билинейная форма $\mathcal{A}(x,y)$ при этом называется *полярной билинейной формой* к квадратичной форме $\mathcal{A}(x,x)$.

Теорема 101.6. *Полярная билинейная форма для любой квадратичной формы определена однозначно.*

Это утверждение вытекает из того, что если $\mathcal{A}(x,y)$ – полярная билинейная форма для квадратичной формы $\mathcal{A}(x,x)$, то для любых $x, y \in V$

$$\mathcal{A}(x,y) = \tfrac{1}{2}(\mathcal{A}(x+y, x+y) - \mathcal{A}(x,x) - \mathcal{A}(y,y)). \quad ■$$

Таким образом, полярная билинейная форма $\mathcal{A}(x,y)$ однозначно восстанавливается по квадратичной форме $\mathcal{A}(x,x)$. Единственность полярной билинейной формы $\mathcal{A}(x,y)$ позволяет переносить ее характеристики на квадратичную форму $\mathcal{A}(x,x)$.

Матрицей квадратичной формы $\mathcal{A}(x,x)$ в базисе e называется матрица полярной к ней билинейной формы $\mathcal{A}(x,y)$ в этом базисе.

Из свойств билинейной формы вытекают следующие свойства квадратичных форм.

1°. Матрица квадратичной формы симметрична.

2°. Любая симметрическая матрица является матрицей единственной квадратичной формы в заданном базисе.

3°. Матрицы квадратичной формы в базисах e и $f = eQ$ связаны соотношением

$$A_f = Q^T A_e Q. \tag{101.6}$$

4°. В базисе e квадратичная форма $\mathcal{A}(x,x)$ с матрицей $A_e = (a_{ij})$ может быть записана в следующем виде: $\forall x = \sum_{i=1}^n x_i e_i$

$$\mathcal{A}(x,x) = \sum_{i,j=1}^n a_{ij}x_i x_j, \quad a_{ij} = a_{ji}, \tag{101.7}$$

или, в компактной форме,

$$\mathcal{A}(x,x) = x_e^T A_e x_e, \quad A_e^T = A_e. \tag{101.8}$$

Представление квадратичной формы в виде (101.7) или (101.8) называется *общим видом квадратичной формы $\mathcal{A}(x,x)$ в базисе e.* Выражение $f(x_1, \ldots, x_n) = \sum_{i,j=1}^n a_{ij}x_i x_j$, где $a_{ij} = a_{ji}$, называется *квадратичной формой от переменных $x_1, \ldots, x_n$.*

5°. *Рангом квадратичной формы* называется ранг ее матрицы в произвольном базисе. Очевидно, $\mathrm{rg}\,\mathcal{A}(x,x) = \mathrm{rg}\,\mathcal{A}(x,y)$. Квадратичная форма $\mathcal{A}(x,x)$ называется *вырожденной*, если $\mathrm{rg}\,\mathcal{A}(x,x) < \dim V$, и *невырожденной*, если $\mathrm{rg}\,\mathcal{A}(x,x) = \dim V$.

Канонический вид квадратичной формы. Выбирая подходящим образом базис V, можно менять вид матрицы квадратичной формы и, следовательно, ее общий вид (101.7).

Базис $e = (e_1, \ldots, e_n)$ называется *каноническим базисом квадратичной формы $\mathcal{A}(x,x)$*, если матрица квадратичной формы в этом базисе диагональна: $A_e = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

В каноническом базисе квадратичная форма $\mathcal{A}(x,x)$ согласно (101.7) имеет вид $\mathcal{A}(x,x) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$, который называется *каноническим видом* квадратичной формы, а числа $\lambda_1, \ldots, \lambda_n$ – ее *каноническими коэффициентами.* Канонический вид называют также *суммой квадратов.* Очевидно, что число ненулевых квадратов совпадает с рангом $\mathcal{A}(x,x)$. Итак, если e – канонический базис и $r = \mathrm{rg}\,\mathcal{A}(x,x)$, то

$$\mathcal{A}(x,x) = \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2, \quad \forall x = \sum_{i=1}^n x_i e_i. \tag{101.9}$$

Теорема 101.7. *Для любой квадратичной формы существует канонический базис.*

Теорема 96.2. *Собственные векторы нормального оператора, отвечающие различным собственным значениям, попарно ортогональны.*

Доказательство. Пусть $\mathcal{A}x = \lambda x$, $\mathcal{A}y = \mu y$, $\lambda \neq \mu$. Тогда $(\mathcal{A}x, y) = (\lambda x, y) = \lambda(x,y)$. С другой стороны, $(\mathcal{A}x, y) = (x, \mathcal{A}^*y) = \{$ в силу теоремы 96.1 $\} = (x, \overline{\mu}y) = \mu(x,y)$. Следовательно, $\lambda(x,y) = \mu(x,y)$ и так как $\lambda \neq \mu$, то $(x,y) = 0$. ■

Теорема 96.3 (теорема Шура). *Для любого оператора, действующего в унитарном пространстве, существует ортонормированный базис, в котором он имеет треугольную матрицу.*

Доказательство этой теоремы отличается от доказательства теоремы 89.2 только тем, что на каждом шаге строится ортонормированный базис инвариантного подпространства. ■

Ортонормированный базис унитарного (евклидова) пространства, в котором матрица линейного оператора имеет треугольную форму, называется *базисом Шура* для этого оператора.

Теорема 96.4 (критерий нормальности). *Оператор, действующий в унитарном пространстве, нормален тогда и только тогда, когда существует ортонормированный базис из собственных векторов этого оператора.*

Доказательство. Необходимость. Пусть $\mathcal{A}$ – нормальный оператор и e – его базис Шура. Тогда

$$A_e = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix}, \quad (A_e)^H = \begin{bmatrix} \overline{a_{11}} & 0 & \ldots & 0 \\ \overline{a_{12}} & \overline{a_{22}} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \overline{a_{1n}} & \overline{a_{2n}} & \ldots & \overline{a_{nn}} \end{bmatrix}$$

и, согласно (95.2) и замечанию 1, $A_e(A_e)^H = (A_e)^H A_e$. Сравнив диагональные элементы матриц, стоящих в левой и правой частях последнего соотношения, получим равенства

$$|a_{11}|^2 + |a_{12}|^2 + \ldots + |a_{1n}|^2 = |a_{11}|^2, \ |a_{22}|^2 + \ldots + |a_{2n}|^2 = \\ = |a_{22}|^2, \ \ldots, |a_{n-1,n-1}|^2 + |a_{n-1,n}|^2 = |a_{n-1,n-1}|^2,$$

из которых следует, что

$a_{12} = a_{13} = \ldots = a_{1n} = 0$, $a_{23} = a_{24} = \ldots = a_{2n} = 0$, $\ldots$, $a_{n-1,n} = 0$.

Следовательно, матрица A_e имеет диагональную форму. Таким образом, базис Шура является ортонормированным базисом из собственных векторов оператора $\mathcal{A}$.

Достаточность. Пусть e – ортонормированный базис из собственных векторов оператора $\mathcal{A}$, тогда

$$A_e = \begin{bmatrix} \lambda_1 & & O \\ & \ddots & \\ O & & \lambda_n \end{bmatrix}, \quad (A_e)^H = \begin{bmatrix} \overline{\lambda_1} & & O \\ & \ddots & \\ O & & \overline{\lambda_n} \end{bmatrix}.$$

Из перестановочности диагональных матриц следует, что A_e – нормальная матрица. Это означает (замечание 1), что $\mathcal{A}$ – нормальный оператор. ■

Следствие 3. В унитарном пространстве нормальный оператор $\mathcal{A}$ и его сопряженный $\mathcal{A}^*$ имеют общий ортонормированный базис из собственных векторов.

Теорема 96.5. *Если любой собственный вектор оператора $\mathcal{A}$, действующего в унитарном пространстве V, является собственным вектором сопряженного оператора $\mathcal{A}^*$, то $\mathcal{A}$ – нормальный оператор.*

Доказательство. Любой оператор, действующий в комплексном пространстве, обладает хотя бы одним собственным вектором (теорема 86.6). Пусть $\dim V = n$ и e_1 – собственный вектор оператора $\mathcal{A}$, тогда e_1 – собственный вектор оператора $\mathcal{A}^*$ и согласно теореме 95.4 подпространство $L_{n-1} = \mathcal{L}^\perp(e_1)$ инвариантно относительно оператора $\mathcal{A}$. В этом подпространстве (в силу теоремы 86.6) существует собственный вектор $e_2 \in L_{n-1}$ оператора $\mathcal{A}$, при этом $e_2 \perp e_1$. Вектор e_2 также является собственным вектором $\mathcal{A}^*$, и по-прежнему $\mathcal{L}(e_2)$ инвариантно относительно $\mathcal{A}^*$ и, следовательно, $L_{n-2} = \mathcal{L}^\perp(e_2)$ (имеется в виду ортогональное дополнение $\mathcal{L}(e_2)$ до L_{n-1}) инвариантно относительно оператора $\mathcal{A}$. Поступая аналогично, построим ортогональную систему собственных векторов $e_1, \ldots, e_n$ оператора $\mathcal{A}$. В таком случае векторы $e_1/|e_1|, \ldots, e_n/|e_n|$ образуют ортонормированный базис пространства V, состоящий из собственных векторов оператора $\mathcal{A}$, и в силу теоремы 96.4 $\mathcal{A}$ – нормальный оператор. ■

Замечание 2. Отметим, что доказанная теорема является обратной к теореме 96.1 для случая унитарного пространства.

Унитарно подобные матрицы. Подобные комплексные (вещественные) матрицы A и $B = Q^{-1}AQ$ называются *унитарно* (соответственно *ортогонально*) *подобными*, если матрица преобразования подобия Q унитарна (соответственно ортогональна), т.е. если $Q^HQ = QQ^H = I$ (соответственно $Q^TQ = QQ^T = I$). Очевидно, что соотношение подобия для унитарно подобных матриц A и B имеет вид $B = Q^HAQ$, а для ортогонально подобных – $B = Q^TAQ$.

Замечание 3. Из теоремы 70.5 и свойства 8° §82 следует, что две комплексные (вещественные) квадратные матрицы одинакового порядка унитарно (соответственно ортогонально) подобны тогда и только тогда, когда они являются матрицами одного и того же оператора в унитарном (соответственно евклидовом) пространстве в ортонормированных базисах.

Теорема 96.6. *Квадратная комплексная матрица является нормальной тогда и только тогда, когда она унитарно подобна диагональной матрице.*

Доказательство может быть проведено по известной схеме (88.6), если использовать ортонормированные базисы e, f и учесть замечание 3. ■

Доказанная теорема представляет собой матричную формулировку теоремы 96.4.

$C \neq O$ (равенство $C = O$ означает, что канонический базис уже построен).

1. Если $a_{ii} \neq 0$, то выполним i-й шаг метода Лагранжа.

2. Пусть $a_{ii} = 0$.

а) Если среди диагональных элементов матрицы C существует элемент $a_{jj} \neq 0$, $j > i$, то перенумеруем переменные (т.е. векторы базиса): $x_i' = x_j$, $x_j' = x_i$ и $x_k' = x_k$ при $k \neq i, j$. Тогда в матрице A_{i-1} поменяются местами строки (столбцы) с номерами i и j, поэтому в позиции (i,i) окажется ненулевой элемент $a_{ii}' = a_{jj}$, с помощью которого выполним i-й шаг метода Лагранжа.

б) Пусть все диагональные элементы матрицы C равны нулю, тогда в ней существует внедиагональный элемент $a_{kj} \neq 0$, $k, j \geq i$, $k \neq i$. Это означает, что в квадратичной форме от переменных $x_i, \ldots, x_n$ отсутствуют квадраты $x_i^2, \ldots, x_n^2$, но содержится член вида $2a_{kj}x_kx_j$. Перейдем к новым координатам, положив $x_k = x_k' + x_j'$, $x_j = x_k' - x_j'$ и $x_s' = x_s$ при $s \neq k, j$. Тогда квадратичная форма будет иметь квадраты $x_k'^2$, $x_j'^2$ и мы окажемся в ситуации "а". Отметим, что все преобразования координат были невырожденными. ■

Теорема 101.8. *Если в матрице квадратичной формы $\mathcal{A}(x,x)$ ранга r первые r угловых миноров отличны от нуля: $\Delta_k \neq 0$, $k = \overline{1,r}$, то существует базис e, в котором матрица квадратичной формы имеет диагональный вид $A_e = \mathrm{diag}(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0)$, где*

$$\lambda_k = \Delta_k/\Delta_{k-1}, \ k = \overline{1,r}. \tag{101.13}$$

Доказательство. Для квадратичной формы $\mathcal{A}(x,x)$ в условиях теоремы выполнимы первые r шагов метода Лагранжа. После r-го шага матрица A_r квадратичной формы имеет вид

$$A_r = \begin{bmatrix} \lambda_1 & & & O \\ & \ddots & & \\ & & \lambda_r & \\ O & & & \boxed{C} \end{bmatrix},$$

где, согласно (101.12), $\lambda_k = \Delta_k/\Delta_{k-1}$, $k = \overline{1,r}$, а C – некоторая матрица. Так как $\mathrm{rg}\,A_r = r$ и $\lambda_k \neq 0$, $k = \overline{1,r}$, то $\mathrm{rg}\,C = 0$ и $C = O$. Следовательно, матрица A_r имеет искомый вид. ■

Соотношения (101.13) называются *формулами Якоби.*

§102. Квадратичные формы в вещественном пространстве

Ни канонический базис, ни канонический вид квадратичной формы не определены однозначно. Например, любая перестановка векторов канонического базиса приводит вновь к каноническому базису.

Что общего у различных канонических видов одной и той же квадратичной формы? В (101.9) мы отмечали, что число ненулевых канонических коэффициентов равно рангу квадратичной формы и не зависит от выбора канонического базиса. В вещественном пространстве можно говорить и о знаках канонических коэффициентов.

Пусть квадратичная форма $\mathcal{A}(x,x)$ ранга r в вещественном пространстве V приведена к каноническому виду (101.9). Число π положительных квадратов в (101.9) и число $\nu = r - \pi$ называются *положительным* и *отрицательным индексами инерции* квадратичной формы $\mathcal{A}(x,x)$, их разность $\sigma = \pi - \nu$ называется *сигнатурой* $\mathcal{A}(x,x)$. Корректность определения вытекает из следующей теоремы.

