§ 8. Линейные пространства

8.1. Вывод свойств линейного пространства из аксиом

Пусть K — поле, V — непустое множество с операцией сложения ($V \times V \to V$, $(a, b) \mapsto a + b$) и операциями умножения на элементы $c \in K$ ($V \to V$, $v \mapsto cv$), удовлетворяющими следующим условиям:

I.1) ассоциативность сложения (т. е. $(u + v) + w = u + (v + w)$ для всех $u, v, w \in V$);

I.2) коммутативность сложения (т. е. $u + v = v + u$ для всех $u, v \in V$);

I.3) существование нейтрального элемента 0 для операции сложения (т. е. $v + 0 = v$ для всех $v \in V$);

I.4) существование противоположного элемента $-v$ для всякого $v \in V$ (т. е. $v + (-v) = 0$);

II.1) $1 \cdot v = v$ для всех $v \in V$;

II.2) $(rs)v = r(sv)$ для всех $r, s \in K$, $v \in V$;

III.1) $r(v_1 + v_2) = rv_1 + rv_2$ для всех $r \in K$, $v_1, v_2 \in V$;

III.2) $(r + s)v = rv + sv$ для всех $r, s \in K$, $v \in V$.

Тогда V называется линейным пространством над полем K (обозначение: $V = {}_KV$).

Приведём вывод ряда следствий из этих аксиом линейного пространства (хотя, конечно, в каждом конкретном случае они достаточно очевидны).

1) Уравнение $u + x = v$ для $u, v \in {}_KV$ имеет, причём единственное, решение $x = (-u) + v$.

Действительно, прибавляя $-u$ к левой и правой части, получаем, что $x = (-u) + v$. С другой стороны, $u + (-u) + v = v$.

2) Если $x + x = x$ для $x \in {}_KV$, то $x = 0$.

Действительно, прибавляя к левой и правой части противоположный элемент $-x$, получаем, что $x = (-x) + x + x = (-x) + x = 0$.

3) $0v = 0$ для любого $v \in {}_KV$.

Действительно, если $x = 0v$ (здесь $0 \in K$), то $x + x = 0v + 0v = (0 + 0)v = 0v = x$, и поэтому $x = 0 \in {}_KV$.

4) $r0 = 0$ для $r \in K$, $0 \in V$.

Действительно, если $x = r0$, то $x + x = r0 + r0 = r(0 + 0) = r0 = x$, и поэтому $x = 0$.

5) $(-1)v = -v$ для всех $v \in V$.

Действительно, $(-1)v + v = (-1 + 1)v = 0v = 0$, т. е. $(-1)v = -v$.

6) $rv = 0$ для $r \in K$, $v \in V$ тогда и только тогда, когда либо $r = 0$, либо $v = 0$.

Действительно, если $r \neq 0$, то в поле K существует элемент $r^{-1} \in K$, и поэтому $v = 1v = r^{-1}rv = r^{-1}0 = 0$.

7) $r(u - v) = ru - rv$ для всех $r \in K$, $u, v \in V$.

Действительно, $r(u - v) + rv = r(u - v + v) = ru$, т. е. $r(u - v) = ru - rv$.

8) $-(-v) = v$ для всех $v \in V$.

Действительно, $v + (-v) = 0$, и поэтому $-(-v) = v$.

8.2. Линейная зависимость в линейных пространствах

Пусть ${}_KV$ — линейное пространство над полем K. Если $v_1, \ldots, v_r \in V$, $k_1, \ldots, k_r \in K$, то элемент

$$k_1v_1 + \ldots + k_rv_r \in V$$

называется *линейной комбинацией* элементов $v_1, \ldots, v_r$ с коэффициентами $k_1, \ldots, k_r \in K$.

Систему элементов $v_1, \ldots, v_r \in {}_KV$ назовём *линейно зависимой*, если найдутся элементы $k_1, \ldots, k_r \in K$ такие, что

а) не все k_i равны нулю (т. е. хотя бы один элемент k_i отличен от нуля);

б) $k_1v_1 + k_2v_2 + \ldots + k_rv_r = 0$.

Для краткости в этой ситуации мы будем говорить, что «*нетривиальная*» линейная комбинация элементов $v_1, \ldots, v_r$ равна нулю (конечно, *тривиальная* линейная комбинация всегда равна нулю, $0v_1 + \ldots + 0v_r = 0$).

Система элементов $v_1, \ldots, v_r \in {}_KV$ называется *линейно независимой*, если она не является линейно зависимой, это означает, что из равенства

$$k_1v_1 + \ldots + k_rv_r = 0, \quad k_1, \ldots, k_r \in K,$$

следует, что
$$k_1 = k_2 = \ldots = k_r = 0.$$

Теорема 8.1. *Система элементов* $v_1, \ldots, v_r \in {}_K V$ *линейно зависима тогда и только тогда, когда для некоторого* i, $1 \leqslant i \leqslant r$,
$$v_i = \sum_{j \neq i} l_j v_j, \quad l_j \in K$$
(*т. е. элемент* v_i *является линейной* **комбинацией** *остальных элементов системы* $v_1, \ldots, v_r$).

Доказательство.

1) Пусть система $v_1, \ldots, v_r$ линейно зависима, т. е.
$$k_1 v_1 + \ldots + k_r v_r = 0, \quad k_i \neq 0.$$
Тогда
$$v_i = \sum_{j \neq i} \frac{(-k_j)}{k_i} v_j.$$

2) Если
$$v_i = \sum_{j \neq i} l_j v_j,$$
то
$$\sum_{j \neq i} l_j v_j + (-1) v_i = v_i + (-1) v_i = 0,$$
т. е. система $v_1, \ldots, v_r$ линейно зависима, поскольку $-1 \neq 0$. □

Пример 8.2. Если в системе элементов $v_1, \ldots, v_r \in {}_K V$ есть нулевой элемент, **скажем,** $\boldsymbol{v_i = 0}$**,** то система $v_1 \ldots, v_r$ линейно зависима.

Действительно, $0v_1 + \ldots + 1v_i + \ldots + 0v_r = 0$, или, другим способом, $v_i = 0 = \sum\limits_{j \neq i} \boldsymbol{0 v_j}$.

Пример 8.3. Если $v_i = v_j$ для $i \neq j$, то система $v_1, \ldots, v_r \in {}_K V$ линейно **зависима.**

Действительно, $0v_1 + \ldots + 1v_i + \ldots + (-1)v_j + \ldots + 0v_r = 0$, или, иначе, $\boldsymbol{v_i = v_j + \sum\limits_{\substack{k \neq i \\ k \neq j}} 0 v_k}$.

Пример 8.4. Система строк $\varepsilon_1, \ldots, \varepsilon_n \in {}_K K^n$, где
$$\begin{gathered}
\varepsilon_1 = (1, 0, \ldots, 0); \\
\varepsilon_2 = (0, 1, \ldots, 0); \\
\ldots \\
\varepsilon_n = (0, 0, \ldots, 1),
\end{gathered}$$
линейно независима. Кроме того, любая строка $\alpha = (k_1, \ldots, k_n) \in {}_K K^n$ является линейной комбинацией элементов $\varepsilon_1, \ldots, \varepsilon_n$, а именно, $\alpha = (k_1, \ldots, k_n) = k_1 \varepsilon_1 + \ldots + k_n \varepsilon_n$.

Действительно,
$$k_1 \varepsilon_1 + \ldots + k_n \varepsilon_n = (k_1, \ldots, k_n),$$
и поэтому если
$$k_1 \varepsilon_1 + \ldots + k_n \varepsilon_n = (0, \ldots, 0),$$
то
$$k_1 = k_2 = \ldots = k_n = 0,$$
следовательно, система строк $\{\varepsilon_1, \ldots, \varepsilon_n\}$ линейно независима.

Пример 8.5. Пусть $v_1, v_2, v_3 \in {}_{\mathbb{R}}V$ — линейно независимая система в линейном пространстве ${}_{\mathbb{R}}V$. Тогда

$$u_1 = v_1 + v_2, \quad u_2 = v_1 + v_3, \quad u_3 = v_2 + v_3 \text{ —}$$

также линейно независимая система.

Действительно, если

$$k_1u_1 + k_2u_2 + k_3u_3 = 0,$$

то

$$0 = k_1(v_1 + v_2) + k_2(v_1 + v_3) + k_3(v_2 + v_3) = (k_1 + k_2)v_1 + (k_1 + k_3)v_2 + (k_2 + k_3)v_3,$$

поэтому

$$\begin{cases} k_1 + k_2 = 0, \\ k_1 + k_3 = 0, \\ k_2 + k_3 = 0. \end{cases}$$

Следовательно, $k_1 = 0$, $k_2 = 0$, $k_3 = 0$, и система элементов u_1, u_2, u_3 линейно независима.

Упражнения 8.6.

1) Подсистема линейно независимой системы линейно независима.

2) Если подсистема линейно зависима, то линейно зависима и вся система.