Теорема 102.1 (закон инерции). *Положительный и отрицательный индексы инерции вещественной квадратичной формы не зависят от выбора канонического базиса.*

Доказательство. Пусть e и f – канонические базисы квадратичной формы $\mathcal{A}(x,x)$ ранга r и для $x = \sum_{i=1}^n x_i e_i = \sum_{i=1}^n y_i f_i$

$$\begin{array}{l} \mathcal{A}(x,x) = a_1x_1^2 + \ldots + a_px_p^2 - a_{p+1}x_{p+1}^2 - \ldots - a_rx_r^2, \\ \mathcal{A}(x,x) = b_1y_1^2 + \ldots + b_qy_q^2 - b_{q+1}y_{q+1}^2 - \ldots - b_ry_r^2, \end{array} \tag{102.1}$$

где $a_i > 0$, $b_i > 0$, $i = \overline{1,r}$. Докажем, что $p \leq q$. Пусть $p > q$. Рассмотрим подпространства $L_1 = \mathcal{L}(e_1, \ldots, e_p)$, $L_2 = \mathcal{L}(f_{q+1}, \ldots, f_n)$. Согласно (65.2), $\dim L_1 \cap L_2 = p + (n - q) - \dim(L_1 + L_2)$. Так как $\dim(L_1 + L_2) \leq n$, $p > q$, то $\dim L_1 \cap L_2 > 0$. Следовательно, существует ненулевой вектор $x_0 \in L_1 \cap L_2$. Пусть $x_0 = \alpha_1e_1 + \ldots + \alpha_pe_p = \beta_{q+1}f_{q+1} + \ldots + \beta_nf_n$. Тогда согласно (102.1)

$$\mathcal{A}(x_0, x_0) = a_1\alpha_1^2 + \ldots + a_p\alpha_p^2 = -b_{q+1}\beta_{q+1}^2 - \ldots - b_r\beta_r^2. \tag{102.2}$$

Так как $x_0 \neq \theta$, то $a_1\alpha_1^2 + \ldots + a_p\alpha_p^2 > 0$, $-b_{q+1}\beta_{q+1}^2 - \ldots - b_r\beta_r^2 < 0$. Это противоречит (102.2), и, значит, $p \leq q$. Аналогично показывается, что $q \leq p$. Следовательно, $p = q$. ■

Обозначим символами $P(\Delta_0, \Delta_1, \ldots, \Delta_k)$ и $V(\Delta_0, \Delta_1, \ldots, \Delta_k)$ число совпадений и перемен знаков в последовательности $\Delta_0, \Delta_1, \ldots, \Delta_k$.

Теорема 102.2 (сигнатурное правило Якоби). *Пусть Δ_k – угловой минор k-го порядка матрицы квадратичной формы $\mathcal{A}(x,x)$ ранга r и $\Delta_k \neq 0$, $k = \overline{1,r}$. Тогда*

$$\pi = P(\Delta_0, \Delta_1, \ldots, \Delta_r), \quad \nu = V(\Delta_0, \Delta_1, \ldots, \Delta_r).$$

Утверждение теоремы вытекает из формул Якоби (101.13), так как $\lambda_i > 0$, если $\mathrm{sgn}\,\Delta_i = \mathrm{sgn}\,\Delta_{i-1}$, и $\lambda_i < 0$, если $\mathrm{sgn}\,\Delta_i \neq \mathrm{sgn}\,\Delta_{i-1}$. ■

Знакоопределенные квадратичные формы. Квадратичная форма $\mathcal{A}(x,x)$ называется *положительно (отрицательно) определенной*, если $\mathcal{A}(x,x) > 0$ ($\mathcal{A}(x,x) < 0$) $\forall x \neq \theta$. Такие формы называют *знакоопределенными* (или *знакопостоянными*).

или, в матричной формулировке: *квадратная комплексная матрица унитарна тогда и только тогда, когда она унитарно подобна диагональной матрице, у которой все диагональные элементы по модулю равны единице.*

Каноническая форма матрицы ортогонального оператора. Пусть $\mathcal{Q}$ – ортогональный оператор, действующий в евклидовом пространстве E. Согласно теореме 97.2 (п. "необходимость") и равенству (97.1) для его собственных значений λ и определителя (так как $\lambda \in \mathbb{R}$, $\det \mathcal{Q} \in \mathbb{R}$) имеем

$$\lambda = \pm 1, \quad \det \mathcal{Q} = \pm 1. \tag{97.2}$$

В любом ортонормированном базисе e оператор $\mathcal{Q}$ имеет ортогональную матрицу Q_e, следовательно,

$$Q_e^{-1} = Q_e^T. \tag{97.3}$$

Примеры. 1. В одномерном пространстве матрица Q_e ортогонального оператора $\mathcal{Q}$ имеет вид

$$Q_e = [\pm 1]. \tag{97.4}$$

Иными словами, либо $Q_e = I$, либо $Q_e = -I$.

2. В двумерном пространстве матрица Q_e ортогонального оператора $\mathcal{Q}$ имеет вид $Q_e = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, причем $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}^{-1} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$ в силу (97.3), и, с учетом (97.2), (5.3),

$$\pm\begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}. \tag{97.5}$$

а) Если $\det \mathcal{Q} = 1$, то (97.5) означает, что $\begin{bmatrix} \delta & -\beta \\ -\gamma & \alpha \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$, откуда следует, что $Q_e = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$, $\alpha^2 + \beta^2 = 1$. Положив $\alpha = \cos\varphi$, $\beta = -\sin\varphi$, получим

$$Q_e = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}. \tag{97.6}$$

б) Если $\det \mathcal{Q} = -1$, то (97.5) означает, что $\begin{bmatrix} -\delta & \beta \\ \gamma & -\alpha \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$, откуда следует, что $Q_e = \begin{bmatrix} \alpha & \beta \\ \beta & -\alpha \end{bmatrix}$, $\alpha^2 + \beta^2 = 1$; при этом $f(\lambda) = \det(Q_e - \lambda I) = \lambda^2 - 1$. Следовательно, оператор $\mathcal{Q}$ в двумерном пространстве имеет два различных собственных значения: $\lambda_1 = 1$,

$\lambda_2 = -1$, и для него существует ортонормированный базис f, в котором его матрица Q_f имеет вид

$$Q_f = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{97.7}$$

Теорема 97.4. *Для любого ортогонального оператора $\mathcal{Q}$ в евклидовом пространстве существует ортонормированный базис e, в котором его матрица имеет квазидиагональную форму с клетками вида (97.4) и (97.6) на главной диагонали.*

Доказательство. Воспользуемся индукцией по размерности n пространства V. Для $n = 1, 2$ утверждение вытекает из (97.4), (97.6), (97.7). Пусть $n \geq 3$ и теорема верна для ортогональных операторов в пространствах размерности $k \leq (n-1)$. Докажем теорему для размерности n. Согласно теореме 93.2 для оператора $\mathcal{Q}$ существует одномерное либо двумерное инвариантное подпространство L. Если $\dim L = 1$, то $\mathcal{Q}|L$ – ортогональный оператор (согласно следствию) в одномерном пространстве и в ортонормированном базисе e_1 пространства L его матрица имеет вид (97.4). Если же $\dim L = 2$, то $\mathcal{Q}|L$ – ортогональный оператор в двумерном пространстве и в ортонормированном базисе e_1, e_2 пространства L его матрица имеет вид (97.6) или (97.7). Согласно теореме 97.3 ортогональное дополнение $L^\perp$ инвариантно относительно оператора $\mathcal{Q}$. По индуктивному предположению в пространстве $L^\perp$ существует ортонормированный базис $e_2, \ldots, e_n$ (или $e_3, \ldots, e_n$), в котором матрица оператора $\mathcal{Q}|L^\perp$ имеет требуемый вид. Тогда в базисе $e_1, \ldots, e_n$ матрица оператора $\mathcal{Q}$ также будет иметь требуемый вид. ■

После очевидной перестановки векторов построенного базиса матрица ортогонального оператора $\mathcal{Q}$ будет иметь вид

$$Q_e = \begin{bmatrix} 1 & & & & & & & & O \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \boxed{\begin{matrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{matrix}} & & \\ & & & & & & & \ddots & \\ O & & & & & & & & \boxed{\begin{matrix} \cos\varphi_k & -\sin\varphi_k \\ \sin\varphi_k & \cos\varphi_k \end{matrix}} \end{bmatrix}. \tag{97.8}$$

Матрица (97.8) называется *канонической формой матрицы ортогонального оператора.*

Простым вращением называется оператор в евклидовом пространстве, который в некотором ортонормированном базисе имеет матрицу

Квадратичная форма, для которой существуют векторы x и y такие, что $\mathcal{A}(x,x) > 0$, $\mathcal{A}(y,y) < 0$, называются *знакопеременными*.

Пример. Примером положительно определенной квадратичной формы в вещественном пространстве может служить скалярный квадрат в евклидовом пространстве, т.е. отображение $\mathcal{A}: E \to \mathbb{R}$, определенное равенством $\mathcal{A}(x) = (x,x)$, $\forall x \in E$. Это следует из того, что скалярное произведение (x,y) является симметричной билинейной формой (в силу аксиом 1–3 из (68.1)), а скалярный квадрат – положительно определенной квадратичной формой (в силу аксиомы 4).

Теорема 102.3. *Квадратичная форма $\mathcal{A}(x,x)$ положительно (отрицательно) определена тогда и только тогда, когда ее положительный (соответственно отрицательный) индекс инерции совпадает с размерностью пространства.*

Доказательство. Докажем сначала вариант теоремы, относящийся к положительно определенной форме. Пусть e – канонический базис квадратичной формы $\mathcal{A}(x,x)$ и $\lambda_1,\dots,\lambda_n$ – ее канонические коэффициенты. Необходимость. Если $\mathcal{A}(x,x) > 0$, $\forall x \neq \theta$, то, в частности, $\mathcal{A}(e_i,e_i) > 0$, $i = \overline{1,n}$. Но $\mathcal{A}(e_i,e_i) = \lambda_i$, следовательно, все канонические коэффициенты положительны и $\pi = n$.

Достаточность. Если $\lambda_i > 0$, $i = \overline{1,n}$, то согласно (101.9)

$$\mathcal{A}(x,x) = \sum_{i=1}^{n} \lambda_i x_i^2 > 0, \forall x = \sum_{i=1}^{n} x_i e_i \neq \theta.$$

Для доказательства второго утверждения достаточно рассмотреть квадратичную форму $-\mathcal{A}(x,x)$:

$$\mathcal{A}(x,x) < 0 \iff -\mathcal{A}(x,x) > 0 \iff \nu = \dim V. \ \blacksquare$$

Следствие. Определитель матрицы положительно определенной квадратичной формы положителен, так как если e – канонический базис, то в произвольном базисе f, согласно (101.6), $|A_f| = \lambda_1 \cdot \ldots \cdot \lambda_n \cdot |Q|^2 > 0$.

Теорема 102.4 (критерий Сильвестра). *Квадратичная форма $\mathcal{A}(x,x)$ положительно (отрицательно) определена тогда и только тогда, когда угловые миноры Δ_k, $k = \overline{1,n}$, ее матрицы в произвольном базисе положительны (соответственно чередуют знаки, начиная с отрицательного):*

$$\Delta_k > 0,\ k = \overline{1,n} \quad (\Delta_k \Delta_{k-1} < 0,\ k = \overline{0,n}).$$

Доказательство. Докажем первое утверждение.

Необходимость. Пусть $\mathcal{A}(x,x)$ – положительно определенная квадратичная форма, A_e – ее матрица в произвольном базисе $e = (e_1,\dots,e_n)$. Рассмотрим подпространства $L_k = \mathcal{L}(e_1,\dots,e_k)$, $k = \overline{1,n}$, и квадратичные формы $\mathcal{A}(x,x)$ на L_k. Очевидно, $\mathcal{A}(x,x) > 0, \forall x \in L_k$, $x \neq \theta$, и, значит, ее матрица A_k в базисе $e_1,\dots,e_k$ пространства L_k имеет положительный определитель $|A_k|$ (следствие 1). Но $|A_k| = \Delta_k$,

$k = \overline{1,n}$, следовательно, $\Delta_k > 0$, $k = \overline{1,n}$.

Достаточность вытекает из теорем 102.2, 102.3.

Для доказательства второго утверждения рассмотрим квадратичную форму $-\mathcal{A}(x,x)$ с угловыми минорами δ_k. Отрицательная определенность $\mathcal{A}(x,x)$ равносильна положительности формы $-\mathcal{A}(x,x)$, т.е. условию $\delta_k > 0$, $k = \overline{1,n}$, или $(-1)^k\Delta_k > 0$, $k = \overline{1,n}$. Так как $\Delta_0 = 1 > 0$, то это условие означает, что $\Delta_k\Delta_{k-1} < 0$, $k = \overline{1,n}$. $\blacksquare$

Как следует из примера положительно определенной квадратичной формы, скалярное произведение в евклидовом пространстве является билинейной формой, полярной к положительно определенной квадратичной форме. Оказывается, что такими билинейными формами исчерпываются все скалярные произведения в вещественном пространстве.

Теорема 102.5. *Пусть V – вещественное линейное пространство. Отображение $\mathcal{A}: V \times V \to \mathbb{R}$ есть скалярное произведение в пространстве V тогда и только тогда, когда оно является билинейной формой, полярной к положительно определенной квадратичной форме.*

Доказательство. Необходимость рассмотрена в примере. Достаточность. Пусть $\mathcal{A}(x,y)$ – билинейная форма, полярная к положительно определенной квадратичной форме. Тогда отображение $\mathcal{A}: V \times V \to \mathbb{R}$, определенное правилом

$$(x,y) = \mathcal{A}(x,y), \quad \forall x,y \in V, \tag{102.3}$$

является скалярным произведением, так как оно удовлетворяет всем аксиомам (68.1) скалярного произведения. $\blacksquare$

Замечание. Матрица билинейной формы, задающей скалярное произведение, совпадает с матрицей Грама базисных векторов: $A_e = G(e_1,\dots,e_k)$.

§103. Квадратичные формы в комплексном пространстве

В комплексном пространстве полным аналогом вещественных квадратичных форм являются так называемые эрмитовы квадратичные формы.

Полуторалинейные формы. Пусть V – комплексное линейное пространство. Отображение $\mathcal{A}: V \times V \to \mathbb{C}$ называется *полуторалинейной формой*, если для любых $x,y,z \in V$, $\alpha \in \mathbb{C}$:

1) $\mathcal{A}(x+y,z) = \mathcal{A}(x,z) + \mathcal{A}(y,z)$;
2) $\mathcal{A}(\alpha x,y) = \alpha\mathcal{A}(x,y)$;
3) $\mathcal{A}(x,y+z) = \mathcal{A}(x,y) + \mathcal{A}(x,z)$;
4) $\mathcal{A}(x,\alpha y) = \overline{\alpha}\mathcal{A}(x,y)$.

Полуторалинейную форму называют *эрмитовой*, если

$$\mathcal{A}(y,x) = \overline{\mathcal{A}(x,y)}, \forall x,y \in V.$$

являются плоскостями симметрии, а начало координат – центром симметрии двуполостного гиперболоида.

Плоскости $z = h$ пересекают этот гиперболоид:

а) при $|h| > c$ по эллипсам, размеры которых неограниченно возрастают при $|h| \to \infty$;

б) при $|h| = c$ в единственной точке.

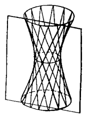


Рис. 3

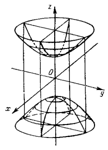


Рис. 4

В слое между плоскостями $z = c$ и $z = -c$ нет ни одной точки гиперболоида. Таким образом, двуполостный гиперболоид состоит из двух симметричных полостей, расположенных в полупространствах $z \geq c$ и $z \leq -c$.

Плоскости $x = h$ и $y = h$ пересекают гиперболоид по гиперболам.

Конус (рис. 5). Поверхность, определяемая уравнением 6, называется *конусом*. Как следует из канонического уравнения, координатные плоскости являются плоскостями симметрии, а начало координат – центром симметрии конуса.

Плоскости $z = h$, $h \neq 0$, пересекают конус по эллипсам, размеры которых неограниченно возрастают при $|h| \to \infty$. Прямая, проходящая через любую точку $M_0(x_0,y_0,z_0)$ конуса и начало координат, является прямолинейной образующей конуса, так как любая точка этой прямой имеет координаты (x_0t, y_0t, z_0t).

Плоскости $y = h$ и $x = h$, где $h \neq 0$, пересекают конус по гиперболам (рис. 6) с полуосями $c|h|/b$, $a|h|/b$ и $c|h|/a$, $b|h|/a$ соответственно; плоскости $x = 0$ и $y = 0$ – по парам пересекающихся прямых.

Плоскими сечениями конуса являются и параболы (рис. 7). Так, параболой будет сечение конуса плоскостью $z = h + cx/a$, где $h \neq 0$, ибо числа x, y являются на этой плоскости аффинными (не прямоугольными) координатами, а уравнение линии, высекаемой на ней

конусом, имеет в этих координатах вид

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{(h + cx/a)^2}{c^2} = 0.$$

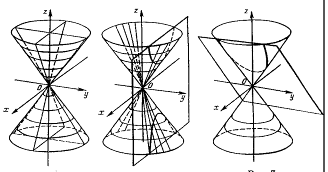


Рис. 5 Рис. 6 Рис. 7

Традиционными преобразованиями это уравнение приводится к уравнению параболы

$$y = 2\frac{hb^2}{ac}\left(x + \frac{ha}{2c}\right).$$

Таким образом, и эллипс, и гипербола, и парабола являются плоскими сечениями конуса. На этом основании эти линии обычно называют *коническими сечениями*.

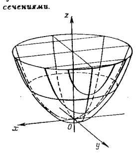


Рис. 8

Доказательство. Существование. Пусть $e = (e_1,\dots,e_n)$ – ортонормированный базис пространства E и A_e – матрица квадратичной формы $\mathcal{A}(x,x)$ в этом базисе. В силу симметричности матрицы A_e и ортонормированности базиса e существует симметрический оператор $\mathcal{H} \in \mathcal{L}(E,E)$, который в базисе e имеет матрицу A_e, так что $H_e = A_e$. Тогда для любого вектора $x = \sum_{i=1}^n x_i e_i$ координаты вектора $\mathcal{H}x$ находятся согласно (77.3): $(\mathcal{H}x)_e = H_e x_e = A_e x_e$. Отсюда в силу (70.7) следует равенство $(\mathcal{H}x,x) = x_e^T A_e x_e$, которое с учетом (101.8) доказывает (104.1).

Единственность. Если $\mathcal{H}_1, \mathcal{H}_2$ – симметрические операторы, удовлетворяющие (104.1), то $(\mathcal{H}_1x,x) = (\mathcal{H}_2x,x)$, $\forall x \in E$, или

$$((\mathcal{H}_1 - \mathcal{H}_2)x, x) = 0,\ \forall x \in E. \tag{104.2}$$

Отсюда следует [2], что $\mathcal{H}_1 = \mathcal{H}_2$. $\blacksquare$

Теорема 104.2. *Для любой квадратичной формы в евклидовом пространстве E существует ортонормированный базис, в котором она имеет канонический вид.*

Доказательство. Пусть $\mathcal{H}$ – симметрический оператор, для которого $(\mathcal{H}x) = \mathcal{A}(x,x)$, $\forall x \in E$. Если $e_1,\dots,e_n$ – ортонормированный базис из собственных векторов $\mathcal{H}$, то для любого вектора $x = \sum_{i=1}^n x_i e_i$, согласно (104.1) и (70.7), имеем

$$\mathcal{A}(x,x) = \left(\sum_{i=1}^{n} x_i\lambda_i e_i, \sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} \lambda_i x_i^2, \tag{104.3}$$

где $\lambda_1,\dots,\lambda_n$ – собственные значения оператора $\mathcal{H}$, отвечающие собственным векторам $e_1,\dots,e_n$. $\blacksquare$

Операция построения ортонормированного базиса, в котором квадратичная форма имеет канонический вид, называется *приведением квадратичной формы к главным осям*. Из теоремы 104.2 следует, что *любая квадратичная форма приводится к главным осям*.

Замечание 1. Доказывая теорему 104.2, мы попутно показали, что канонический базис квадратичной формы $\mathcal{A}(x,x)$ совпадает с ортонормированным базисом из собственных векторов соответствующего симметрического оператора $\mathcal{H}$, а канонические коэффициенты – с отвечающими им собственными значениями. Собственные значения оператора $\mathcal{H}$ являются корнями уравнения $|A_e - \lambda I| = 0$, которые, вообще говоря, уже не зависят от оператора $\mathcal{H}$ и инвариантно связаны только с самой квадратичной формой.

Таким образом, при приведении квадратичной формы к главным осям *канонические коэффициенты определены однозначно*. Это позволяет находить канонический вид квадратичной формы, минуя вычисление канонического базиса.

Что же касается канонического базиса, то он определен с той же степенью произвола, с какой определена полная ортонормированная система из собственных векторов оператора.

[2] Оператор $\mathcal{H}_1 - \mathcal{H}_2$ симметричен, для него существует ортонормированный базис из собственных векторов, а все собственные значения в силу (104.2) равны нулю.

Теорема 104.3 (о паре квадратичных форм). *Для любой пары квадратичных форм $\mathcal{A}(x,x)$ и $\mathcal{B}(x,x)$ в вещественном пространстве V, одна из которых положительно определена, существует общий базис, в котором обе квадратичные формы имеют канонический вид.*

Доказательство. Пусть $\mathcal{B}(x,x) > 0$ и $\mathcal{B}(x,y)$ – билинейная форма, полярная к квадратичной форме $\mathcal{B}(x,x)$. В соответствии с теоремой 102.5 введем скалярное произведение (102.3). Тогда пространство V становится евклидовым и в нем согласно теореме 104.2 существует ортонормированный базис $e_1,\dots,e_n$, в котором квадратичная форма $\mathcal{A}(x,x)$ имеет канонический вид (104.3). При этом в силу (70.7) для любого вектора $x = \sum_{i=1}^n x_i e_i$

$$\mathcal{B}(x,x) = (x,x) = \sum_{i=1}^{n} x_i^2. \ \blacksquare \tag{104.4}$$

Замечание 2. Один из способов поиска общего базиса $e_1,\dots,e_n$, в котором формы $\mathcal{A}(x,x)$ и $\mathcal{B}(x,x)$ имеют канонический вид (104.3) и (104.4), состоит в следующем.

Пусть A и B – матрицы квадратичных форм $\mathcal{A}(x,x)$ и $\mathcal{B}(x,x)$ в некотором базисе f и пусть $e = fQ$. Тогда $Q^TAQ = \Lambda$, где $\Lambda = \mathrm{diag}(\lambda_1,\dots,\lambda_n)$, и $Q^TBQ = I$. Следовательно, $A = (Q^T)^{-1}\Lambda Q^{-1}$, $B = (Q^T)^{-1}Q^{-1}$, $B^{-1} = QQ^T$, $B^{-1}A = Q\Lambda Q^{-1}$ и $B^{-1}AQ = Q\Lambda$. Последнее равенство означает, что столбцы матрицы Q (т.е. координаты векторов искомого канонического базиса e в исходном базисе f) являются собственными векторами матрицы $B^{-1}A$, отвечающими собственным значениям $\lambda_1,\dots,\lambda_n$. Таким образом, канонические коэффициенты $\lambda_1,\dots,\lambda_n$ являются корнями уравнения $|B^{-1}A - \lambda I| = 0$ или, в равносильной записи, уравнения

$$|A - \lambda B| = 0, \tag{104.5}$$

а векторы канонического базиса – нетривиальными решениями однородной системы уравнений

$$Ax = \lambda Bx. \tag{104.6}$$

Теорема 104.3 обеспечивает существование полного набора вещественных корней уравнения (104.5) и для каждого кратного корня – наличие соответствующего числа линейно независимых решений системы (104.6). Вообще говоря, описанный алгоритм (как и доказательство теоремы 104.3) решает поставленную задачу в усиленной форме: положительно определенная форма $\mathcal{B}(x,x)$ приводится к каноническому виду с коэффициентами, равными 1. В действительности это не требуется: без существенных изменений приведенный здесь алгоритм применим и для случая, когда $Q^TBQ = \mathrm{diag}(\mu_1,\dots,\mu_n)$. Таким образом, коэффициенты преобразуемых форм не определяются однозначно. Однако, как легко показать, *отношения λ_i/μ_i канонических коэффициентов не зависят от способа одновременного приведения форм $\mathcal{A}(x,x)$ и $\mathcal{B}(x,x)$ к каноническому виду.*

В вычислительной алгебре задачу (104.6) называют *обобщенной проблемой собственных значений.*

Замечание 3. Теоремы 104.1–104.3 с очевидными изменениями (см. §103) переносятся и на эрмитовы квадратичные формы в комплексном пространстве.

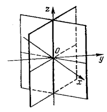


Рис. 14

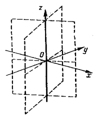


Рис. 15

Поверхности, определяемые уравнениями 15–17, называют соответственно *парой параллельных плоскостей* (рис. 16), *парой мнимых параллельных плоскостей* и *парой совпадающих плоскостей* (рис. 17).

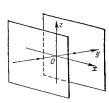


Рис. 16

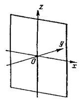


Рис. 17

Глава XIX. Линейные нормированные пространства

§107. Норма вектора

Пусть V – линейное пространство, вещественное или комплексное. *Нормой* в линейном пространстве V называется отображение $\|\cdot\|: V \to \mathbb{R}$, ставящее в соответствие каждому вектору $x \in V$ число $\|x\| \in \mathbb{R}$ и удовлетворяющее аксиомам: $\forall x,y \in V$, $\alpha \in \mathbb{R}(\mathbb{C})$

1) $\|x\| \geq 0$,
   $\|x\| = 0$ тогда и только тогда, когда $x = \theta$;
2) $\|\alpha x\| = |\alpha| \cdot \|x\|$;
3) $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника).

Линейное пространство V с заданной на нем нормой $\|\cdot\|$ называется *линейным нормированным пространством*. Число $\|x\|$ называется *нормой вектора x*.

Примеры. 1. В евклидовом (унитарном) пространстве норма может быть введена как длина, т.е. $\|x\| = |x|$. Эта норма называется *евклидовой нормой* и обозначается символом $\|x\|_E$. Итак,

$$\|x\|_E = \sqrt{(x,x)}. \tag{107.1}$$

Справедливость аксиом нормы вытекает из свойств длины (§ 69).

2. В арифметических пространствах $\mathbb{R}^n$, $\mathbb{C}^n$ обычно рассматривают следующие нормы вектора $x = (x_1,\dots,x_n)$:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|; \quad \|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}; \quad \|x\|_\infty = \max_{1\leq i\leq n} |x_i|,$$

которые можно записать единообразно в виде

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \text{ где } p = 1,2,\infty$$

(здесь $\|x\|_\infty = \lim\limits_{p\to\infty} \|x\|_p$). Отметим, что $\|\cdot\|_2$ совпадает с $\|\cdot\|_E$ для естественного скалярного произведения (70.7).

3. Аналогичные нормы рассматриваются в произвольном конечномерном линейном пространстве V (вещественном или комплексном): если $e_1,\dots,e_n$ – базис V и $x = \sum_{i=1}^n x_i e_i$, то

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad \text{где } p = 1,2,\infty. \tag{107.2}$$

Матрица S ортогональна (так как ее строки ортонормированы), поэтому преобразование базиса с помощью матрицы перехода S приводит к новому ортонормированному базису. При этом преобразование координат осуществляется по закону $x_k''' = x_k''$ при $k = \overline{1,r}$, $x_r''' = \alpha^{-1}\sum_{k=r+1}^n b_k' x_k''$, $x_k''' = \sum_{i=r+1}^n s_{ki}x_i''$ при $k = \overline{r+2,n}$, и приводит уравнение (105.5) к виду $\lambda_1 x_1'''^2 + \ldots + \lambda_r x_r'''^2 + 2\alpha x_{r+1}''' + c' = 0$. В последнем уравнении можно освободиться от свободного члена c', если выполнить параллельный перенос на вектор, у которого все координаты, кроме $(r+1)$-й (равной $c'/(2\alpha)$), равны нулю.

Итак, с помощью параллельного переноса и перехода к новому ортонормированному базису общее уравнение (105.2) приводится к одному из двух типов уравнений:

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + a_0 = 0, \quad \lambda_1 \ldots \lambda_r \neq 0, \tag{105.6}$$

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + b_0 x_{r+1} = 0, \quad \lambda_1 \ldots \lambda_r b_0 \neq 0, \tag{105.7}$$

которые принято называть *приведенными уравнениями гиперповерхности*.

Инварианты гиперповерхности. Величины и функции, определяемые коэффициентами общего уравнения (105.2), которые не изменяются при преобразованиях координат, называются *инвариантами гиперповерхности относительно этих преобразований*.

Аппарат инвариантов (как и в случае линий второго порядка на плоскости) позволяет по общему уравнению однозначно найти тип приведенного уравнения и все его коэффициенты.

Рассмотрим матрицу

$$B = \left[\begin{array}{c|c} A & b \\ \hline b^T & c \end{array}\right], \tag{105.8}$$

полученную окаймлением матрицы A квадратичной формы столбцом и строкой из коэффициентов линейной формы и свободным членом c.