Замечание 8.7. Для системы строк в K^n

$$\alpha_1 = (a_{11}, \ldots, a_{1n}),$$
$$\ldots$$
$$\alpha_r = (a_{r1}, \ldots, a_{rn})$$

вопрос о её линейной зависимости равносилен существованию ненулевого решения $(k_1, \ldots, k_r)$ следующей однородной системы линейных уравнений:

$$\begin{cases} a_{11}x_1 + \ldots + a_{r1}x_r = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}x_1 + \ldots + a_{rn}x_r = 0 \end{cases}$$

с транспонированной матрицей A^*, где

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rn} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_r \end{pmatrix}.$$

Таким образом, метод Гаусса даёт нам в этом случае алгоритмическое решение задачи о линейной зависимости строк.

Теорема 8.8. *Пусть $A = (a_{ij}) \in \mathrm{M}_n(K)$ — квадратная матрица. Тогда следующие условия равносильны:*

1) $|A| = 0$;

2) *система строк $A_1, \ldots, A_n$ матрицы A линейно зависима (в пространстве строк K^n);*

3) *система столбцов $\hat{A}_1, \ldots, \hat{A}_n$ матрицы A линейно зависима (в пространстве столбцов $\hat{K}^n$).*

Доказательство.

1) Если строки матрицы A линейно зависимы, скажем, i-я строка A_i является линейной комбинацией остальных, $A_i = \sum\limits_{j\neq i} l_j A_j$, то, как мы показали, $|A| = 0$, т. е. 2) $\Longrightarrow$ 1).

2) Пусть $|A| = 0$. Тогда

$$k_1 A_1 + \ldots + k_n A_n = 0$$

в том и только в том случае, если $(k_1, \ldots, k_n)$ является решением однородной системы линейных уравнений с матрицей A^*. Так как $|A^*| = |A| = 0$, то существует ненулевое решение $(k_1, \ldots, k_n)$, т. е. система строк $A_1, \ldots, A_n$ матрицы A линейно зависима. Итак, 1) $\Longrightarrow$ 2).

3) Так как $|A^*| = |A|$, то 1) $\iff$ 3). □

Теорема 8.9. *Любая система из m строк в K^n при $m > n$ линейно зависима.*

Доказательство. Если

$$\alpha_1 = (a_{11}, \ldots, a_{1n}),$$
$$\ldots$$
$$\alpha_m = (a_{m1}, \ldots, a_{mn}),$$

то равенство $k_1\alpha_1 + \ldots + k_m\alpha_m = 0$ равносильно тому, что $(k_1, \ldots, k_m)$ является решением следующей однородной системы линейных уравнений:

$$\begin{cases} a_{11}x_1 + \ldots + a_{m1}x_m = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1n}x_1 + \ldots + a_{mn}x_m = 0. \end{cases}$$

Так как число n уравнений меньше числа m переменных, то однородная система обладает ненулевым решением, т. е. система $\alpha_1, \ldots, \alpha_m$ линейно зависима. □

Следствие 8.10. *Если система $\alpha_1, \ldots, \alpha_r \in K^n$ линейно независима, то $r \leqslant n$.*

Лемма 8.11. *Если система элементов $\alpha_1, \ldots, \alpha_r \in {}_K V$ линейного пространства ${}_K V$ над полем K линейно независима, $\beta \in {}_K V$ и система $\alpha_1, \ldots, \alpha_r, \beta$ линейно зависима, то β является линейной комбинацией элементов $\alpha_1, \ldots, \alpha_r$.*

Доказательство. Пусть

$$k_1\alpha_1 + \ldots + k_r\alpha_r + k_{r+1}\beta = 0, \quad k_1, \ldots, k_{r+1} \in K,$$

где не все k_i, $1 \leqslant i \leqslant r+1$, равны нулю. Если бы $k_{r+1} = 0$, то нетривиальная линейная комбинация $k_1\alpha_1 + \ldots + k_r\alpha_r = 0$, равная нулю, означала бы, что система $\alpha_1, \ldots, \alpha_r$ линейно зависима, что противоречит предположению.

Итак, $k_{r+1} \neq 0$, и поэтому

$$\beta = \frac{-k_1}{k_{r+1}}\alpha_1 + \ldots + \frac{-k_r}{k_{r+1}}\alpha_r. \qquad \square$$

Лемма 8.12 (единственность представления элемента линейного пространства ${}_K V$ в виде линейной комбинации линейно независимой системы элементов). *Пусть $\{\alpha_1, \ldots, \alpha_r\}$ — линейно независимая система элементов линейного пространства ${}_K V$ и*

$$\beta = k_1\alpha_1 + \ldots + k_r\alpha_r = k_1'\alpha_1 + \ldots + k_r'\alpha_r, \quad k_i, k_i' \in K.$$

Тогда $k_1 = k_1', \ldots, k_r = k_r'$.

Доказательство. Действительно,

$$(k_1 - k_1')\alpha_1 + \ldots + (k_r - k_r')\alpha_r = 0,$$

и поэтому $k_1 - k_1' = 0, \ldots, k_r - k_r' = 0$. □

8.3. Максимальные линейно независимые подсистемы систем элементов линейных пространств, базис линейного пространства

Пусть $S \subseteq {}_K V$. Наиболее важные для нас случаи:

а) S — конечное подмножество элементов в ${}_K V$;

б) $S = {}_K V$.

Подсистема $v_1, \ldots, v_r \in S \subseteq {}_K V$ называется *максимальной линейно независимой подсистемой* в S, если:

1) $v_1, \ldots, v_r$ — линейно независимая система;

2) $v_1, \ldots, v_r, v$ — линейно зависимая система для всякого $v \in S$,

или, что эквивалентно,

2′) любой элемент $v \in S$ является линейной комбинацией элементов $v_1, \ldots, v_r$.

Максимальная линейно независимая подсистема $v_1, \ldots, v_r$ в $S = {}_K V$ (если в ${}_K V$ существует такая *конечная* система) называется *базисом* линейного пространства ${}_K V$. Линейное пространство ${}_K V$ с конечным базисом $v_1, \ldots, v_r$ называется *конечномерным линейным пространством* (при этом будет показано, что любой другой базис линейного пространства содержит то же самое число элементов).

Пример 8.13. Как мы уже видели, система строк

$$\begin{aligned} \varepsilon_1 &= (1, 0, \ldots, 0), \\ \varepsilon_2 &= (0, 1, \ldots, 0), \\ &\ldots \\ \varepsilon_n &= (0, 0, \ldots, 1) \end{aligned}$$

является базисом линейного пространства строк K^n.

Лемма 8.14. *Любую линейно независимую подсистему $v_1, \ldots, v_r$ в $S \subseteq K^n$ можно дополнить до максимальной линейно независимой подсистемы в $S \subseteq K^n$.*

Доказательство. Если $v_1, \ldots, v_r$ — максимальная линейно независимая подсистема в $S \subseteq K^n$, то все доказано. Если нет, то найдётся элемент $v \in S$ такой, что $v_1, v_2, \ldots, v_r, v = v_{r+1}$ — линейно независимая подсистема в S. После конечного числа шагов процесс остановится, так как любые системы из $n+1$ элементов в линейном пространстве K^n оказываются линейно зависимыми. □

Следствие 8.15. *Любой ненулевой элемент $0 \neq v \in S \subseteq K^n$ дополняем до максимальной линейно независимой подсистемы в S.*

Следствие 8.16. *В $S = \mathbb{R}^n$ (или $S = K^n$ для бесконечного поля K) бесконечно много различных базисов. Если поле K конечно, $|K| = q$ (например, $K = \mathbb{Z}_2$), то число элементов в K^n равно q^n, и поэтому число базисов в K^n конечно. Найдите их число.*

Замечание 8.17. Пусть строки $a_1, \ldots, a_s \in K^n$ линейно независимы, $s < n$. Тогда существуют такие строки $a_{s+1}, \ldots, a_n \in K^n$, что $\{a_1, \ldots, a_n\}$ — базис линейного пространства K^n. Практическое нахождение строк $a_{s+1}, \ldots, a_n$ можно осуществить следующим образом. Запишем строки $a_1, \ldots, a_s$ по столбцам и приведём полученную матрицу к ступенчатому виду: $\varphi(a_1^*, \ldots, a_s^*) = A_{\text{ступ}}$, где $(a_1^*, \ldots, a_s^*), A_{\text{ступ}} \in \mathrm{M}_{n,s}(K)$, φ — последовательность элементарных преобразований строк. Так как строки $a_1, \ldots, a_s$ линейно независимы, то в $A_{\text{ступ}}$ имеется ровно s ненулевых строк (первые s

строк). Пусть $\hat{b}_{s+1}, \ldots, \hat{b}_n \in \hat{K}^n$ — столбцы, на i-м месте которых стоит 1, а остальные элементы равны 0, $i = s+1, \ldots, n$. Припишем эти столбцы справа к матрице $A_{\text{ступ}}$. Пусть $B \in \mathrm{M}_n(K)$ — полученная матрица. Применяя к матрице B последовательность элементарных преобразований строк, обратную к φ, приходим к матрице $\tilde{B}$. При этом $(\tilde{B})^*$ — матрица, в которой первые s строк — это $a_1, \ldots, a_s$, а последующие строки дополняют их до базиса линейного пространства K^n.