Теорема 105.1. *Характеристические многочлены матриц A и B являются инвариантами гиперповерхности относительно ортогонального преобразования координат.*

Доказательство. Утверждение теоремы следует из того, что соотношения $A_f = Q^T A_e Q$ и $B_f = \widetilde{Q}^T B_e \widetilde{Q}$ (см. доказательство теоремы 58.2) в случае ортогональности Q означают подобие матриц A_e и A_f, B_e и B_f, а следовательно, и совпадение характеристических многочленов (теорема 86.2). $\blacksquare$

Следствие. Инвариантами гиперповерхности относительно ортогонального преобразования координат являются:

а) канонические коэффициенты $\lambda_1,\dots,\lambda_n$ (как корни характеристического многочлена A);

б) величины $\mathrm{tr}\,A$, $\det A$ (как коэффициенты характеристического многочлена);

в) числа $\mathrm{rg}\,A$, $\mathrm{rg}\,B$.

Теорема 105.2. *Величины* $\mathrm{tr}\,A$, $\det A$, $\det B$, $\mathrm{rg}\,A$, $\mathrm{rg}\,B$ *являются инвариантами гиперповерхности относительно параллельного переноса.*

Доказательство. Инвариантность величин $\mathrm{tr}\,A$, $\det A$, $\mathrm{rg}\,A$ следует из того, что при параллельном переносе матрица A не изменяется, а инвариантность $\det B$, $\mathrm{rg}\,B$ – из того, что матрица B при параллельном переносе подвергается элементарным преобразованиям (см. доказательство теоремы 58.2). $\blacksquare$

Классификация гиперповерхностей. Классификация осуществляется на основе приведенных уравнений (105.6) и (105.7).

1. Пусть $r = n$. Тогда уравнение (105.2) приводится только к уравнению типа (105.6), которое в этом случае имеет вид

$$\lambda_1 x_1^2 + \ldots + \lambda_n x_n^2 + a_0 = 0, \tag{105.9}$$

при этом $\det A = \lambda_1 \ldots \lambda_n$, $\det B = \lambda_1 \ldots \lambda_n a_0$, $a_0 = \det B/\det A$. Все коэффициенты в (105.9) определены однозначно вследствие инвариантности чисел $\lambda_1,\dots,\lambda_n$, $\det A$, $\det B$. В силу закона инерции (теорема 102.1) число положительных и число отрицательных коэффициентов $\lambda_1,\dots,\lambda_n$ определены однозначно, так что уравнение (105.9) однозначно преобразуется в уравнение

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_k^2}{a_k^2} - \frac{x_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{x_n^2}{a_n^2} = 1, \tag{105.10}$$

если $\det B \neq 0$; а в случае, когда $\det B = 0$, – в уравнение

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_k^2}{a_k^2} - \frac{x_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{x_n^2}{a_n^2} = 0. \tag{105.11}$$

Гиперповерхность, определяемая уравнением (105.10), при $k = n$ называется *эллипсоидом*, при $k = 0$ – *мнимым эллипсоидом*, при $0 < k < n$ – *гиперболоидом*. Гиперповерхность, определяемая уравнением (105.11), называется *конусом*.

2. Пусть $r = n - 1$. Тогда уравнение (105.2) приводится к уравнению типа (105.6), если $\det B = 0$, и к уравнению (105.7), если $\det B \neq 0$. Эти уравнения однозначно преобразуются в первом случае в уравнение

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_k^2}{a_k^2} - \frac{x_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{x_{n-1}^2}{a_{n-1}^2} = c, \tag{105.12}$$

где $c = 1$, если $\mathrm{rg}\,B = n$, и $c = 0$, если $\mathrm{rg}\,B = n - 1$; а во втором случае в уравнение

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_k^2}{a_k^2} - \frac{x_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{x_{n-1}^2}{a_{n-1}^2} = 2px_n, \quad p > 0. \tag{105.13}$$

нормы эквивалентны.

Доказательство. Выберем в пространстве V произвольный базис $e = (e_1,\dots,e_n)$ и введем скалярное произведение согласно (70.7). Тогда V – евклидово (унитарное) пространство и в нем можно рассматривать евклидову норму (107.1). Так как в этом скалярном произведении e – ортонормированный базис, то $\|x\|_E = |x| = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$. Покажем, что любая норма $\|\cdot\|$ в линейном пространстве V эквивалентна евклидовой норме $\|\cdot\|_E$. Тем самым теорема будет доказана.

1. Докажем сначала существование числа $c_1 > 0$ такого, что для любого вектора $x \in V$

$$\|x\| \leq c_1\|x\|_E. \tag{107.3}$$

Пусть $x \in V$ и $x = \sum_{i=1}^n x_i e_i$. Тогда $\|x\| \leq \sum_{i=1}^n |x_i|\,\|e_i\|$ и в силу неравенства Коши–Буняковского

$$\|x\| \leq \left(\sum_{i=1}^{n} |x_i|^2 \sum_{i=1}^{n} \|e_i\|^2\right)^{1/2} = c_1\|x\|_E, \text{ где } c_1 = \left(\sum_{i=1}^{n} \|e_i\|^2\right)^{1/2} > 0.$$

2. Докажем теперь существование числа $c_2 > 0$ такого, что для любого вектора $x \in V$

$$\|x\|_E \leq c_2\|x\|. \tag{107.4}$$

а) Рассмотрим единичную сферу по норме $\|\cdot\|$: $S = \{x \in V \mid \|x\| = 1\}$ и покажем, что множество координат всех векторов сферы S ограничено. Пусть это не так. Тогда для любого $m \in \mathbb{N}$ существует вектор $x^{(m)}$ такой, что $\|x^{(m)}\| = 1$ и $\left|x_k^{(m)}\right| > m$ для некоторого $k \in \{1,\dots,n\}$. Следовательно,

$$\|x^{(m)}\|_E = \left(\sum_{i=1}^{n} \left|x_i^{(m)}\right|^2\right)^{1/2} \geq \left|x_k^{(m)}\right| > m. \tag{107.5}$$

Рассмотрим последовательность векторов

$$y^{(m)} = x^{(m)} \big/ \|x^{(m)}\|_E\ . \tag{107.6}$$

Очевидно, что $\|y^{(m)}\|_E = 1$ и, значит, $y^{(m)} \in S_E(1)$. В силу утверждения 3 из последовательности $\{y^{(m)}\}$ можно выделить подпоследовательность $\{y^{(m_j)}\}$, сходящуюся по $\|\cdot\|_E$ к вектору $a \in S_E(1)$, т.е.

$$\|y^{(m_j)} - a\|_E \to 0, \tag{107.7}$$

$$\|a\|_E = 1. \tag{107.8}$$

Из (107.7) с учетом (107.3) следует, что

$$\|y^{(m_j)} - a\| \to 0. \tag{107.9}$$

Таким образом, последовательность $\{y^{(m_j)}\}$ сходится к вектору a по норме $\|\cdot\|$.

С другой стороны, в силу (107.6), (107.5) с учетом $\|x^{(m_j)}\| = 1$

$$\|y^{(m_j)}\| = \|x^{(m_j)}\|/\|x^{(m_j)}\|_E < 1/m_j.$$

Следовательно, $\|y^{(m_j)}\| \to 0$. Отсюда и из (107.9) в силу единственности предела вытекает, что $a = \theta$, что противоречит (107.8).

Таким образом, существует число M такое, что для всех $x = \sum_{i=1}^n x_i e_i \in S$ выполняется неравенство

$$|x_i| \leq M, \quad i = \overline{1,n}. \tag{107.10}$$

б) Пусть $x = \sum_{i=1}^n x_i e_i$ – любой вектор пространства V, тогда вектор $x/\|x\| \in S$ и для его координат имеет место неравенство (107.10). Значит, $|x_i|/\|x\| \leq M$, т.е. $|x_i| \leq M\|x\|$, $i = \overline{1,n}$, и $\|x\|_E = \sqrt{|x_1|^2 + \ldots + |x_n|^2} \leq M\sqrt{n}\,\|x\| = c_2\|x\|$, где $c_2 = M\sqrt{n} > 0$. $\blacksquare$

Следствие. В конечномерном пространстве из сходимости по одной норме следует сходимость по любой другой норме, так как

$$\|x^{(k)} - a\|_1 \leq c_1\|x^{(k)} - a\|_2.$$

§108. Линейные операторы в нормированных пространствах

Пусть V и W – линейные нормированные пространства (оба вещественные или оба комплексные). Линейное пространство $\mathcal{L}(V,W)$, как и любое линейное пространство, можно сделать нормированным, введя на нем норму оператора. Однако для нормы оператора нет той свободы выбора, которая имеет место для векторных норм, так как оператор тесно связан с векторами и, следовательно, его норма тоже должна быть связана с векторными нормами пространств V и W.

Согласованные нормы. Норма оператора $\mathcal{A} \in \mathcal{L}(V,W)$ называется *согласованной* с векторными нормами $\|\cdot\|_V$, $\|\cdot\|_W$ пространств V и W, если для любого вектора $x \in V$

$$\|\mathcal{A}x\|_W \leq \|\mathcal{A}\| \cdot \|x\|_V.$$

Замечание 1. В дальнейшем в символе $\|\cdot\|_V$ индекс V будет опускаться, если из контекста ясно, о каком пространстве идет речь.

Теорема 108.1. *Собственное значение линейного оператора $\mathcal{A} \in \mathcal{L}(V,V)$ не превосходит по абсолютной величине любую его согласованную норму.*

Доказательство. Если $\mathcal{A}x = \lambda x$, то для любой согласованной нормы оператора имеем $\|\mathcal{A}x\| = |\lambda| \cdot \|x\|$ и $\|\mathcal{A}x\| \leq \|\mathcal{A}\| \cdot \|x\|$, откуда следует, что $|\lambda| \leq \|\mathcal{A}\|$. $\blacksquare$

(106.4): $I_3 = 0$, $K_4 \neq 0$;
(106.5): $I_3 = 0$, $K_4 = 0$, $I_2 \neq 0$;
(106.6): $I_3 = 0$, $K_4 = 0$, $I_2 = 0$, $K_3 \neq 0$;
(106.7): $I_3 = 0$, $K_4 = 0$, $I_2 = 0$, $K_3 = 0$, $I_1 \neq 0$.

Эти условия необходимы и достаточны для каждого из перечисленных уравнений, так как взаимно исключают друг друга. При этом все коэффициенты уравнений определены однозначно, ибо $\lambda_1, \lambda_2, \lambda_3$ – корни характеристического уравнения: $-\lambda^3 + I_1\lambda^2 - I_2\lambda + I_3 = 0$; $a_0 = \dfrac{K_4}{I_3}$; $b_0 = \left(-\dfrac{K_4}{I_2}\right)^{1/2}$; $c_0 = \dfrac{K_3}{I_2}$; $p_0 = \left(-\dfrac{K_3}{I_1}\right)^{1/2}$; $q_0 = \dfrac{K_2}{I_1}$.

Канонические уравнения.

Теорема 106.1. *Для любой алгебраической поверхности второго порядка существует прямоугольная декартова система координат $Oxyz$, в которой уравнение этой поверхности имеет один из следующих видов:*

1) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$, *эллипсоид;*

2) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = -1$, *мнимый эллипсоид;*

3) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 0$, *вырожденный эллипсоид;*

4) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$, *однополостный гиперболоид;*

5) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = -1$, *двуполостный гиперболоид;*

6) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0$, *конус;*

7) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 2z$, *эллиптический параболоид;*

8) $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 2z$, *гиперболический параболоид;*

9) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$, *эллиптический цилиндр;*

10) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = -1$, *мнимый эллиптический цилиндр;*

11) $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$, *гиперболический цилиндр;*

12) $y^2 = 2px\ (p > 0)$, *параболический цилиндр;*

13) $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 0$, *пара пересекающихся плоскостей;*

14) $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 0$, *пара мнимых пересекающихся плоскостей;*

15) $y^2 = a^2\ (a > 0)$, *пара параллельных плоскостей;*

16) $y^2 = -a^2\ (a > 0)$, *пара мнимых параллельных плоскостей;*

17) $y^2 = 0$, *пара совпадающих плоскостей.*

Доказательство. Будем исходить из уравнений (106.3)–(106.7). Рассмотрим уравнение (106.3). В зависимости от сочетаний знаков λ_1, λ_2, λ_3, a_0 оно может быть записано по-разному.

1. Пусть знаки чисел λ_1, λ_2, λ_3 совпадают. При этом:

а) если $a_0\lambda_1 < 0$, то уравнение (106.3) может быть записано как уравнение 1, где $a^2 = -a_0/\lambda_1$, $b^2 = -a_0/\lambda_2$, $c^2 = -a_0/\lambda_3$;

б) если $a_0\lambda_1 > 0$ – как уравнение 2;

в) если $a_0 = 0$ – как уравнение 3.

2. Пусть знаки чисел λ_1, λ_2, λ_3 не совпадают. Не ограничивая общности рассуждений, можно считать, что $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 < 0$. При этом уравнение (106.3) может быть записано как уравнение 4 (если $a_0\lambda_3 > 0$), как уравнение 5 (если $a_0\lambda_3 < 0$) и как уравнение 6 (если $a_0 = 0$).

Аналогично уравнение (106.4) в зависимости от знаков λ_1, λ_2 может быть записано (после возможных перенумераций неизвестных и изменения направления одной из осей координат) как уравнение 7 (если $\lambda_1\lambda_2 > 0$) и как уравнение 8 (если $\lambda_1\lambda_2 < 0$).

Уравнение (106.5) может быть записано как уравнение 9 (если $\lambda_1\lambda_2 > 0$, $c_0\lambda_1 < 0$), как уравнение 10 (если $\lambda_1\lambda_2 > 0$, $c_0\lambda_1 > 0$), как уравнение 14 (если $\lambda_1\lambda_2 > 0$, $c_0 = 0$), как уравнение 11 (если $\lambda_1\lambda_2 < 0$, $c_0 \neq 0$) и как уравнение 13 (если $\lambda_1\lambda_2 < 0$, $c_0 = 0$).

Уравнение (106.6) после стандартных преобразований может быть записано как уравнение 12.

Уравнение (106.7) может быть записано как уравнение 15 (если $q_0\lambda_1 < 0$), как уравнение 16 (если $q_0\lambda_1 > 0$) и как уравнение 17 (если $q_0 = 0$).

Таким образом, уравнениями 1–17 исчерпываются все поверхности второго порядка. $\blacksquare$

§109. Матричные нормы оператора

Спектральная норма линейного оператора является, по существу, единственной подчиненной нормой, не связанной явно с базисами пространств V и W. Если же в пространствах V и W зафиксированы базисы, то возможность введения операторных норм существенно расширяется.

Пусть $e = (e_1,\dots,e_n)$ и $f = (f_1,\dots,f_m)$ – базисы пространств V и W. Введем в V и W векторную норму $\|\cdot\|_p$ согласно (107.2). Будем считать, что в пространствах V и W введены нормы одинакового типа. Обозначим через $\|\mathcal{A}\|_p$ норму оператора, подчиненную векторным нормам $\|\cdot\|_p$, через $A = (a_{ij})$ – матрицу оператора $\mathcal{A}$ в базисах e и f.

Теорема 109.1. *Для любого оператора $\mathcal{A} \in \mathcal{L}(V,W)$*

$$\|\mathcal{A}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|.$$

Доказательство. Пусть $x = \sum_{j=1}^n x_j e_j$, тогда

$$\mathcal{A}x = \sum_{j=1}^{n} x_j \mathcal{A}e_j = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij} f_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} x_j\Big) f_i.$$

Согласно (107.2)

$$\|\mathcal{A}x\|_1 = \sum_{i=1}^{m} \Big|\sum_{j=1}^{n} a_{ij}x_j\Big| \leq \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} |x_j| \sum_{i=1}^{m} |a_{ij}|. \tag{109.1}$$

Пусть k-й столбец матрицы A имеет максимальную столбцовую сумму, т.е. $\sum_{i=1}^m |a_{ik}| = \max\limits_{1\leq j\leq n} \sum_{i=1}^m |a_{ij}|$. Тогда из (109.1) следует, что $\|\mathcal{A}x\|_1 \leq \sum_{j=1}^n |x_j| \sum_{i=1}^m |a_{ik}| = \|x\|_1 \sum_{i=1}^m |a_{ik}|$. Это означает, что $\|\mathcal{A}x\|_1 \leq \sum_{i=1}^m |a_{ik}|$ для любого вектора x, у которого $\|x\|_1 = 1$, и $\|\mathcal{A}x\|_1 = \sum_{i=1}^m |a_{ik}|$ для вектора $x = e_k$ (очевидно, $\|e_k\|_1 = 1$). Следовательно, $\sum_{i=1}^m |a_{ik}| = \sup\limits_{\|x\|_1=1} \|\mathcal{A}x\|_1$. $\blacksquare$

Теорема 109.2. *Для любого оператора $\mathcal{A} \in \mathcal{L}(V,W)$*

$$\|\mathcal{A}\|_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|.$$

Доказательство аналогично доказательству теоремы 109.1. Оно предоставляется читателю. $\blacksquare$

Замечание. Векторные нормы $\|\cdot\|_2$ пространств V и W порождают спектральную норму оператора $\|\mathcal{A}\|_2$, так как норма $\|\cdot\|_2$ совпадает с евклидовой нормой $\|\cdot\|_E$, если в пространствах V и W ввести скалярное произведение так, чтобы базисы e и f стали ортонормированными.

Евклидова норма оператора. Можно строить норму линейного оператора и как норму вектора в линейном пространстве $\mathcal{L}(V,W)$. Одна из таких норм представляет интерес.

Евклидовой нормой оператора $\mathcal{A}$ называется число

$$\|\mathcal{A}\|_E = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2\right)^{1/2}. \tag{109.2}$$

Корректность этого определения следует из легко проверяемых для $\|\mathcal{A}\|_E$ аксиом нормы.

Как видно из (109.2), евклидова норма оператора легко вычисляема (по сравнению, например, с $\|\mathcal{A}\|_2$). При этом она обладает многими свойствами подчиненных норм. Перечислим их.

1. *Свойство согласованности:* $\|\mathcal{A}\|_E \leq \|\mathcal{A}\|_E\|x\|_E$, так как в силу неравенства Коши–Буняковского $\|\mathcal{A}x\|_E^2 = \sum_{i=1}^m \left|\sum_{j=1}^n a_{ij}x_j\right|^2 \leq \sum_{i=1}^m \left(\sum_{j=1}^n |a_{ij}|^2 \sum_{j=1}^n |x_j|^2\right) = \sum_{j=1}^n |x_j|^2 \cdot \|\mathcal{A}\|_E^2 = \|\mathcal{A}\|_E^2\|x\|_E^2$.

2. *Свойство мультипликативности:* $\|\mathcal{AB}\|_E \leq \|\mathcal{A}\|_E\|\mathcal{B}\|_E$, ибо $\|\mathcal{AB}\|_E^2 = \sum_{i,j}\left|\sum_k a_{ik}b_{kj}\right|^2 \leq \sum_{i,j}\left(\sum_k |a_{ik}||b_{kj}|\right)^2 \leq \sum_{i,j}\left(\sum_k |a_{ik}|^2 \cdot \sum_k |b_{kj}|^2\right) = \sum_{i,k} |a_{ik}|^2 \sum_{j,k} |b_{kj}|^2 = \|\mathcal{A}\|_E^2\|\mathcal{B}\|_E^2$.

3. $\|\mathcal{A}\|_E^2 = \mathrm{tr}\,\mathcal{A}^*\mathcal{A} = \mathrm{tr}\,\mathcal{A}\mathcal{A}^*$. Эти равенства легко проверяются непосредственно.

4. $\|\mathcal{A}\|_E^2 = \rho_1^2 + \ldots + \rho_n^2$, где $\rho_1,\dots,\rho_n$ – сингулярные числа оператора $\mathcal{A}$. Это следует из свойства 3 и теоремы 92.2 (следствие).

5. $\|\mathcal{A}\|_E \geq \|\mathcal{A}\|_2$ в силу теоремы 108.4.

6. $\|\mathcal{A}\|_E$ не изменяется при умножении оператора $\mathcal{A}$ на ортогональные (унитарные) операторы. Это следует из свойства 4 и теоремы 108.5.

§110. Экстремальные свойства собственных значений самосопряженных операторов

Пусть $\mathcal{A}$ – самосопряженный оператор в евклидовом (унитарном) пространстве V. Построим в пространстве V ортонормированный базис

$$e_1,\dots,e_n \tag{110.1}$$

из собственных векторов оператора $\mathcal{A}$, отвечающих собственным значениям

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n. \tag{110.2}$$

Такой нумерации собственных векторов и собственных значений будем придерживаться всюду в этом параграфе. Под нормой $\|\cdot\|$ будем

Эллиптический параболоид (рис. 8). Поверхность, определяемая уравнением 7, называется *эллиптическим параболоидом*. Как следует из канонического уравнения, плоскости $x = 0$ и $y = 0$ являются плоскостями симметрии. Центра симметрии нет. Плоскости $z = h$, $h > 0$, пересекают параболоид по эллипсам, размеры которых неограниченно возрастают при $h \to \infty$. Плоскости $y = h$ и $x = h$ пересекают параболоид по параболам, ветви которых направлены вверх.

Гиперболический параболоид (рис. 9). Поверхность, определяемая уравнением 8, называется *гиперболическим параболоидом*. Плоскости $x = 0$ и $y = 0$ являются плоскостями симметрии. Центра симметрии нет. Плоскости $y = h$ и $x = h$ пересекают параболоид по параболам, ветви первой параболы направлены вверх, а второй – вниз.

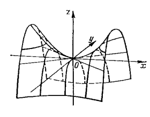


Рис. 9

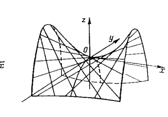


Рис. 10

Плоскости $z = h$ пересекают параболоид:

а) при $h < 0$ по гиперболам $-\frac{x^2}{a_1^2} + \frac{y^2}{b_1^2} = 1$, где $a_1^2 = -2a^2h$, $b_1^2 = -2b^2h$;

б) при $h > 0$ по гиперболам $\frac{x^2}{a_2^2} - \frac{y^2}{b_2^2} = 1$, где $a_2^2 = 2a^2h$, $b_2^2 = 2b^2h$;

в) при $h = 0$ по паре пересекающихся прямых.

Как и однополостный гиперболоид, гиперболический параболоид покрыт двумя различными семействами прямолинейных образующих (рис. 10). Те же рассуждения, что и в случае однополостного гиперболоида, показывают, что *через любую точку $M_0(x_0, y_0, z_0)$ гиперболического параболоида проходят две различные прямые, целиком лежащие на этой поверхности.* Эти прямые определяются плоскостями

$$L: \begin{cases} \alpha z = \beta\left(\frac{x}{a} + \frac{y}{b}\right), \\ 2\beta = \alpha\left(\frac{x}{a} - \frac{y}{b}\right), \\ \alpha^2 + \beta^2 \neq 0 \end{cases} \quad \text{и} \quad L^*: \begin{cases} \gamma z = \delta\left(\frac{x}{a} - \frac{y}{b}\right), \\ 2\delta = \gamma\left(\frac{x}{a} + \frac{y}{b}\right), \\ \gamma^2 + \delta^2 \neq 0, \end{cases}$$

где коэффициенты α, β, γ, δ находятся из условия, что эти прямые проходят через точку $M_0(x_0, y_0, z_0)$.

Цилиндрические поверхности. Все остальные поверхности называются цилиндрами. Поверхности, определяемые уравнениями 9, 11 и 12, называются соответственно *эллиптическим цилиндром* (рис. 11), *гиперболическим цилиндром* (рис. 12) и *параболическим цилиндром* (рис. 13). Уравнения этих поверхностей не зависят от z, поэтому все сечения плоскостями $z = h$ совпадают. Для цилиндрических поверхностей достаточно найти сечение плоскостью $z = 0$, чтобы выяснить форму поверхности. Отметим, наконец, что прямая, проходящая через точку $M_0(x_0, y_0, z_0)$ поверхности и параллельная оси Oz, является прямолинейной образующей, так как любая точка этой прямой имеет координаты $(x_0, y_0, z_0 + t)$.

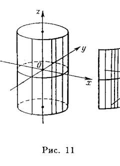


Рис. 11

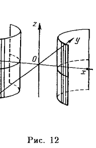


Рис. 12

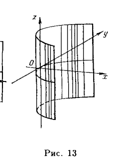


Рис. 13

Нет ни одной точки пространства, координаты которой удовлетворяют уравнению 10. Принято называть его уравнением *мнимого эллиптического цилиндра*. Уравнение 13 определяет *пару пересекающихся плоскостей* (рис. 14).

Точки поверхности, определяемой уравнением 14, составляют прямую. Это уравнение называют уравнением *пары мнимых пересекающихся плоскостей* (рис. 15).

Примеры. 1. Скалярное произведение (x, y) в унитарном пространстве является эрмитовой полуторалинейной формой.

2. Примером полуторалинейной формы в n-мерном комплексном пространстве V с базисом $e_1, \ldots, e_n$ может служить функция, определенная равенством: $\forall x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^n y_i e_i$

$$\mathcal{A}(x, y) = \sum_{i,j=1}^n a_{ij} x_i \overline{y_j}, \tag{103.1}$$

где a_{ij} $(i, j = \overline{1, n})$ – фиксированные комплексные числа.

В действительности выражение (103.1) дает *общий вид полуторалинейной формы в базисе e*. Это доказывается так же, как и аналогичное соотношение (101.2) для билинейной формы. Следует отметить, что для полуторалинейных форм остаются в силе и все другие утверждения (а также их доказательства), касающиеся билинейных форм. Незначительные изменения в них носят очевидный характер. В частности:

– компактная форма (103.1) имеет вид $\mathcal{A}(x, y) = x_e^T A_e \overline{y_e}$ или $\mathcal{A}(x, y) = y_e^H A_e^T x_e$;

– если e и $f = eQ$ – два базиса пространства, то $A_f = Q^T A_e \overline{Q}$;

– полуторалинейная форма эрмитова тогда и только тогда, когда ее матрица в любом базисе эрмитова.

Теорема 103.1. *Полуторалинейная форма эрмитова тогда и только тогда, когда $\mathcal{A}(x, x) \in \mathbb{R}$, $\forall x \in V$.*

Доказательство. Необходимость очевидна, так как

$$\mathcal{A}(x, x) = \overline{\mathcal{A}(x, x)}, \ \forall x \in V.$$

Достаточность вытекает из равенства: $\forall x, y \in V$

$$\mathcal{A}(x, y) = \tfrac{1}{4}\big(\mathcal{A}(x + y, x + y) - \mathcal{A}(x - y, x - y) + i\mathcal{A}(x + iy, x + iy) - i\mathcal{A}(x - iy, x - iy)\big), \tag{103.2}$$

которое справедливо для любой полуторалинейной формы $\mathcal{A}(x, y)$. ■

Эрмитовы формы. Пусть $\mathcal{A}(x, y)$ – эрмитова полуторалинейная форма в комплексном пространстве V. *Эрмитовой квадратичной формой* (или, короче, *эрмитовой формой*) называется отображение $\mathcal{A}: V \to \mathbb{C}$, которое каждому вектору $x \in V$ ставит в соответствие число $\mathcal{A}(x, x)$. Полуторалинейная форма $\mathcal{A}(x, y)$ при этом называется *полярной полуторалинейной формой* к эрмитовой форме $\mathcal{A}(x)$.

Эрмитова квадратичная форма обладает всеми свойствами (с очевидными изменениями) обычных квадратичных форм (§101). Отметим некоторые из них.

1. Существует взаимно однозначное соответствие между эрмитовыми полуторалинейными и эрмитовыми квадратичными формами пространства. Это следует из (103.2).

2. Матрица эрмитовой квадратичной формы в любом базисе эрмитова.

3. Общий вид эрмитовой квадратичной формы: $\forall x = \sum_{i=1}^n x_i e_i$

$$\mathcal{A}(x, x) = \sum_{i,j=1}^n a_{ij} x_i \overline{x_j}, \quad a_{ij} = \overline{a_{ji}} \ (i, j = \overline{1, n}),$$

или, в компактной форме,

$$\mathcal{A}(x, x) = x_e^T A_e \overline{x_e} = x_e^H A_e^T x_e, \quad A_e^H = A_e.$$

4. Канонический вид эрмитовой квадратичной формы:

$$\mathcal{A}(x, x) = \sum_{k=1}^r \lambda_k |x_k|^2, \quad \forall x = \sum_{i=1}^n x_i e_i,$$

где $r = \operatorname{rg} \mathcal{A}(x, x)$, $\lambda_k \in \mathbb{R}$ $(k = \overline{1, r})$.

5. Метод Лагранжа приведения к каноническому виду применим и для эрмитовой формы. Изменение алгоритма, описанного в доказательстве теоремы 101.7, состоит лишь в том, что на каждом шаге выделяется полный квадрат модуля (например, на первом шаге: $f(x_1, \ldots, x_n) = a_{11}|x_1 + \sum_{k=2}^n \frac{a_{1k}}{a_{11}} x_k|^2 + g(x_2, \ldots, x_n)$).