8.4. Замечание о линейной выражаемости конечных систем элементов в линейном пространстве

Пусть ${}_KV$ — линейное пространство, $S_1 \subseteq {}_KV$, $S_2 \subseteq {}_KV$. Будем говорить, что система S_2 элементов $u_1, \ldots, u_s$ *линейно выражается* через систему S_1 элементов $v_1, \ldots, v_r$, если каждый элемент $u_i \in S_2$, $1 \leqslant i \leqslant s$, является линейной комбинацией элементов $v_1, \ldots, v_r$ системы S_1,

$$u_i = \sum_{j=1}^{r} m_{ij} v_j, \quad m_{ij} \in K.$$

Если к тому же система S_3 элементов $w_1, \ldots, w_t$ линейно выражается через систему S_2.

$$w_k = \sum_{i=1}^{s} l_{ki} u_i, \quad l_{ki} \in K, \quad 1 \leqslant k \leqslant t,$$

то

$$w_k = \sum_{i=1}^{s} l_{ki} u_i = \sum_{i=1}^{s} \sum_{j=1}^{r} (l_{ki} m_{ij}) v_j = \sum_{j=1}^{r} \left(\sum_{i=1}^{s} l_{ki} m_{ij} \right) v_j.$$

т. е. система S_3 линейно выражается через систему S_1.

Системы S_1 и S_2 называются *эквивалентными*, если они линейно выражаются друг через друга (обозначение: $S_1 \sim S_2$).

Следствие 8.18. *Отношение «быть эквивалентными системами», $S_1 \sim S_2$, является отношением эквивалентности.*

Следствие 8.19. *Если элемент $v \in {}_KV$ является линейной комбинацией элементов $v_1, \ldots, v_r$ системы S_1, $S_1 \sim S_2$, где S_2 — система элементов $u_1, \ldots, u_s$, то элемент v является линейной комбинацией элементов $u_1, \ldots, u_s$ системы S_2.*

Следствие 8.20. *Любая (конечная) система элементов $S \subseteq {}_KV$ эквивалентна своей максимальной линейно независимой подсистеме.*

Следствие 8.21. *Любые две (конечные) максимально независимые подсистемы любой системы $S \subseteq {}_KV$ эквивалентны.*

Замечание 8.22. Если $A, B \in \mathrm{M}_{m,n}(K)$ и матрица B получена из матрицы A конечным числом элементарных преобразований 1-го, 2-го и 3-го типов, то каждая строка матрицы B является линейной комбинацией строк матрицы A (поскольку от матрицы B мы можем вернуться к матрице A с помощью элементарных преобразований строк 1-го, 2-го и 3-го типов, то каждая строка матрицы A является линейной комбинацией строк матрицы B). Таким образом, в линейном пространстве строк K^n системы строк $A_1, \ldots, A_m$ матрицы A и $B_1, \ldots, B_m$ матрицы B линейно выражаются друг через друга.

Теорема 8.23 (основная теорема о линейной зависимости). *Пусть в линейном пространстве ${}_KV$ линейно независимая система элементов $v_1, \ldots, v_r$ линейно выражается через другую систему элементов $u_1, \ldots, u_s$. Тогда $r \leqslant s$.*

Доказательство. Допустим противное: пусть $r > s$. В силу нашего предположения

$$v_1 = a_{11}u_1 + \ldots + a_{1s}u_s,$$
$$\ldots$$
$$v_r = a_{r1}u_1 + \ldots + a_{rs}u_s, \quad a_{ij} \in K.$$

Так как $r > s$, то r строк

$$(a_{11}, \ldots, a_{1s}),$$
$$\ldots$$
$$(a_{r1}, \ldots, a_{rs})$$

в линейном пространстве строк K^s линейно зависимы: найдётся их линейная комбинация с коэффициентами $k_1, \ldots, k_r$, где $k_i \neq 0$ для некоторого i, равная нулевой строке $(0, \ldots 0) \in K^s$. Но тогда и линейная комбинация элементов $v_1, \ldots, v_r$ с этими же коэффициентами $k_1, \ldots, k_r$, равна нулю, $k_1v_1 + \ldots + k_rv_r = 0$. Таким образом, система элементов $v_1, \ldots, v_r$ линейно зависима, что приводит нас к противоречию. □

Следствие 8.24. *Две эквивалентные конечные линейно независимые системы в линейном пространстве $_KV$ содержат равное число элементов.*

Следствие 8.25. *Для системы $S \subseteq {_KV}$, где $_KV$ — конечномерное линейное пространство, любые две (конечные) максимальные линейно независимые подсистемы содержат одинаковое число элементов* $\mathrm{r}(S)$, *называемое* рангом *системы S.*

Следствие 8.26. *Если $S = {_KV}$ и $_KV$ — конечномерное линейное пространство, то любые два базиса в $_KV$ состоят из одного и того же числа элементов n, это число n называется* размерностью *линейного пространства $_KV$, обозначение:* $\dim_K V = n$.

Как мы видели ранее, одним из базисов в линейном пространстве строк $_KK^n$ является система строк

$$\varepsilon_1 = (1, 0, \ldots, 0),$$
$$\ldots$$
$$\varepsilon_n = (0, 0, \ldots, 1),$$

и поэтому $\dim_K K^n = n$.

Следствие 8.27. *Если в конечномерном линейном пространстве $_KV$ одна система элементов S_1 линейно выражается через другую систему S_2, то $\mathrm{r}(S_1) \leqslant \mathrm{r}(S_2)$.*

Следствие 8.28. *Если в линейном пространстве $_KV$ система M из m элементов имеет ранг r, то любая её подсистема S из s элементов ($s \leqslant m$) имеет ранг не меньше чем $r + s - m$.*

Доказательство. Действительно, если R — максимальная линейно независимая подсистема в M, $|R| = r$, то $R \setminus (R \cap S) \subset M \setminus S$, и поэтому $|R \setminus (R \cap S)| \leqslant m - s$. Следовательно, $|R \cap S| \geqslant r - (m - s) = r + s - m$. □

Следствие 8.29. *Для системы строк $v_1, \ldots, v_r \in K^n$ следующие условия эквивалентны:*

1) *система строк $v_1, \ldots, v_r$ является базисом линейного пространства строк K^n (т. е. максимальной линейно независимой подсистемой строк в K^n; и тогда $r = n$);*

2) *каждая строка* $v \in K^n$ *единственным образом представляется в виде линейной комбинации*

$$v = \lambda_1 v_1 + \ldots + \lambda_r v_r, \quad \lambda_1, \ldots, \lambda_r \in K$$

(*и тогда* $r = n$);

3) $r = n$ *и система строк* $v_1, \ldots, v_n$ *линейно независима*;

4) $r = n$ *и каждая строка* $v \in K^n$ *представима в виде линейной комбинации*

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n, \quad \lambda_1, \ldots, \lambda_n \in K.$$

Доказательство. Мы уже показали, что 1) $\Longrightarrow$ 2). Покажем, что 2) $\Longrightarrow$ 1). Если $v_1, \ldots, v_r$ — линейно зависимая система строк, $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$ с некоторым $\lambda_i \neq 0$, то нулевая строка имеет два различных представления

$$0 = 0 \cdot v_1 + \ldots + 0 \cdot v_r = \lambda_1 v_1 + \ldots + \lambda_r v_r, \quad \lambda_i \neq 0.$$

При этом $r = n$, так как любые базисы в K^n содержат n элементов.

Ясно, что 1) $\Longrightarrow$ 3). Покажем, что 3) $\Longrightarrow$ 1). Для любой строки $v \in K^n$ система строк $v_1, \ldots, v_n, v$ линейно зависима $(n + 1 > n)$. Так как $v_1, \ldots, v_n$ — линейно независимая система, то $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$ для некоторых $\lambda_1, \ldots, \lambda_n \in K$.

Ясно, что 1) $\Longrightarrow$ 4). Покажем, что 4) $\Longrightarrow$ 1). Допустим, что $v_1, \ldots, v_n$ — линейно зависимая система. Тогда её максимально линейно независимая подсистема $v_{i_1}, \ldots, v_{ir}$, $r < n$, является максимальной линейно независимой подсистемой в K^n, что противоречит $r = n$. □

8.5. Изоморфизм линейных пространств

Пусть ${}_KU$, ${}_KV$ — линейные пространства над полем K. Биективное отображение

$$f\colon {}_KU \to {}_KV,$$

для которого

$$f(u_1 + u_2) = f(u_1) + f(u_2),$$
$$f(ku) = kf(u)$$

для всех $u_1, u_2, u \in {}_KU$, $k \in K$, называется *изоморфизмом* линейных пространств ${}_KU$ и ${}_KV$ (в этом случае будем говорить, что линейные пространства ${}_KU$ и ${}_KV$ *изоморфны*, обозначение: ${}_KU \cong {}_KV$).

Упражнение 8.30. Отношение ${}_KU \cong {}_KV$ является отношением эквивалентности.

Лемма 8.31. *Если* $f\colon {}_KU \to {}_KV$ — *изоморфизм линейных пространств,* $\dim {}_KU = n$, $\{e_1, \ldots, e_n\}$ — *базис в* ${}_KU$, *то* $\{f(e_1), \ldots, f(e_n)\}$ — *базис в* ${}_KV$, *и поэтому* $\dim {}_KV = n = \dim {}_KU$.