6. Остаются справедливыми и формулы Якоби (101.13).

7. Как следует из теоремы 103.1, эрмитова квадратичная форма принимает только вещественные значения. Это свойство сближает ее с вещественной квадратичной формой (§102). В частности, для эрмитовых квадратичных форм справедливы закон инерции, сигнатурное правило Якоби, критерий Сильвестра. Практически без изменения остаются и доказательства этих фактов.

8. И, наконец, отображение $\mathcal{A}: V \times V \to \mathbb{C}$ есть скалярное произведение в комплексном пространстве V тогда и только тогда, когда оно является полуторалинейной формой, полярной к положительно определенной эрмитовой форме.

§104. Квадратичные формы в евклидовом (унитарном) пространстве

Мы отмечали (§102), что ни канонический базис, ни канонический вид квадратичной формы не определены однозначно – однозначно лишь число всех ненулевых квадратов (в произвольном линейном пространстве) и число положительных и отрицательных квадратов (в вещественном линейном пространстве). В евклидовом пространстве (в унитарном пространстве для эрмитовых квадратичных форм) положение иное, если рассматривать ортонормированные базисы.

Теорема 104.1. *Для любой квадратичной формы $\mathcal{A}(x, x)$ в евклидовом пространстве E существует, и притом единственный, симметрический оператор $\mathcal{H} \in \mathcal{L}(E, E)$ такой, что*

$$\mathcal{A}(x, x) = (\mathcal{H}x, x), \quad \forall x \in E. \tag{104.1}$$

4. В пространствах матриц $\mathbb{R}^{m \times n}$ и $\mathbb{C}^{m \times n}$ рассматривают следующие нормы матрицы $A = (a_{ij})$:

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^m |a_{ij}|; \quad \|A\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^n |a_{ij}|;$$

$$\|A\|_E = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{1/2}.$$

Сходимость по норме.

Утверждение 1. *В нормированном пространстве V отображение $\rho: V \times V \to \mathbb{R}$, определенное равенством*

$$\rho(x, y) = \|x - y\|, \quad \forall x, y \in V,$$

является метрикой.

Аксиомы метрики непосредственно вытекают из аксиом нормы. ■

Таким образом, в нормированном пространстве можно ввести расстояние между векторами и, значит, пользоваться предельным переходом. Последовательность векторов $\{x^{(k)}\}$ в нормированном пространстве V называется *сходящейся по норме* к вектору $a \in V$, если $\lim_{k \to \infty} \|x^{(k)} - a\| = 0$, при этом вектор a называется *пределом последовательности $\{x^{(k)}\}$ по норме $\|\cdot\|$*.

Обозначение: $\lim_{k \to \infty} x^{(k)} = a$ или $x^{(k)} \to a$.

Утверждение 2. *Сходящаяся по норме последовательность имеет единственный предел.*

Доказательство повторяет доказательство аналогичной теоремы для числовой последовательности [7] и основано на аксиоме треугольника: $\|a - b\| = \|a - x^{(k)} + x^{(k)} - b\| \le \|x^{(k)} - a\| + \|x^{(k)} - b\|$, где a и b – два предела последовательности $x^{(k)}$. ■

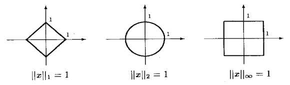
Пусть $x_0 \in V$ и $r > 0$. Множество $S(x_0, r) = \{x \in V \mid \|x - x_0\| = r\}$ называется *сферой радиуса r с центром x_0 по норме $\|\cdot\|$*, а множество $B(x_0, r) = \{x \in V \mid \|x - x_0\| \le r\}$ – замкнутым *шаром радиуса r с центром x_0 по норме $\|\cdot\|$*.

На рис. 1 точками плоскости изображены сферы радиуса единица с центром $x_0 = (0, 0)$ по нормам $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ в арифметическом пространстве $\mathbb{R}^2$.

В дальнейшем сферы и шары по евклидовой норме будем обозначать символами $B_E(x_0, r)$ и $S_E(x_0, r)$.

Утверждение 3. *Из любой последовательности векторов $x^{(k)} \in B_E(x_0, r)$ (или $S_E(x_0, r)$) можно выделить подпоследовательность, сходящуюся по норме $\|\cdot\|_E$ к вектору $a \in B_E(x_0, r)$ ($S_E(x_0, r)$ соответственно).*

Доказательство. Без ограничения общности можно считать, что $x_0 = \theta$. Доказательство проведем для сферы $S_E(r) = \{x \in$



$\|x\|_1 = 1$ $\|x\|_2 = 1$ $\|x\|_\infty = 1$

Рис.1

$V \mid \|x\|_E = r\}$. Пусть $e_1, \ldots, e_n$ – ортонормированный базис пространства V и $x^{(k)} = \sum_{i=1}^n x_i^{(k)} e_i \in S_E(r)$. Тогда

$$\|x^{(k)}\|_E = \left(\sum_{i=1}^n |x_i^{(k)}|^2\right)^{1/2} = r.$$

Это означает ограниченность координат векторов $x^{(k)}$ рассматриваемой последовательности. Согласно теореме Больцано–Вейерштрасса [7] из этой последовательности можно выделить (последовательно покоординатно) подпоследовательность $\{x^{(k_m)}\}$. Пусть $x^{(k_m)}$ имеет координаты $x_1^{(k_m)}, \ldots, x_n^{(k_m)}$, сходящиеся соответственно к $a_1, \ldots, a_n$. Положим $a = \sum_{i=1}^n a_i e_i$. Тогда

$$\|x^{(k_m)} - a\|_E = \left(\sum_{i=1}^n \left|x_i^{(k_m)} - a_i\right|^2\right)^{1/2} \to 0,$$

следовательно, подпоследовательность $\{x^{(k_m)}\}$ сходится к вектору a по евклидовой норме.

Покажем, что $a \in S_E(r)$. Действительно, в очевидном неравенстве $\big|\|x\| - \|y\|\big| \le \|x - y\|$ положим $x = x^{(k_m)}$, $y = a$. Тогда $\big|\|x^{(k_m)}\|_E - \|a\|_E\big| \le \|x^{(k_m)} - a\|_E$, откуда следует, что

$$\|x^{(k_m)}\|_E - \|x^{(k_m)} - a\|_E \le \|a\|_E \le \|x^{(k_m)}\|_E + \|x^{(k_m)} - a\|_E$$

или, с учетом того, что $\|x^{(k_m)} - a\|_E \to 0$,

$$\|x^{(k_m)}\|_E - \varepsilon \le \|a\|_E \le \|x^{(k_m)}\|_E + \varepsilon, \quad \forall \varepsilon > 0,$$

если m достаточно велико. Следовательно, $\|a\|_E = r$ и $a \in S_E(r)$. ■

Возникает естественный вопрос: может ли одна и та же последовательность быть сходящейся по одной норме и не сходящейся по другой? Ответим на этот вопрос для конечномерных пространств.

Эквивалентность норм в конечномерном пространстве. Две нормы $\|\cdot\|_1$ и $\|\cdot\|_2$ в линейном пространстве V называются *эквивалентными*, если существуют такие числа $c_1 > 0$, $c_2 > 0$, что для любого вектора $x \in V$ выполняются неравенства

$$\|x\|_1 \le c_1 \|x\|_2 \quad \text{и} \quad \|x\|_2 \le c_2 \|x\|_1.$$

Теорема 107.1. *В конечномерном пространстве любые две*

Глава XVIII. Поверхности второго порядка

§105. Гиперповерхности второго порядка в евклидовом пространстве

Общее уравнение. Пусть $\mathcal{A}(x, x)$ – ненулевая квадратичная форма, $g(x)$ – линейная форма, заданные в евклидовом пространстве E, и c – вещественная константа. Множество всех векторов $x \in E$, удовлетворяющих условию

$$\mathcal{A}(x, x) + 2g(x) + c = 0, \tag{105.1}$$

называется *гиперповерхностью второго порядка в евклидовом пространстве E*.[1] Уравнение (105.1) называется *общим уравнением гиперповерхности второго порядка*.

Пусть $e = (e_1, \ldots, e_n)$ – базис пространства E, $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ – матрица квадратичной формы $\mathcal{A}(x, x)$ в этом базисе, $b_i = g(e_i)$, $i = \overline{1, n}$, – коэффициенты линейной формы и $x = \sum_{i=1}^n x_i e_i$. Тогда в соответствии с (101.7), (81.1) общее уравнение (105.1) может быть записано в виде

$$\sum_{i,j=1}^n a_{ij} x_i x_j + 2\sum_{i=1}^n b_i x_i + c = 0, \quad a_{ij} = a_{ji}, \tag{105.2}$$

или, в компактной форме,

$$X^T A X + 2b^T X + c = 0, \quad A = A^T, \tag{105.3}$$

где $X = (x_1, \ldots, x_n)^T$, $b = (b_1, \ldots, b_n)^T$.

Приведенные уравнения. Исследование гиперповерхности второго порядка в n-мерном пространстве проводится по той же схеме, по которой исследуются линии второго порядка на плоскости (§58, 59). Отметим основные моменты этого исследования с общих позиций.

Пусть e – ортонормированный базис E и пусть в этом базисе уравнение гиперповерхности имеет вид (105.2).

1. Приведем квадратичную форму $\mathcal{A}(x, x)$ к главным осям, т.е. найдем ортонормированный базис f, в котором квадратичная форма $\mathcal{A}(x, x)$ имеет канонический вид. Отметим, что канонические коэффициенты $\lambda_1, \ldots, \lambda_n$ будут при этом определены однозначно (§104,

[1] В случае, когда $\dim E = 2$, геометрический образ, определяемый уравнением (105.1), называется линией второго порядка. Однако всюду в дальнейшем мы употребляем общий термин "поверхность", не оговаривая особо этот случай.

замечание 1). Не нарушая общности рассуждений, можно считать, что $\lambda_k \neq 0$ при $k \le r$ (где $r = \operatorname{rg} \mathcal{A}(x, x)$) и $\lambda_k = 0$ при $k > r$.

При переходе от базиса e к базису $f = eQ$ уравнение (105.2) ортогональным преобразованием координат (теорема 70.5) преобразуется в уравнение

$$\sum_{k=1}^r \lambda_k x_k'^2 + 2\sum_{k=1}^n b_k' x_k' + c = 0. \tag{105.4}$$

2. Пусть $a \in E$. Отображение $\varphi: E \to E$, определенное равенством $\varphi(x) = x + a$, $\forall x \in E$, называется *параллельным переносом пространства E на вектор a*. Очевидно, $\operatorname{im} \varphi = E$.

Если в уравнении (105.4) $\lambda_k \neq 0$ для некоторого k, то $\lambda_k x_k'^2 + 2b_k' x_k' = \lambda_k (x_k' + b_k'/\lambda_k)^2 - b_k'^2/\lambda_k$. Положив $x_k'' = x_k' + b_k'/\lambda_k$, можно добиться исчезновения переменной x_k'' в первой степени. Выполнение таких преобразований для всех $k = \overline{1, r}$ равносильно параллельному переносу на вектор с координатами $(b_1'/\lambda_1, \ldots, b_r'/\lambda_r, 0, \ldots, 0)$. В результате этого переноса уравнение (105.4) преобразуется в уравнение

$$\lambda_1 x_1''^2 + \ldots + \lambda_r x_r''^2 + 2\sum_{k=r+1}^n b_k' x_k'' + c' = 0, \tag{105.5}$$

где $c' = c - \sum_{k=1}^r b_k'^2/\lambda_k$.

3. Исследуем уравнение (105.5). Возможны два случая:

а) $b_{r+1}' = \ldots = b_n' = 0$;

б) $\exists b_k' \neq 0$, $k > r$.

В случае "а" уравнение (105.5) имеет вид

$$\lambda_1 x_1''^2 + \ldots + \lambda_r x_r''^2 + c' = 0.$$

В случае "б" выполним еще одно ортогональное преобразование координат. Матрица S этого преобразования строится следующим образом. Рассмотрим арифметический вектор $(b_{r+1}', \ldots, b_n') \in \mathbb{R}^{n-r}$. По условию он ненулевой и его можно нормировать: если $\alpha = \left(\sum_{k=r+1}^n b_k'^2\right)^{1/2}$, то вектор $s_1 = (b_{r+1}'/\alpha, \ldots, b_n'/\alpha)$ представляет собой ортонормированную систему из одного вектора. Дополним этот вектор до ортонормированного базиса $s_1, s_2, \ldots, s_{n-r}$ пространства $\mathbb{R}^{n-r}$. В обозначениях $s_i = (s_{i,r+1}, \ldots, s_{in})$, $i = \overline{1, n-r}$, искомая матрица S имеет вид

$$S = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & O & \\ & & 1 & & & \\ & & & s_{1,r+1} & \ldots & s_{1n} \\ & O & & \ldots & \ldots & \ldots \\ & & & s_{n-r,r+1} & \ldots & s_{n-r,n} \end{bmatrix}.$$

Ограниченный оператор. Линейный оператор $\mathcal{A} \in \mathcal{L}(V, W)$ называется *ограниченным*, если единичный шар в V он переводит в ограниченное по норме пространства W множество, т.е. если существует число $c > 0$ такое, что для всех векторов $x \in V$, для которых $\|x\| = 1$, выполняется неравенство $\|\mathcal{A}x\| \le c$, или, в равносильной формулировке, если существует число $c > 0$ такое, что $\|\mathcal{A}x\| \le c\|x\|$ для всех векторов $x \in V$.