Доказательство.

1) Если $v \in {}_KV$, то $f(u) = v$ для некоторого $u \in {}_KU$. Пусть $u = k_1 e_1 + \ldots + k_n e_n$, где $k_1, \ldots, k_n \in K$. Тогда

$$v = f(u) = k_1 f(e_1) + \ldots + k_n f(e_n).$$

2) Пусть $k_1 f(e_1) + \ldots + k_n f(e_n) = 0$ для $k_1, \ldots, k_n \in K$. Тогда

$$0 = k_1 f(e_1) + \ldots + k_n f(e_n) = f(k_1 e_1 + \ldots + k_n e_n),$$

и поэтому

$$k_1 e_1 + \ldots + k_n e_n = 0,$$

следовательно, $k_1 = k_2 = \ldots = k_n = 0$.

Итак, в силу 1) и 2), $\{f(e_1), \ldots, f(e_n)\}$ — базис линейного пространства ${}_KV$. □

Лемма 8.32. *Если* $\dim_K V = n$ *и* $\{e_1, \ldots, e_n\}$ — *базис линейного пространства* ${}_K V$, *то, сопоставляя каждому элементу* $v = k_1 e_1 + \ldots + k_n e_n \in {}_K V$ *однозначно определённую строчку его координат* $(k_1, \ldots, k_n)$ *в базисе* $\{e_1, \ldots, e_n\}$, *получаем изоморфизм линейных пространств* ${}_K V \cong K^n$, *таким образом, каждое* n*-мерное линейное пространство* ${}_K V$ *над полем* K *изоморфно линейному пространству строк* K^n.

Доказательство. Соответствие

$$\Delta \colon {}_K V \in v = k_1 e_1 + \ldots + k_n e_n \mapsto (k_1, \ldots, k_n) \in K^n$$

является биекцией, для которой

$$\begin{aligned}
\Delta(v + v') &= \Delta\big((k_1 e_1 + \ldots + k_n e_n) + (k_1' e_1 + \ldots + k_n' e_n)\big) = \\
&= \Delta\big((k_1 + k_1')e_1 + \ldots + (k_n + k_n')e_n\big) = \\
&= (k_1 + k_1', \ldots, k_n + k_n') = (k_1, \ldots, k_n) + (k_1', \ldots, k_n') = \\
&= \Delta(v) + \Delta(v');\\
\Delta(kv) &= \Delta\big(k(k_1 e_1 + \ldots + k_n e_n)\big) = \Delta\big((kk_1)e_1 + \ldots + (kk_n)e_n\big) = \\
&= (kk_1, \ldots, kk_n) = k(k_1, \ldots, k_n) = k\Delta(v).
\end{aligned}$$

$\square$

Теорема 8.33. *Конечномерные линейные пространства* ${}_K U$ *и* ${}_K V$ *изоморфны тогда и только тогда, когда* $\dim_K U = \dim_K V = n$, *и в этом случае* ${}_K U \cong K^n \cong {}_K V$.

Доказательство теоремы следует из лемм 8.31 и 8.32. $\square$

Упражнение 8.34. Покажите, что следующие линейные пространства являются бесконечномерными линейными пространствами (это означает, что в них нет базиса из конечного числа элементов):

1) ${}_{\mathbb{R}}C[0,1]$ — линейное пространство вещественных непрерывных функций на отрезке $[0,1]$;
2) ${}_K K[x]$ — линейное пространство многочленов от переменной x с коэффициентами из поля K;
3) ${}_K K^{\mathbb{N}}$ — линейное пространство всех счётных последовательностей $(k_1, k_2, \ldots, k_n, \ldots)$ элементов из поля K.

Упражнение 8.35. Докажите, что

а) $\dim_K \mathrm{M}_{m,n}(K) = mn$;

б) $\dim_{\mathbb{R}} \{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = A\} = \dfrac{n(n+1)}{2}$;

в) $\dim_{\mathbb{R}} \{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = -A\} = \dfrac{n(n-1)}{2}$.

8.6. Замена базиса линейного пространства и преобразование координат

Пусть V — конечномерное линейное пространство над полем K, $\dim V = n < \infty$, $\{v_1, \ldots, v_n\}$ — базис в V, $\{v_1', \ldots, v_n'\}$ — другой базис в V,

$$v_j' = c_{1j} v_1 + c_{2j} v_2 + \ldots + c_{nj} v_n, \quad j = 1, \ldots, n, \quad c_{ij} \in K$$

(запись по столбцу!). $C = (c_{ij}) \in \mathrm{M}_n(K)$ — *матрица перехода* от первого базиса ко второму.

Замечание 8.36. Так как умножение в поле K коммутативно, то левое линейное пространство ${}_KV$ можно рассматривать и как правое линейное пространство V_K, полагая $v\lambda = \lambda v$ для всех $\lambda \in K$, $v \in V$. Тогда определение матрицы перехода может быть записано в матричном виде как

$$(v'_1, \ldots, v'_n) = (v_1, \ldots, v_n)C.$$

Ограничиваясь левыми линейными пространствами, мы можем использовать эквивалентную форму записи:

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

или, кратко, $\mathcal{E}' = C^*\mathcal{E}$, где

$$\mathcal{E} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \ \mathcal{E}' = \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} \in \mathrm{M}_{n,1}(V).$$

Если $V = K^n$, то $v_1, \ldots, v_n, v'_1, \ldots, v'_n \in K^n$, $\mathcal{E}, \mathcal{E}' \in \mathrm{M}_n(K)$ и $\mathcal{E}' = C^*\mathcal{E}$ означает равенство квадратных $(n \times n)$-матриц.

8.7. Обратимость матрицы перехода

1) Если $|C| = 0$, то $|C^*| = 0$ и строки матрицы C^* линейно зависимы. Поэтому из

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad \text{т. е.} \quad \mathcal{E}' = C^*\mathcal{E},$$

следует, что $v'_1, \ldots, v'_n$ — линейно зависимая система в V, что приводит к противоречию с тем, что $v'_1, \ldots, v'_n$ — базис. Итак, мы показали, что $|C| \neq 0$ и существует обратная матрица C^{-1} (тогда $(C^*)^{-1} = (C^{-1})^*$).

2) Другое доказательство обратимости матрицы C даёт интерпретация матрицы $B = C^{-1}$ как матрицы перехода от второго базиса к первому.

Действительно, элементы $v_1, \ldots, v_n$ также выражаются как линейные комбинации элементов базиса $\{v'_1, \ldots, v'_n\}$:

$$v_i = b_{1i}v'_1 + \ldots + b_{ni}v'_n, \quad i = 1, \ldots, n, \quad b_{ij} \in K,$$

$B = (b_{ij}) \in \mathrm{M}_n(K)$. Тогда $\mathcal{E} = B^*\mathcal{E}'$. Так как $\mathcal{E}' = C^*\mathcal{E}$, то

$$\mathcal{E} = B^*(C^*\mathcal{E}) = (B^*C^*)\mathcal{E} = (CB)^*\mathcal{E}.$$

Так как $\{v_1, \ldots, v_n\}$ — базис в V, то $(CB)^* = E$, следовательно, $CB = E$, и поэтому $B = C^{-1}$. □

3) Для любой обратимой матрицы $C \in \mathrm{M}_n(K)$, $|C| \neq 0$, и любого базиса $\{v_1, \ldots, v_n\}$ конечномерного линейного пространства ${}_KV$, $\dim_K V = n$, элементы $v'_1, \ldots, v'_n \in {}_KV$, где

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

образуют базис линейного пространства ${}_KV$.

Действительно, в этом случае

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (C^*)^{-1} \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix}, \quad (C^*)^{-1} = (C^{-1})^*,$$

т. е. n линейно независимых элементов $v_1, \ldots, v_n$ линейно выражаются через $v'_1, \ldots, v'_n$. По основной лемме о линейной зависимости элементы $v'_1, \ldots, v'_n$ линейно независимы. Так как $\dim_K V = n$, то $\{v'_1, \ldots, v'_n\}$ — базис линейного пространства $_K V$. □

8.8. Замена координат элемента линейного пространства при замене базиса

Пусть $\{v_1, \ldots, v_n\}$, $\{v'_1, \ldots, v'_n\}$ — два базиса линейного пространства $_K V$, $\dim_K V = n$, $C \in \mathrm{M}_n(K)$, $|C| \neq 0$, — матрица перехода от первого базиса ко второму,

$$\begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix},$$

$x = x_1 v_1 + \ldots + x_n v_n = x'_1 v'_1 + \ldots + x'_n v'_n \in {}_K V$. Так как

$$x = x_1 v_1 + \ldots + x_n v_n = (x_1, \ldots, x_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = (x_1, \ldots, x_n)(C^{-1})^* \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix} = (x'_1, \ldots, x'_n) \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix},$$

то

$$(x'_1, \ldots, x'_n) = (x_1, \ldots, x_n)(C^{-1})^*,$$

или

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = C^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

что эквивалентно

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$ □

Пример 8.37. Пусть $V = \mathbb{R}^3$, $v_1 = (2, 1, -3)$, $v_2 = (3, 2, -5)$, $v_3 = (1, -1, 1)$. Необходимо выяснить, образуют ли элементы v_1, v_2, v_3 базис в $\mathbb{R}^3$, и если да, то найти координаты строки $x = (6, 2, -7)$ в базисе $\{v_1, v_2, v_3\}$.

Решение.

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = C^* \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix},$$

где $\{e_1, e_2, e_3\}$ — стандартный базис в $\mathbb{R}^3$,

$$C = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & -1 \\ -3 & -5 & 1 \end{pmatrix}.$$

Строки v_1, v_2, v_3 образуют базис в $\mathbb{R}^3$ тогда и только тогда, когда матрица C обратима. Если матрица C обратима, то столбец координат строки x в базисе $\{v_1, v_2, v_3\}$ равен

$$C^{-1} \begin{pmatrix} 6 \\ 2 \\ -7 \end{pmatrix}.$$

Для вычисления этого столбца применим алгоритм вычисления матрицы $A^{-1}B$ (см. с. 66), в процессе работы которого проверяется, обратима ли матрица $A = C$:

$$\left(\begin{array}{rrr|r} 2 & 3 & 1 & 6 \\ 1 & 2 & -1 & 2 \\ -3 & -5 & 1 & -7 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 2 & 3 & 1 & 6 \\ -3 & -5 & 1 & -7 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & -1 & 3 & 2 \\ -3 & -5 & 1 & -7 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & 1 & -3 & -2 \\ 0 & 1 & -2 & -1 \end{array}\right) \to$$

$$\to \left(\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & -1 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 2 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Таким образом, матрица C обратима, $(1,1,1)$ — координаты строки x в базисе $\{v_1, v_2, v_3\}$, $x = v_1 + v_2 + v_3$.

Этот же результат можно было получить, используя формулу $(6,2,-7)(C^*)^{-1} = (1,1,1)$,

$$\left(\begin{array}{rrr} 2 & 1 & -3 \\ 3 & 2 & -5 \\ 1 & -1 & 1 \\ \hline 6 & 2 & -7 \end{array}\right) \to \left(\begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 1 & 1 \end{array}\right)$$

(здесь применяем элементарные преобразования столбцов).

8.9. Линейные подпространства линейных пространств

Пусть K — поле, ${}_K V$ — линейное пространство над полем K. Непустое подмножество $\varnothing \neq U \subseteq {}_K V$ называется *линейным подпространством* линейного пространства ${}_K V$, если:

1) $u_1 + u_2 \in U$ для всех $u_1, u_2 \in U$;

2) $ku \in U$ для всех $k \in K$, $u \in U$.

Ясно, что ${}_K U$ — линейное пространство относительно тех же операций сложения элементов и умножения на элементы из поля K, что и в линейном пространстве ${}_K V$.

Если U — линейное подпространство в конечномерном линейном пространстве ${}_K V$, $n = \dim_K V < \infty$, то $\dim_K U \leqslant \dim_K V$. Действительно, если элементы $u_1, \ldots, u_s \in {}_K U$ линейно независимы в ${}_K U$, то эти элементы линейно независимы и в линейном пространстве ${}_K V$, $s \leqslant n$, поэтому $\dim_K U \leqslant \dim_K V$.

Если ${}_K U$ — линейное подпространство линейного пространства ${}_K V$, ${}_K U \subseteq {}_K V$ и $\dim_K U = \dim_K V = n$, то ${}_K U = {}_K V$. Действительно, если $\{u_1, \ldots, u_n\}$ — базис линейного пространства ${}_K U \subseteq {}_K V$, то эти n элементов $u_1, \ldots, u_n$ линейно независимы в ${}_K V$ и $\dim_K V = n$, поэтому $\{u_1, \ldots, u_n\}$ — базис линейного пространства ${}_K V$. Итак, каждый элемент $v \in V$ имеет вид $v = k_1 u_1 + \ldots + k_n u_n \in {}_K U$, $k_i \in K$, т. е. ${}_K V = {}_K U$.

8.10. Пересечение линейных подпространств

Лемма 8.38. *Пересечение*

$$U = \bigcap_{i \in I} U_i$$

любого семейства линейных подпространств $\{U_i \subset {}_K V \mid i \in I\}$ линейного пространства ${}_K V$ является линейным подпространством.

Доказательство. Если $u, u_1, u_2 \in U = \bigcap\limits_{i \in I} U_i$, $k \in K$, то $u, u_1, u_2 \in U_i$ для любого $i \in I$, поэтому $u_1 + u_2, ku \in U_i$ для любого $i \in I$, т. е. $u_1 + u_2, ku \in U = \bigcap\limits_{i \in I} U_i$. □

Следствие 8.39. *Если U_1 и U_2 — линейные подпространства линейного пространства ${}_K V$, то $U_1 \cap U_2$ — линейное подпространство в ${}_K V$ (наибольшее подпространство среди подпространств, лежащих одновременно в U_1 и в U_2).*

8.11. Сумма линейных подпространств

Если U_1 и U_2 — линейные подпространства линейного пространства ${}_K V$, то *сумма линейных подпространств*

$$U_1 + U_2 = \{u_1 + u_2 \mid u_1 \in U_1,\ u_2 \in U_2\}$$

также является линейным подпространством. Действительно, если $u_1 + u_2, u_1' + u_2' \in U_1 + U_2$, $u_1, u_1' \in U_1$, $u_2, u_2' \in U_2$, $k \in K$, то

$$(u_1 + u_2) + (u_1' + u_2') = (u_1 + u_1') + (u_2 + u_2') \in U_1 + U_2;$$
$$k(u_1 + u_2) = ku_1 + ku_2 \in U_1 + U_2.$$
□

Замечание 8.40. $U_1 + U_2$ — наименьшее линейное подпространство среди линейных подпространств, содержащих одновременно U_1 и U_2. Более того,

$$U_1 + U_2 = \bigcap_{\substack{U \subseteq_K V \\ U_1 \subseteq U,\ U_2 \subseteq U}} U.$$

Замечание 8.41. Если U, U_1, U_2, U_3 — линейные подпространства в ${}_K V$, то

$$U \cap U = U, \quad U + U = U,$$
$$U_1 \cap U_2 = U_2 \cap U_1, \quad U_1 + U_2 = U_2 + U_1,$$
$$U_1 \cap (U_2 \cap U_3) = (U_1 \cap U_2) \cap U_3,$$
$$U_1 + (U_2 + U_3) = (U_1 + U_2) + U_3,$$
$$U_1 \cap (U_1 + U_2) = U_1, \quad U_1 + (U_1 \cap U_2) = U_1.$$

8.12. Линейная оболочка элементов линейного пространства

Пусть ${}_K V$ — линейное пространство, $v_1, \ldots, v_m \in {}_K V$. Рассмотрим

$$\langle v_1, \ldots, v_m \rangle = \{k_1 v_1 + \ldots + k_m v_m \mid k_1, \ldots, k_m \in K\} -$$

совокупность всех линейных комбинаций $k_1 v_1 + \ldots + k_m v_m$ элементов $v_1, \ldots, v_m$ с коэффициентами $k_1, \ldots, k_m \in K$, называемую *линейной оболочкой* элементов $v_1, \ldots, v_m$. Линейная оболочка $\langle v_1, \ldots, v_m \rangle$ является наименьшим линейным подпространством, содержащим элементы $v_1, \ldots, v_m$. Действительно,

$$(k_1 v_1 + \ldots + k_m v_m) + (l_1 v_1 + \ldots + l_m v_m) = (k_1 + l_1) v_1 + \ldots + (k_m + l_m) v_m;$$

$$k(k_1 v_1 + \ldots + k_m v_m) = (kk_1) v_1 + \ldots + (kk_m) v_m;$$

если U — линейное подпространство в ${}_K V$, $v_1, \ldots, v_m \in U$, то $k_1 v_1 + \ldots + k_m v_m \in U$, следовательно, $\langle v_1, \ldots, v_m \rangle \subseteq U$. Более того,

$$\langle v_1, \ldots, v_m \rangle = \bigcap_{\substack{U \subseteq_K V \\ v_1, \ldots, v_m \in U}} U.$$

Замечание 8.42. Если $0 \neq v \in {}_KV$, то $\langle v\rangle = Kv = \{kv \mid k \in K\}$, $\dim\langle v\rangle = 1$; если $v = 0$, $\langle v\rangle = Kv = \{0\}$.

Замечание 8.43. $\langle v_1, \ldots, v_m\rangle = Kv_1 + \ldots + Kv_m$.

Замечание 8.44. $\dim_K\langle v_1, \ldots, v_m\rangle = \mathrm{r}\{v_1, \ldots, v_m\}$; любая максимальная линейно независимая подсистема в $\{v_1, \ldots, v_m\}$ является базисом линейного подпространства $\langle v_1, \ldots, v_m\rangle$.

Основная лемма о линейной зависимости может быть сформулирована в следующей эквивалентной форме.