Теорема 108.2. *В конечномерных нормированных пространствах V и W любой линейный оператор $\mathcal{A} \in \mathcal{L}(V, W)$ ограничен.*

Доказательство. Пусть $e_1, \ldots, e_n$ – базис пространства V и $x = \sum_{i=1}^n x_i e_i$. Тогда согласно аксиомам нормы и неравенству Коши–Буняковского

$$\|\mathcal{A}x\| \le \sum_{i=1}^n |x_i| \cdot \|\mathcal{A}e_i\| \le \left(\sum_{i=1}^n \|\mathcal{A}e_i\|^2\right)^{1/2} \left(\sum_{i=1}^n |x_i|^2\right)^{1/2} = M\|x\|_E,$$

где $M = \left(\sum_{i=1}^n \|\mathcal{A}e_i\|^2\right)^{1/2}$. Отсюда с учетом (107.4) получим, что $\|\mathcal{A}x\| \le c\|x\|$, где $c = Mc_2 > 0$. ■

Подчиненная норма. Пусть V, W – конечномерные пространства и $\mathcal{A} \in \mathcal{L}(V, W)$. Из теоремы 108.2 следует ограниченность оператора $\mathcal{A}$, т.е. существование числа $c > 0$ такого, что $\|\mathcal{A}x\| \le c\|x\|$, $\forall x \in V$, или, в равносильной формулировке, $\|\mathcal{A}x\|/\|x\| \le c$, $\forall x \neq \theta$. Это означает, что числовое множество

$$\left\{ \frac{\|\mathcal{A}x\|}{\|x\|} \ \middle| \ x \neq \theta \right\}$$

ограничено сверху и, следовательно, для него существует точная верхняя грань. Положим

$$\mu(\mathcal{A}) = \sup_{x \neq \theta} \frac{\|\mathcal{A}x\|}{\|x\|}. \tag{108.1}$$

Теорема 108.3. *Отображение $\mu(\mathcal{A})$ является нормой в пространстве $\mathcal{L}(V, W)$.*

Доказательство. Очевидно, что $\mu(\mathcal{A}) \ge 0$ для любого оператора $\mathcal{A} \in \mathcal{L}(V, W)$, при этом равенство $\mu(\mathcal{A}) = 0$ означает, что $\|\mathcal{A}x\| = 0$, $\forall x \in V$, т.е. $\mathcal{A}x = \theta$, $\forall x \in V$ или $\mathcal{A} = \mathcal{O}$. Аксиомы 2 и 3 вытекают из свойств точной верхней грани. ■

Норма $\mu(\mathcal{A})$ называется *нормой оператора $\mathcal{A}$, подчиненной (порожденной) векторным нормам пространств V и W*.

Обозначение: $\|\mathcal{A}\|$. Итак,

$$\|\mathcal{A}\| = \sup_{x \neq \theta} \frac{\|\mathcal{A}x\|}{\|x\|} = \sup_{\|x\|=1} \|\mathcal{A}x\|.$$

Свойства подчиненной нормы. 1. Подчиненная норма обладает свойством согласованности: $\|\mathcal{A}x\| \le \|\mathcal{A}\| \cdot \|x\|$, $\forall x \in V$, так как, согласно (108.1), $\|\mathcal{A}x\|/\|x\| \le \mu(\mathcal{A}) = \|\mathcal{A}\|$.

2. Подчиненная норма – наименьшая из всех согласованных норм, так как является точной верхней гранью множества, тогда как любая согласованная норма – верхней гранью этого множества.

3. Подчиненная норма обладает *свойством мультипликативности*, т.е. $\|\mathcal{AB}\| \le \|\mathcal{A}\| \cdot \|\mathcal{B}\|$, так как $\|\mathcal{AB}\| = \sup_{\|x\|=1} \|\mathcal{AB}x\| \le \|\mathcal{A}\| \cdot \sup_{\|x\|=1} \|\mathcal{B}x\| = \|\mathcal{A}\| \cdot \|\mathcal{B}\|$.

Замечание 2. Подчиненная норма определяется не только самим оператором, но и векторными нормами пространств V и W. Меняя векторные нормы, можно изменить и норму оператора. Рассмотрим наиболее часто употребляемые нормы.

Спектральная норма. Пусть V, W – евклидовы (унитарные) пространства. Норма линейного оператора $\mathcal{A} \in \mathcal{L}(V, W)$, порожденная евклидовыми нормами вектора, называется *спектральной нормой*. Обозначение: $\|\mathcal{A}\|_2$. Итак,

$$\|\mathcal{A}\|_2 = \sup_{\|x\|_E=1} \|\mathcal{A}x\|_E = \sup_{(x,x)=1} \sqrt{(\mathcal{A}x, \mathcal{A}x)}.$$

Теорема 108.4. *Спектральная норма оператора равна максимальному сингулярному числу этого оператора.*

Доказательство. Пусть $e_1, \ldots, e_n$ – ортонормированный базис из собственных векторов оператора $\mathcal{A}^*\mathcal{A}$, а $\rho_1, \ldots, \rho_n$ – соответствующие сингулярные числа оператора $\mathcal{A}$ (§100), $x = \sum_{i=1}^n x_i e_i$, $\rho_k = \max_{1 \le i \le n} \rho_i$. Тогда $\|\mathcal{A}x\|_E^2 = (\mathcal{A}x, \mathcal{A}x) = (\mathcal{A}^*\mathcal{A}x, x) = \left(\sum_{i=1}^n \rho_i^2 x_i e_i, \sum_{j=1}^n x_j e_j\right) = \sum_{i=1}^n \rho_i^2 |x_i|^2 \le \rho_k^2 \sum_{i=1}^n |x_i|^2 = \rho_k^2 \|x\|_E^2$. Отсюда следует, что $\|\mathcal{A}x\|_E \le \rho_k$, если $\|x\|_E = 1$, и $\|\mathcal{A}x\|_E = \rho_k$, если $x = e_k$ (очевидно, $\|e_k\|_E = 1$). Значит, $\rho_k = \max_{\|x\|_E=1} \|\mathcal{A}x\|_E = \sup_{\|x\|_E=1} \|\mathcal{A}x\|_E = \|\mathcal{A}\|_2$. ■

Следствие 1. *Спектральная норма нормального оператора равна абсолютному значению максимального по модулю собственного значения этого оператора.*

Теорема 108.5. *Сингулярные числа линейного оператора в евклидовом (унитарном) пространстве не изменяются при умножении оператора на ортогональный (унитарный) оператор.*

Доказательство. Пусть $\mathcal{B} = \mathcal{UA}V$, где $\mathcal{U}^*\mathcal{U} = \mathcal{I}$, $\mathcal{V}^*\mathcal{V} = \mathcal{I}$. Тогда $\mathcal{B}^*\mathcal{B} = \mathcal{V}^*\mathcal{A}^*\mathcal{A}\mathcal{V}$, следовательно, матрицы операторов $\mathcal{B}^*\mathcal{B}$ и $\mathcal{A}^*\mathcal{A}$ подобны и их собственные значения совпадают. ■

Следствие 2. *Спектральная норма линейного оператора не изменяется при умножении оператора на ортогональный (унитарный) оператор.*

Гиперповерхность, определяемая уравнением (105.13), называется *параболоидом*.

3. Пусть $0 < r < n - 1$. Тогда уравнение (105.2) приводится к уравнению типа (105.6), если $\operatorname{rg} B \le r + 1$, и к уравнению типа (105.7), если $\operatorname{rg} B = r + 2$. Эти уравнения однозначно преобразуются в первом случае в уравнение

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_k^2}{a_k^2} - \frac{x_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{x_r^2}{a_r^2} = c, \tag{105.14}$$

где $c = 1$, если $\operatorname{rg} B = r + 1$, и $c = 0$, если $\operatorname{rg} B = r$; во втором случае в уравнение

$$\frac{x_1^2}{a_1^2} + \ldots + \frac{x_k^2}{a_k^2} - \frac{x_{k+1}^2}{a_{k+1}^2} - \ldots - \frac{x_r^2}{a_r^2} = 2px_{r+1}, \quad p > 0. \tag{105.15}$$

Гиперповерхности, определяемые уравнениями (105.12), (105.14), (105.15), называются *цилиндрами*. Так как уравнения, определяющие эти уравнения, взаимно однозначны друг с другом, то *общее уравнение (105.2) определяет одну и только одну из перечисленных поверхностей.*

§106. Алгебраические поверхности второго порядка

Общее уравнение. Под общим уравнением алгебраической поверхности второго порядка в системе координат $Oxyz$ пространства понимают уравнение вида

$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2b_1x + 2b_2y + 2b_3z + c = 0, \tag{106.1}$$

где не все коэффициенты a_{ij} $(i, j = \overline{1, 3})$ равны нулю, $a_{ij} = a_{ji}$ $(i, j = \overline{1, 3})$. Группа членов второго порядка образует здесь квадратичную форму от переменных x, y, z (или, в векторной терминологии, от радиус-вектора точки (x, y, z)), а группа линейных членов – линейную форму от этих переменных. Поэтому уравнение (106.1) является уравнением гиперповерхности второго порядка в евклидовом пространстве V_3. Обозначения, соответствующие общей теории гиперповерхностей (§105), в данном случае имеют вид: $X = (x, y, z)^T$, $b = (b_1, b_2, b_3)^T$,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{12} & a_{22} & a_{23} & b_2 \\ a_{13} & a_{23} & a_{33} & b_3 \\ b_1 & b_2 & b_3 & c \end{bmatrix}.$$

В этих обозначениях общее уравнение (106.1) может быть записано в компактной форме:

$$X^T A X + 2b^T X + c = 0, \quad A = A^T. \tag{106.2}$$

Приведенные уравнения. Пусть общее уравнение (106.1) задано в прямоугольной декартовой системе координат. Согласно общей теории гиперповерхностей (§105) уравнение (106.1) с помощью ортогонального преобразования координат (т.е. простым вращением и простым отражением) и параллельного переноса (т.е. переноса начала) приводится к одному и только одному из двух типов приведенных уравнений (105.6) и (105.7). В соответствии с общей схемой классификации гиперповерхностей эти уравнения можно разбить на следующие пять простейших уравнений в зависимости от значения $r = \operatorname{rg} A$.

1. При $r = 3$:

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + a_0 = 0, \quad \lambda_1 \lambda_2 \lambda_3 \neq 0. \tag{106.3}$$

2. При $r = 2$:

$$\lambda_1 x^2 + \lambda_2 y^2 + 2b_0 z = 0, \quad \lambda_1 \lambda_2 b_0 \neq 0 \tag{106.4}$$

или

$$\lambda_1 x^2 + \lambda_2 y^2 + c_0 = 0, \quad \lambda_1 \lambda_2 \neq 0. \tag{106.5}$$

3. При $r = 1$:

$$\lambda_1 x^2 + 2p_0 y = 0, \quad \lambda_1 p_0 \neq 0 \tag{106.6}$$

или

$$\lambda_1 x^2 + q_0 = 0, \quad \lambda_1 \neq 0. \tag{106.7}$$

Отметим, что все коэффициенты уравнений (106.3)–(106.7) определены однозначно общими инвариантами гиперповерхностей: $I_1 = \operatorname{tr} A$, $I_2 = \det A$, $K_4 = \det B$ и специальными инвариантами алгебраической поверхности второго порядка: I_2 – сумма главных миноров второго порядка матрицы A и K_2, K_3 – суммы главных миноров второго и третьего порядков матрицы B.

Величины I_2, K_2, K_3 – инварианты ортогонального преобразования координат, так как являются коэффициентами характеристических многочленов матриц A, B (докажите!). Однако при параллельном переносе величина K_3 остается неизменной, только если $I_3 = K_4 = 0$, а число K_2 – если $I_2 = I_3 = K_3 = K_4 = 0$. В этом несложно убедиться непосредственно аналитически тому, как это было сделано для линий второго порядка (теорема 58.2).

Для уравнений (106.3)–(106.7) легко вычислить все инварианты, а именно:

(106.3): $I_3 \neq 0$;

понимать евклидову норму $\|\cdot\|_E$, так что если $x = \sum_{i=1}^n x_i e_i$, то $\|x\| = \sqrt{(x, x)} = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$.

Теорема 110.1. *Для самосопряженного оператора $\mathcal{A}$*

$$\lambda_1 = \max_{\|x\|=1} (\mathcal{A}x, x), \qquad \lambda_n = \min_{\|x\|=1} (\mathcal{A}x, x).$$

Доказательство. Для любого вектора $x = \sum_{i=1}^n x_i e_i \in V$ имеем $\mathcal{A}x = \sum_{i=1}^n x_i \lambda_i e_i \in V$, и, с учетом ортонормированности базиса, $(\mathcal{A}x, x) = \sum_{i=1}^n \lambda_i |x_i|^2$. В силу (110.2) отсюда следует, что $\lambda_1 \ge (\mathcal{A}x, x) \ge \lambda_n$, если $\|x\| = 1$; причем $(\mathcal{A}e_1, e_1) = \lambda_1$, $(\mathcal{A}e_n, e_n) = \lambda_n$ и $\|e_1\| = 1$, $\|e_n\| = 1$. Следовательно, λ_1 и λ_n – наибольшее и наименьшее значения $(\mathcal{A}x, x)$ на единичной евклидовой сфере. ■

Замечание. Эта теорема дает экстремальные свойства и квадратичной формы в евклидовом (унитарном) пространстве: *на единичной сфере квадратичная форма $\mathcal{A}(x, x)$ принимает экстремальные значения на тех векторах, которые являются собственными векторами самосопряженного оператора $\mathcal{H}$* (теорема 104.1).

Кстати, в терминах квадратичной формы $f(x_1, \ldots, x_n)$ от переменных $x_1, \ldots, x_n$ задача поиска экстремальных значений квадратичной формы представляет собой классическую задачу на условный экстремум. Одним из методов ее решения является метод Лагранжа [7].

Теорема 110.2. *Если L – линейная оболочка собственных векторов $e_{i_1}, \ldots, e_{i_k}$ $(i_1 < \ldots < i_k)$ из (110.1) самосопряженного оператора $\mathcal{A}$, то*

$$\lambda_{i_1} = \max_{\|x\|=1, x \in L} (\mathcal{A}x, x), \qquad \lambda_{i_k} = \min_{\|x\|=1, x \in L} (\mathcal{A}x, x). \tag{110.3}$$

Доказательство аналогично доказательству теоремы 110.1. ■

Теорема 110.3 (теорема Куранта–Фишера). *Для собственных значений самосопряженного оператора $\mathcal{A}$ справедливо представление*

$$\lambda_k = \max_{L_k} \min_{\|x\|=1, x \in L_k} (\mathcal{A}x, x), \tag{110.4}$$

где максимум берется по всевозможным k-мерным подпространствам L_k пространства V.

Доказательство. Пусть L_k – произвольное k-мерное подпространство и W_{n-k+1} – линейная оболочка собственных векторов $e_k, \ldots, e_n$ из (110.1) оператора $\mathcal{A}$. Так как $\dim L_k + \dim W_{n-k+1} = n + 1$, то $L_k \cap W_{n-k+1} \neq \{\theta\}$. Пусть $x_0 \in L_k \cap W_{n-k+1}$ и $\|x_0\| = 1$. Тогда, согласно (110.3), $(\mathcal{A}x_0, x_0) \le \lambda_k$, поэтому $\min_{\|x\|=1, x \in L_k} (\mathcal{A}x_0, x_0) \le \lambda_k$, так что и $\max_{L_k} \min_{\|x\|=1, x \in L_k} (\mathcal{A}x_0, x_0) \le \lambda_k$. Равенство в (110.4) достигается для $L_k = \mathcal{L}(e_1, \ldots, e_k)$. ■

§111. Линейные операторные уравнения

Рассмотрим проблему решения систем линейных алгебраических уравнений с точки зрения свойств линейного оператора.