Теорема 8.45 (о замене). *Пусть $v_1, \ldots, v_s \in {}_KV$ — линейно независимая система, $u_1, \ldots, u_r \in \langle v_1, \ldots, v_s\rangle$, $\{u_1, \ldots, u_r\}$ — линейно независимая система элементов. Тогда $r \leqslant s$ и*

$$\langle v_1, \ldots, v_s\rangle = \langle u_1, \ldots, u_r, v_{i_{r+1}}, \ldots, v_{i_s}\rangle,$$

где

$$1 \leqslant i_{r+1} < \ldots < i_s \leqslant s.$$

Доказательство. Так как $s = \dim_K\langle v_1, \ldots, v_s\rangle$, то $r \leqslant s$. Если $r = s$, то $\langle v_1, \ldots, v_s\rangle = \langle u_1, \ldots, u_r\rangle$. Если $r < s$, то найдётся $v_{i_{r+1}} \notin \langle u_1, \ldots, u_r\rangle$ (индекс i_{r+1} — минимальный с этим свойством). Продолжая этот процесс, построим базис $\{u_1, \ldots, u_r, v_{i_{r+1}}, \ldots, v_{i_s}\}$ в $\langle v_1, \ldots, v_s\rangle$. $\square$

Следствие 8.46. *Пусть U, W — линейные подпространства в ${}_KV$ и $U \subseteq W$, $\dim_K U = l$, $\dim_K W = m$. Тогда $l \leqslant m$ и любой базис подпространства U можно дополнить $m - l$ элементами до базиса подпространства W. В частности, если $U \subseteq W$ и $l = m$, то $U = W$.*

Теорема 8.47 (формула размерности). *Пусть U, W — линейные подпространства в ${}_KV$, $\dim_K V = n < \infty$. Тогда*

$$\dim_K U + \dim_K W = \dim_K(U \cap W) + \dim_K(U + W),$$

или, что эквивалентно,

$$\dim_K(U + W) = \dim_K U + \dim_K W - \dim_K(U \cap W).$$

Доказательство. Пусть $\dim_K(U \cap W) = d$, $\dim_K U = s$, $\dim_K W = t$. Ясно, что $0 \leqslant d \leqslant s$, $d \leqslant t$. При $d = 0$ утверждение очевидно (объединение базисов в U и W даёт базис в $U + W$). Выберем базис $v_1, \ldots, v_d$ линейного пространства $U \cap W$ и дополним его до базиса $v_1, \ldots, v_d, u_1, \ldots, u_{s-d}$ линейного пространства U и до базиса $v_1, \ldots, v_d, w_1, \ldots, w_{t-d}$ линейного пространства W. Ясно, что

$$U + W = \langle v_1, \ldots, v_d, u_1, \ldots, u_{s-d}, w_1, \ldots, w_{t-d}\rangle.$$

Если

$$\lambda_1 v_1 + \ldots + \lambda_d v_d + \mu_1 u_1 + \ldots + \mu_{s-d} u_{s-d} + \gamma_1 w_1 + \ldots + \gamma_{t-d} w_{t-d} = 0,$$

то

$$\sum_{i=1}^{d} \lambda_i v_i + \sum_{j=1}^{s-d} \mu_j u_j = -\sum_{k=1}^{t-d} \gamma_k w_k \in U \cap W,$$

поэтому $\mu_1 = \ldots = \mu_{s-d} = 0$, $\gamma_1 = \ldots = \gamma_{t-d} = 0$. Следовательно, $\lambda_1 = \ldots = \lambda_d = 0$. Таким образом,

$$\{v_1, \ldots, v_d, u_1, \ldots, u_{s-d}, w_1, \ldots, w_{t-d}\} \text{ —}$$

базис линейного подпространства $U + W$, откуда

$$s + t = d + (s - d) + d + (t - d) = d + \big(d + (s - d) + (t - d)\big),$$

поэтому

$$\dim_K U + \dim_K W = \dim_K U \cap W + \dim_K(U + W).$$ $\square$

Теорема 8.48 (о существовании прямого дополнения подпространства). *Пусть* $\dim_K V = n < \infty$, U *— линейное подпространство в* $_KV$. *Тогда существует линейное подпространство* W *в* $_KV$ *такое, что*

$$U + W = V, \quad U \cap W = \{0\},$$

(*называемое прямым дополнением подпространства* U *в* $_KV$; *в этом случае также говорят, что линейное пространство* $_KV$ *является прямой суммой линейных подпространств* U *и* W, *обозначение*: $_KV = U \oplus W$).

Доказательство. Если $\dim_K U = r$ и $\{u_1, \ldots, u_r\}$ — базис в $_KU$, то дополним его до базиса линейного пространства $_KV$: $u_1, \ldots, u_r, v_1, \ldots, v_{n-r}$. Пусть $W = \langle v_1, \ldots, v_{n-r}\rangle$. Тогда $_KV = U + W$, $U \cap W = \{0\}$. □

Замечание 8.49. Конечно, прямое дополнение определено неоднозначно, однако все прямые дополнения линейного пространства изоморфны (а именно, все они имеют размерность $\dim_K V - \dim_K U$).

Замечание 8.50. Если $_KV = U \oplus W$, то представление элемента $v \in V$ в виде $v = u - w$, $u \in U$, $w \in W$, определено однозначно (действительно, если $v = u + w = u' + w'$, $u' \in U$, $w' \in W$, то $u - u' = w' - w \in U \cap W = \{0\}$, следовательно, $u = u'$, $w = w'$), и поэтому линейное пространство $_KV = U \oplus W$ изоморфно *внешней прямой* сумме $\{(u, w) \mid u \in U,\ w \in W\}$ линейных пространств $_KU$ и $_KW$ с естественными операциями сложения пар и их умножения на $c \in K$.

Пример 8.51 (прямого разложения). Пусть

$$V = \mathrm{M}_n(\mathbb{R}), \quad U = \{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = A\},$$
$$W = \{A \in \mathrm{M}_n(\mathbb{R}) \mid A^* = -A\}.$$

Тогда

$$_{\mathbb{R}}V = U \oplus W.$$

Действительно, $A = \dfrac{A + A^*}{2} + \dfrac{A - A^*}{2}$. Если $A = A^* = -A$, то $A = 0 \in \mathrm{M}_n(\mathbb{R})$. □

8.13. Линейные отображения линейных пространств

Пусть K — поле, $_KU$, $_KV$ — линейные пространства над полем K. Отображение $f\colon {_KU} \to {_KV}$ называется *линейным отображением*, если

$$f(x + y) = f(x) + f(y), \quad f(\alpha x) = \alpha x$$

для всех $x, y \in U$, $\alpha \in K$.

Если $_KU = {_KV}$ и $f\colon {_KU} \to {_KV}$ — линейное отображение, то отображение f называется линейным оператором на линейном пространстве $_KU$. Если $V = K$, то линейное отображение $f\colon {_KU} \to {_KK}$ называется линейным функционалом на линейном пространстве $_KU$.

Введём операцию сложения для линейных отображений $f, g\colon U \to V$:

$$(f + g)(x) = f(x) + g(x)$$

для всех $x \in U$. Для $\alpha \in K$ положим

$$(\alpha f)(x) = \alpha f(x)$$

для всех $x \in U$. С этими операциями множество всех линейных отображений $U \to V$ становится линейным пространством (обозначение: $\mathrm{Hom}_K(U, V)$).

Пример 8.52. Пусть $U = \hat{K}_m$, $V = \hat{K}_n$, $A \in \mathrm{M}_{n,m}(K)$. Для $\hat{X} \in \hat{K}_m$ положим $\hat{Y} = A\hat{X}$. Тогда так определённое отображение ${}_KU \to {}_KV$ является линейным отображением.

Примеры 8.53 (линейных функционалов).

1) $K = \mathbb{R}$, ${}_{\mathbb{R}}V = C[0,1]$, $t_0 \in [0,1]$, для $g = g(t) \in C[0,1]$ положим $f(g) = g(t_0)$.

2) $K = \mathbb{R}$, ${}_{\mathbb{R}}V = C[0,1]$. Для $g = g(t) \in C[0,1]$ положим $f(g) = \int\limits_0^1 g(t)\,dt$.

3) $V = K^n$, $\alpha_1, \ldots, \alpha_n \in K$. Для $x = (x_1, \ldots, x_n) \in K^n$ положим $f(x) = \alpha_1 x_1 + \ldots + \alpha_n x_n$.

8.14. Сопряжённое пространство

Пусть ${}_KV$ — линейное пространство над полем K, ${}_KV^*$ — множество всех **линейных функционалов** на V. Для $f, g \in V^*$ имеем

$$(f+g)(x) = f(x) + g(x) \quad \text{для всех} \quad x \in V,$$
$$(\alpha f)(x) = \alpha f(x) \quad \text{для всех} \quad \alpha \in K, \quad x \in V.$$

С этими операциями сложения и умножения на элементы поля K множество ${}_KV^*$ превращается в линейное пространство над полем K. Линейное пространство ${}_KV^*$ называется сопряжённым пространством к линейному пространству ${}_KV$.