Пусть V, W – евклидовы (унитарные) пространства, $\mathcal{A} \in \mathcal{L}(V, W)$, $u \in W$. Уравнение

$$\mathcal{A}z = u \tag{111.1}$$

называется *линейным операторным уравнением*, вектор u – *правой частью*, вектор z – *решением*. Очевидно, в матричной записи операторное уравнение превращается в систему линейных алгебраических уравнений и, следовательно, все свойства систем уравнений можно переносить на операторные уравнения и наоборот. Отметим некоторые из этих свойств в терминах операторных уравнений.

Однородное уравнение

$$\mathcal{A}^* w = \theta \tag{111.2}$$

называется *сопряженным* к уравнению (111.1).

Теорема 111.1 (альтернатива Фредгольма). *Либо основное уравнение (111.1) имеет решение при любой правой части $u \in W$, либо сопряженное к нему уравнение имеет нетривиальное решение.*

Доказательство. Пусть $r = \operatorname{rg} \mathcal{A}$, $n = \dim W$. Возможны два случая: либо $r = n$, либо $r < n$. Условие $r = n$ равносильно условию $\operatorname{im} \mathcal{A} = W$, которое означает, что уравнение (111.1) имеет решение при любом $u \in W$. При этом так как $\operatorname{rg} \mathcal{A} = \operatorname{rg} \mathcal{A}^*$, то $\ker \mathcal{A}^* = \{\theta\}$ и уравнение (111.2) не имеет ненулевого решения. Условие $r < n$ равносильно условию $\dim(\ker \mathcal{A}^*) > 0$, которое означает существование ненулевого вектора $w \in \ker \mathcal{A}^*$, т.е. ненулевого решения (111.2). При этом $\operatorname{im} \mathcal{A} \neq W$ и уравнение (111.1) имеет решение не для любого $u \in W$. ■

Замечание 1. Альтернатива Фредгольма для оператора $\mathcal{A}$, действующего в одном пространстве V, означает, что либо основное уравнение имеет единственное решение при любом $u \in V$, либо сопряженное к нему уравнение имеет нетривиальное решение.

Теорема 111.2 (теорема Фредгольма). *Операторное уравнение (111.1) имеет решение тогда и только тогда, когда его правая часть ортогональна всем решениям сопряженного уравнения.*

Доказательство. Уравнение (111.1) имеет решение тогда и только тогда, когда $u \in \operatorname{im} \mathcal{A}$ или, с учетом (95.4), когда $u \in \ker^\perp \mathcal{A}^*$. Это равносильно ортогональности вектора u всем векторам $\ker \mathcal{A}^*$, т.е. решениям уравнения (111.2). ■

Нормальное решение. Пусть уравнение (111.1) разрешимо, т.е. имеет хотя бы одно решение. Обозначим через H множество всех его решений. *Нормальным решением* уравнения (111.1) называется такое

Уравнения 1–17 называются *каноническими уравнениями алгебраических поверхностей второго порядка*.

Геометрические свойства. Канонические уравнения позволяют исследовать геометрические свойства поверхностей. Чтобы представить форму поверхности, проще всего изучить ее сечения плоскостями, параллельными координатным плоскостям.

Эллипсоид (рис. 1). Поверхность, определяемая уравнением 1, называется *эллипсоидом*. Числа a, b, c в каноническом уравнении называются *полуосями* эллипсоида. Как следует из уравнения, координатные плоскости являются плоскостями симметрии, а начало координат – центром симметрии эллипсоида. Эллипсоид – ограниченная поверхность, заключенная в параллелепипеде $|x| \le a$, $|y| \le b$, $|z| \le c$. Линии пересечения эллипсоида с любой плоскостью являются поэтому ограниченными линиями второго порядка, т.е. эллипсами.

Мнимый эллипсоид. Нет ни одной точки пространства, координаты которой удовлетворяют уравнению 2. Принято говорить об этом уравнении как об уравнении *мнимого эллипсоида*.

Вырожденный эллипсоид. Уравнению 3 удовлетворяет только начало координат. Это уравнение принято называть уравнением *вырожденного эллипсоида* (или *мнимого конуса*).

Однополостный гиперболоид (рис. 2). Поверхность, определяемая уравнением 4, называется *однополостным гиперболоидом*.

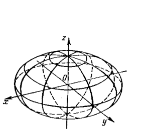


Рис. 1

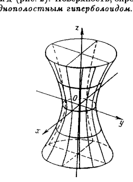


Рис. 2

Как следует из канонического уравнения, координатные плоскости являются плоскостями симметрии, а начало координат – центром симметрии однополостного гиперболоида. Сечения гиперболоида плоскостями $z = h$ представляют собой эллипсы с полуосями $a(1 + h^2/c^2)^{1/2}$, $b(1 + h^2/c^2)^{1/2}$, которые неограниченно возрастают при $h \to \infty$. Эл-

липс, получающийся при $h = 0$, называется *горловым эллипсом* гиперболоида.

Плоскость $y = h$ пересекает гиперболоид:

а) при $|h| < b$ по гиперболе с полуосями $a(1 - h^2/b^2)^{1/2}$, $c(1 - h^2/b^2)^{1/2}$, которые убывают от a и c до нуля при $|h| \to b$;

б) при $|h| = b$ по паре пересекающихся прямых;

в) при $|h| > b$ по гиперболе с полуосями $c(-1 + h^2/b^2)^{1/2}$, $a(-1 + h^2/b^2)^{1/2}$, которые неограниченно возрастают при $|h| \to \infty$.

Сечения гиперболоида плоскостями $x = h$ обладают аналогичными свойствами.

Важной особенностью однополостного гиперболоида является наличие прямых, целиком лежащих на этой поверхности.

Теорема 106.2. *Через каждую точку однополостного гиперболоида проходят две различные прямые, целиком расположенные на этой поверхности.*

Доказательство. Рассмотрим прямые L и L^*, заданные как пересечение плоскостей

$$L: \begin{cases} \alpha\left(\frac{x}{a} - \frac{z}{c}\right) = \beta\left(1 - \frac{y}{b}\right), \\ \beta\left(\frac{x}{a} + \frac{z}{c}\right) = \alpha\left(1 + \frac{y}{b}\right), \\ \alpha^2 + \beta^2 \neq 0 \end{cases} \quad \text{и} \quad L^*: \begin{cases} \gamma\left(\frac{x}{a} - \frac{z}{c}\right) = \delta\left(1 + \frac{y}{b}\right), \\ \delta\left(\frac{x}{a} + \frac{z}{c}\right) = \gamma\left(1 - \frac{y}{b}\right), \\ \gamma^2 + \delta^2 \neq 0. \end{cases}$$

Прямые L и L^* целиком лежат на гиперболоиде (чтобы убедиться в этом, достаточно почленно перемножить уравнения плоскостей). При этом через каждую точку $M_0(x_0, y_0, z_0)$ гиперболоида проходит единственная прямая из семейства L и единственная прямая из семейства L^*: эти прямые (т.е. пары чисел α, β и γ, δ) находятся из однородных систем уравнений

$$\begin{cases} \alpha\left(\frac{x_0}{a} - \frac{z_0}{c}\right) = \beta\left(1 - \frac{y_0}{b}\right), \\ \beta\left(\frac{x_0}{a} + \frac{z_0}{c}\right) = \alpha\left(1 + \frac{y_0}{b}\right), \end{cases} \quad \text{и} \quad \begin{cases} \gamma\left(\frac{x_0}{a} - \frac{z_0}{c}\right) = \delta\left(1 + \frac{y_0}{b}\right), \\ \delta\left(\frac{x_0}{a} + \frac{z_0}{c}\right) = \gamma\left(1 - \frac{y_0}{b}\right), \end{cases}$$

матрицы которых вырождены (т.е. системы имеют нетривиальные решения) и имеют ранг, равный 1 (т.е. все решения каждой из систем пропорциональны и определяют единственную прямую). Остается добавить, что прямые L и L^* не совпадают (достаточно проверить неколлинеарность их направляющих векторов). ■

Прямые, все точки которых лежат на поверхности, называются *прямолинейными образующими* этой поверхности. Итак, однополостный гиперболоид покрыт двумя различными семействами прямолинейных образующих (рис. 3).

Двуполостный гиперболоид (рис. 4). Поверхность, определяемая уравнением 5, называется *двуполостным гиперболоидом*. Как следует из канонического уравнения, координатные плоскости его решение z_0, что

$$\|z_0\|_E = \inf_{z \in H} \|z\|_E.$$

Другими словами, нормальное решение – это решение наименьшей длины. Корректность определения вытекает из следующей теоремы.

Теорема 111.3. *Для любого разрешимого уравнения* (111.1) *нормальное решение существует и единственно.*

Доказательство. Существование. Пусть z – решение уравнения (111.1). Совокупность H всех решений этого уравнения является линейным многообразием $H = z + \ker \mathcal{A}$, так как для множеств H и $z + \ker \mathcal{A}$, как легко показать, имеет место двустороннее вложение. Из свойств линейного многообразия в евклидовом (унитарном) пространстве (теорема 73.1) следует существование и единственность нормального вектора сдвига z_0, ортогонального направляющему подпространству $\ker \mathcal{A}$. Причем вектор z_0 имеет наименьшую длину среди всех решений уравнения (111.1). Таким образом, z_0 – нормальное решение уравнения (111.1).

Единственность. Пусть z_1 – нормальное решение (111.1). Тогда $z_1 \in H$, т.е. $z_1 = z_0 + w$, где $w \in \ker \mathcal{A}$, $(w, z_0) = 0$, и $\|z_1\|_E^2 = \|z_0\|_E^2 + \|w\|_E^2$. Так как $\|z_0\|_E = \|z_1\|_E$, то $w = \theta$ и $z_1 = z_0$. ■

Замечание 2. Доказательство теоремы дает правило для отыскания нормального решения (теорема 73.2): z_0 – перпендикуляр, опущенный из любого решения z уравнения (111.1) на $\ker \mathcal{A}$.

Псевдорешение. Рассматривается уравнение (111.1), не обязательно разрешимое. Вектор $r = \mathcal{A}z - u$ называется ***невязкой*** вектора z, функция $F(z) = \|\mathcal{A}u - z\|_E^2$ – ***функционалом невязки.***

Очевидно, вектор z является решением (111.1) тогда и только тогда, когда его невязка $r = \theta$, т.е. когда $\|r\|_E = 0$. Поскольку нулевое значение нормы является наименьшим, то решение z можно рассматривать как вектор, невязка которого имеет наименьшую норму, или, как принято говорить, ***минимизирующий функционал невязки.***

Задача отыскания векторов, минимизирующих функционал невязки, имеет смысл и тогда, когда уравнение (111.1) неразрешимо (например, вследствие погрешностей измерения $\mathcal{A}$ и u). В этом случае, если z минимизирует функционал невязки, то расстояние $\rho(\mathcal{A}z, u) = \|\mathcal{A}z - u\|_E$ минимально и, следовательно, при таком z левая часть уравнения $\mathcal{A}z$ "ближе" всего к правой части u. Для многих задач вычислительной математики это свойство служит оценкой качества приближенного решения z.

Вектор $z^+ \in V$ называется ***псевдорешением*** уравнения (111.1), если

$$\|\mathcal{A}z^+ - u\|_E^2 = \inf_{z \in V} \|\mathcal{A}z - u\|_E^2. \tag{111.3}$$

Другими словами, псевдорешение – это вектор пространства V, минимизирующий функционал невязки.

Очевидно, если уравнение (111.1) разрешимо, то псевдорешение совпадает с решением в обычном смысле.

Теорема 111.4. *Псевдорешение существует для любого операторного уравнения* (111.1).

Доказательство. Согласно определению (111.3), $\|\mathcal{A}z^+ - u\|_E = \inf_{z \in V} \|\mathcal{A}z - u\|_E = \inf_{z \in V} |\mathcal{A}z - u| = \inf_{z \in V} \rho(\mathcal{A}z, u) = \inf_{y \in \operatorname{im} \mathcal{A}} \rho(y, u)$.

Это означает, что вектор $\mathcal{A}z^+$ – это вектор из $\operatorname{im} \mathcal{A}$, который ближе всего расположен к вектору u. В силу теоремы 74.2 о кратчайшем расстоянии отсюда следует, что $\mathcal{A}z^+$ – ортогональная проекция вектора u на $\operatorname{im} \mathcal{A}$. Пусть $u = g + h$, где $g \in \operatorname{im} \mathcal{A}$, $h \perp \operatorname{im} \mathcal{A}$. Тогда z^+ является решением в обычном смысле уравнения

$$\mathcal{A}z = g \tag{111.4}$$

(очевидно, оно имеет решение, так как $g \in \operatorname{im} \mathcal{A}$). ■

Уравнение

$$\mathcal{A}^*\mathcal{A}z = \mathcal{A}^*u \tag{111.5}$$

называется ***нормальным уравнением*** для уравнения (111.1).

Теорема 111.5. *Вектор z^+ пространства V является псевдорешением уравнения* (111.1) *тогда и только тогда, когда z^+ – решение нормального уравнения* (111.5).

Доказательство. Выше было показано, что z^+ – псевдорешение (111.1) тогда и только тогда, когда z^+ – решение в обычном смысле уравнения (111.4). Покажем, что уравнения (111.4) и (111.5) равносильны. Действительно, $\mathcal{A}^*u = \mathcal{A}^*g$, так как $u = g + h$, где $h \in \operatorname{im}^\perp \mathcal{A} = \ker \mathcal{A}^*$. При этом если z – решение (111.4), то $\mathcal{A}z = g$, поэтому $\mathcal{A}^*\mathcal{A}z = \mathcal{A}^*u$ и z является решением (111.5). Если же z – решение (111.5), то $\mathcal{A}^*\mathcal{A}z = \mathcal{A}^*g$ или $\mathcal{A}^*(\mathcal{A}z - g) = \theta$. Значит, $\mathcal{A}z - g \in \ker \mathcal{A}^* = \operatorname{im}^\perp \mathcal{A}$. Но $\mathcal{A}z \in \operatorname{im} \mathcal{A}$, $g \in \operatorname{im} \mathcal{A}$, следовательно, $\mathcal{A}z - g \in \operatorname{im} \mathcal{A}$. Отсюда с учетом соотношения $\mathcal{A}z - g \in \operatorname{im}^\perp \mathcal{A}$, полученного выше, следует, что $\mathcal{A}z - g = \theta$, т.е. z является решением (111.4). ■

Псевдорешение наименьшей длины называется ***нормальным псевдорешением.*** Из теорем 111.3, 111.4 следует, что нормальное псевдорешение существует и единственно для любого уравнения (111.1).