Пусть $\dim V = n$, $e_1, \ldots, e_n$ — базис линейного пространства V, $f \in V^*$, $x \in V$. Тогда $x = x_1 e_1 + \ldots + x_n e_n$, $x_i \in K$, $1 \leqslant i \leqslant n$, $f(x) = x_1 f(e_1) + \ldots + x_n f(e_n)$. Функционал f однозначно определяется элементами $f(e_1), \ldots, f(e_n) \in K$. Отображение

$$V^* \to K^n, \quad f \to \big(f(e_1), \ldots, f(e_n)\big)$$

является изоморфизмом линейных пространств. Таким образом, $V^* \cong V$ (так как $\dim V^* = n = \dim V$).

Рассмотрим следующие функционалы $f_i \in V^*$, $1 \leqslant i \leqslant n$:

$$f_i(e_i) = 1, \quad f_i(e_j) = 0 \ \text{ при } \ j \neq i.$$

Так как

$$\left(\sum_{i=1}^{n} \alpha_i f_i\right)(e_j) = \alpha_j, \quad 1 \leqslant j \leqslant n,$$

то функционалы f_i, $1 \leqslant i \leqslant n$, линейное независимые. Для любых функционала $f \in V^*$ и элемента $x \in V$ имеем

$$f(x) = \sum_{i=1}^{n} x_i f(e_i),$$

следовательно,

$$f = \sum_{i=1}^{n} f(e_i) f_i.$$

Итак, $f_1, \ldots, f_n$ — базис линейного пространства V^*, этот базис называется *дуальным базисом* к базису $e_1, \ldots, e_n$ линейного пространства V. Элемент $f \in V^*$ имеет координаты $\big(f(e_1), \ldots, f(e_n)\big)$ в базисе $f_1, \ldots, f_n$.

Если

$$\mathcal{E} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}, \quad \mathcal{E}' = \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} -$$

базисы линейного пространства V, $C \in \mathrm{GL}_n(K)$ — матрица перехода от базиса $\mathcal{E}$ к базису $\mathcal{E}'$, $\mathcal{E}' = C^*\mathcal{E}$, то, так как f — линейное отображение, имеем

$$\begin{pmatrix} f(e_1') \\ \vdots \\ f(e_n') \end{pmatrix} = C^* \begin{pmatrix} f(e_1) \\ \vdots \\ f(e_n) \end{pmatrix}.$$

Это равенство задаёт правило замены координат в линейном пространстве V^* при переходе от базиса $f_1, \ldots, f_n$ (дуального к базису $e_1, \ldots, e_n$ линейного пространства V) к базису $f_1', \ldots, f_n'$ (дуального к базису $e_1', \ldots, e_n'$ линейного пространства V).

Рассмотрим теперь второе сопряжённое пространство V^{**} и построим канонический изоморфизм $V \to V^{**}$. Любой элемент $x \in V$ рассмотрим как следующий линейный функционал на линейном пространстве V^* (т. е. как элемент $\tilde{x} \in V^{**}$):

$$\tilde{x}\colon V^* \to K, \quad \tilde{x}(f) = f(x) \text{ для всех } f \in V^*. \tag{7}$$

Построенное отображение

$$V \ni x \to \tilde{x} \in V^{**}$$

задаёт инъективное линейное отображение $V \to V^{**}$. Действительно, если $\tilde{x} = 0$, то $f(x) = \tilde{x}(f) = 0$ для всех $f \in V^*$, поэтому $x = 0$. Так как $\dim V = \dim V^* = \dim V^{**}$, то это отображение является изоморфизмом линейных пространств.

8.15. Теорема о ранге матрицы

Пусть $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ — прямоугольная $(m \times n)$-матрица с элементами a_{ij} из поля K. Определитель $M_{i_1,\ldots,i_k;j_1,\ldots,j_k}$ квадратной $(k \times k)$-матрицы, состоящей из элементов на пересечении k строк с номерами $i_1, \ldots, i_k$ и k столбцов с номерами $j_1, \ldots, j_k$, называется *минором k-го порядка* матрицы A. Наивысший порядок *ненулевого* минора матрицы A обозначим через $\mathrm{r}(A)$.

Теорема 8.54 (о ранге матрицы). *Следующие четыре числовые характеристики матрицы $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ совпадают:*

1) $\mathrm{r}(A_1, \ldots, A_m)$ *(ранг системы строк, в K^n);*

2) $\mathrm{r}(\hat{A}_1, \ldots, \hat{A}_n)$ *(ранг системы столбцов, в $\hat{K}^n$);*

3) $\mathrm{r}(A)$ *(наивысший порядок ненулевого минора);*

4) *число ненулевых строк r в ступенчатом виде $\bar{A}$ матрицы A.*

(Это совпадающее число называется рангом матрицы A и будет обозначаться через $\mathrm{r}(A)$).

Доказательство разобьём на четыре леммы.

Лемма 8.55. *Пусть матрица $\tilde{A}$ получена из матрицы A элементарным преобразованием строк (столбцов) 1-го или 2-го типа, тогда $\mathrm{r}(A) = \mathrm{r}(\tilde{A})$. Если $\bar{A}$ — ступенчатая форма, к которой приводится матрица A, то $\mathrm{r}(A) = \mathrm{r}(\bar{A})$.*

Доказательство проведём для преобразований строк (для столбцов всё аналогично).

Случай 1. $A_i' = A_i + cA_j$, $c \in K$, $i \neq j$. Для $k > \mathrm{r}(A)$ рассмотрим минор $\tilde{M} = \tilde{M}_{i_1,\ldots,i_k;j_1,\ldots,j_k}$ в $\tilde{A}$.

а) Если $i \notin \{i_1, \ldots, i_k\}$, то $\tilde{M} = M_{i_1,\ldots,i_k;j_1,\ldots,j_k} = 0$.

б) Если $i, j \in \{i_1, \ldots, i_k\}$, то $\tilde{M} = M_{i_1,\ldots,i_k;j_1,\ldots,j_k} = 0$.

в) Если $i \in \{i_1, \ldots, i_k\} \not\ni j$, то разложим определитель $\tilde{M}$ по i-й строке $A_i' = A_i + cA_j$ в сумму двух определителей: $\tilde{M} = M + c\tilde{\Delta} = 0$, так как $M = M_{i_1,\ldots,i_k;j_1,\ldots,j_k} = 0$, поскольку $k > \mathrm{r}(A)$, определитель $\tilde{\Delta}$ в качестве i-й строчки имеет часть строки A_j, но $j \notin \{i_1, \ldots, i_k\}$, и поэтому $\tilde{\Delta}$ отличается от минора матрицы порядка k перестановкой двух строк, и поэтому $\tilde{\Delta} = 0$. Итак, $\mathrm{r}(\tilde{A}) \leqslant \mathrm{r}(A)$. Поскольку от A к $\tilde{A}$ можно вернуться элементарным преобразованием строк, то $\mathrm{r}(A) \leqslant \mathrm{r}(\tilde{A})$.

Случай 2. $A_i \leftrightarrow A_j$ разбирается аналогично $(i, j \in \{i_1, \ldots, i_k\};\ i, j \notin \{i_1, \ldots, i_k\};\ i \in \{i_1, \ldots, i_k\} \not\ni j)$. □

Лемма 8.56 (о сохранении линейных соотношений между столбцами при элементарных преобразованиях строк). *Пусть от матрицы A к матрице A' мы перешли элементарными преобразованиями строк, тогда столбцы матриц A и A' имеют одни и те же линейные соотношения, а именно, $k_1\hat{A}_1 + \ldots + k_n\hat{A}_n = 0$ тогда и только тогда, когда $k_1\hat{A}'_1 + \ldots + k_n\hat{A}'_n = 0$.*

Доказательство. Ясно, что элементарные преобразования 1-го и 2-го типа для строк сохраняют линейное соотношение для столбцов и эти преобразования обратимы. □

Следствие 8.57. *Система столбцов $\hat{A}_{j_1}, \ldots, \hat{A}_{j_r}$ матрицы A линейно зависима (соответственно, линейно независима или является максимальной линейно независимой подсистемой в $\hat{A}_1, \ldots, \hat{A}_n \in \hat{K}^m$) тогда и только тогда, когда соответствующая система столбцов (с теми же номерами) $\hat{A}'_{j_1}, \ldots, \hat{A}'_{j_r}$ матрицы A' линейно зависима (соответственно линейно независима или является максимальной линейно независимой подсистемой в $\hat{A}'_1, \ldots, \hat{A}'_n \in \hat{K}^m$).*

Следствие 8.58. $\mathrm{r}\{\hat{A}_1, \ldots, \hat{A}_n\} = \mathrm{r}\{\hat{A}'_1, \ldots, \hat{A}'_n\}$.

Лемма 8.59. *Если $\bar{A}$ — ступенчатая матрица, то наивысший порядок ненулевого минора $\mathrm{r}(\bar{A})$ совпадает с числом r ненулевых строк.*

Доказательство.

1) Минор r-го порядка на пересечении r ненулевых строк и столбцов, проходящих через уголки ступенек, является определителем треугольной матрицы с ненулевыми элементами на главной диагонали, и поэтому отличен от нуля.

2) Все миноры, порядок которых больше r, нулевые, так как имеют нулевую строку. □

Лемма 8.60. *В ступенчатой матрице $\bar{A}$ ранг системы столбцов совпадает с числом r ненулевых строк (а именно, столбцы, проходящие через уголки ступенек, образуют максимальную линейно независимую подсистему столбцов).*

Доказательство.

1) Указанные столбцы линейно независимы, так как проходят через $(r \times r)$-матрицу с ненулевым определителем.

2) Любой столбец ступенчатой матрицы является линейной комбинацией указанных. □

Следствие 8.61 (алгоритм нахождения максимальной линейно независимой подсистемы в системе столбцов прямоугольной матрицы). *От матрицы A перейдём к ступенчатой матрице $\bar{A}$ с помощью элементарных преобразований строк 1-го и 2-го типов, запомним номера столбцов $j_1, \ldots, j_r$, проходящих через уголки ступенек в $\bar{A}$, в матрице A возьмём столбцы с этими номерами $\hat{A}_{j_1}, \ldots, \hat{A}_{j_r}$.*

Пример 8.62. Найти какую-либо максимальную линейно независимую подсистему строк в системе $a_1, a_2, a_3, a_4 \in \mathbb{R}^4$,

$$a_1 = (-1, 4, -3, -2), \quad a_2 = (3, -7, 5, 3),$$
$$a_3 = (3, -2, 1, 0), \quad a_4 = (-4, 1, 0, 1),$$

а остальные строки выразить как линейные комбинации строк этой подсистемы.

Решение. Записываем строки a_1, a_2, a_3, a_4 как столбцы и приводим полученную матрицу к главному ступенчатому виду с помощью элементарных преобразований строк:

$$\begin{pmatrix} -1 & 3 & 3 & -4 \\ 4 & -7 & -2 & 1 \\ -3 & 5 & 1 & 0 \\ -2 & 3 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} -1 & 3 & 3 & -4 \\ 0 & 5 & 10 & -15 \\ 0 & -4 & -8 & 12 \\ 0 & -3 & -6 & 9 \end{pmatrix} \to \begin{pmatrix} -1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 & -5 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Записываем номера столбцов в ступенчатом виде, проходящие через уголки ступенек: 1, 2. Поэтому $\{a_1, a_2\}$ — максимальная линейно независимая подсистема, $a_3 = 3a_1 + 2a_2$, $a_4 = -5a_1 - 3a_2$; ранг системы строк a_1, a_2, a_3, a_4 равен 2.

Завершение доказательства теоремы о ранге:

$$\mathrm{r}(\hat{A}_1, \ldots, \hat{A}_n) \overset{\text{лемма 8.56}}{=} \mathrm{r}(\hat{\bar{A}}_1, \ldots, \hat{\bar{A}}_n) \text{ (ранг столбцов ступенчатой матрицы } \bar{A}) \overset{\text{лемма 8.60}}{=}$$
$$= r \overset{\text{лемма 8.59}}{=} \mathrm{r}(\bar{A}) \overset{\text{лемма 8.55}}{=} \mathrm{r}(A) = \mathrm{r}(A^*) \overset{\text{лемма 8.56}}{=} \mathrm{r}(A_1, \ldots, A_m). \quad \square$$

Теорема 8.63. *Пусть $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$, $B = (b_{ij}) \in \mathrm{M}_{n,r}(K)$. Тогда*

$$\mathrm{r}(AB) \leqslant \mathrm{r}(A), \quad \mathrm{r}(AB) \leqslant \mathrm{r}(B).$$

Доказательство. Пусть $C = (c_{ij}) = AB$. Тогда

$$\begin{aligned} c_{ij} &= a_{i1}b_{1j} + \ldots + a_{in}b_{nj}, \\ C_i &= a_{i1}B_1 + \ldots + a_{in}B_n, \\ \hat{C}_j &= \hat{A}_1 b_{1j} + \ldots + \hat{A}_n b_{nj}, \end{aligned}$$

т. е. строки матрицы C линейно выражаются через строки матрицы B, столбцы матрицы C линейно выражаются через столбцы матрицы A. Поэтому $\mathrm{r}(C) \leqslant \mathrm{r}(B)$ и $\mathrm{r}(C) \leqslant \mathrm{r}(A)$. $\square$

Следствие 8.64. *При умножении на квадратную матрицу A с $|A| \neq 0$ ранг не меняется.*

Доказательство. Так как $|A| \neq 0$, то существует обратная матрица A^{-1}. Поэтому

$$(BA)A^{-1} = B = A^{-1}(AB),$$

и следовательно,

$$\mathrm{r}(B) \leqslant \mathrm{r}(BA), \quad \mathrm{r}(B) \leqslant \mathrm{r}(AB).$$

Ранее мы доказали, что

$$\mathrm{r}(B) \geqslant \mathrm{r}(BA), \quad \mathrm{r}(B) \geqslant \mathrm{r}(AB).$$

Поэтому

$$\mathrm{r}(B) = \mathrm{r}(BA), \quad \mathrm{r}(B) = \mathrm{r}(AB). \quad \square$$

Теорема 8.65 (о факториальном ранге). *Пусть $m, n \in \mathbb{N}$, $A \in \mathrm{M}_{m,n}(K)$. Ранг матрицы $\mathrm{r}(A)$ равен наименьшему числу k такому, что*

$$A = B \cdot C, \quad \text{где } B \in \mathrm{M}_{m,k}(K), \quad C \in \mathrm{M}_{k,n}(K)$$

(это число k называется факториальным рангом матрицы A).

Доказательство. Допустим, что $A = B \cdot C$, где $B \in \mathrm{M}_{m,n}(K)$, $C \in \mathrm{M}_{k,n}(K)$. Тогда система столбцов матрицы A линейно выражается через систему столбцов матрицы B (их k штук). Поэтому $\mathrm{r}(A) \leqslant k$.

Пусть $k = \mathrm{r}(A)$. Выберем строки $A_{i_1}, \ldots, A_{i_k}$, образующие максимальную линейно независимую подсистему строк $A_1, \ldots, A_m$ матрицы A,

$$A_i = \beta_{i1} A_{i_1} + \ldots + \beta_{ik} A_{i_k}, \quad \beta_{ij} \in K, \quad 1 \leqslant i \leqslant m.$$

Рассмотрим матрицы $B \in \mathrm{M}_{m,k}(K)$, $B = (\beta_{ij})$, и $C \in \mathrm{M}_{k,n}(K)$, для которой j-я строка $C_j = A_{i_j}$, $j = 1, \ldots, k$. Тогда $A = B \cdot C$. □

Теорема 8.66 (теорема Кронекера—Капелли: критерий совместности и определённости системы линейных уравнений в терминах рангов матриц). *Пусть $(a_{ij} \mid b_i)$ — система m линейных уравнений с n неизвестными, $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ — матрица коэффициентов,*

$$A' = \left(\begin{array}{c|c} & b_1 \\ A & \vdots \\ & b_m \end{array} \right) -$$

расширенная матрица системы линейных уравнений.

а) *Система линейных уравнений совместна тогда и только тогда, когда ранг матрицы коэффициентов A равен рангу расширенной матрицы $A' = (A, \hat{b})$, $\mathrm{r}(A) = \mathrm{r}(A')$.*

б) *Система линейных уравнений определённая тогда и только тогда, когда $\mathrm{r}(A) = \mathrm{r}(A') = n$.*

Доказательство.

1) Используя определение ранга матрицы с помощью столбцов, видим, что всегда $\mathrm{r}(A) \leqslant \mathrm{r}(A')$.

2) Если $(k_1, \ldots, k_n)$ — решение, то

$$k_1 \hat{A}_1 + \ldots + k_n \hat{A}_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

т. е. столбцы матрицы A' линейно выражаются через столбцы матрицы A, следовательно, $\mathrm{r}(A') \leqslant \mathrm{r}(A)$, и поэтому $\mathrm{r}(A') = \mathrm{r}(A)$.

3) Пусть $\mathrm{r}(A') = \mathrm{r}(A) = r$. Тогда максимальная линейно независимая система столбцов матрицы A содержит r столбцов, и поэтому она является и максимальной линейно независимой системой столбцов матрицы A'. Таким образом, столбец

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

линейно выражается через эту систему столбцов матрицы A, а поэтому и через *все* столбцы матрицы A,

$$k_1 \hat{A}_1 + \ldots + k_n \hat{A}_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Итак, существует решение $(k_1, \ldots, k_n)$ системы линейных уравнений. □

Второе доказательство. Элементарными преобразованиями приведём систему линейных уравнений к ступенчатому виду (ранги матриц не меняются при этом). Совпадение рангов означает отсутствие «экзотических» уравнений в ступенчатом виде, т. е. совместность системы линейных уравнений. □

4) *Доказательство критерия определённости (в терминах рангов).* Если система определена, т. е. $\mathrm{r}(A) = \mathrm{r}(A')$, то она определена тогда и только тогда, когда в ступенчатом виде нет свободных неизвестных, т. е. $\mathrm{r}(A) = \mathrm{r}(A') = n$. □

8.16. Размерность пространства решений однородной системы линейных уравнений

Как мы отметили ранее, совокупность решений $X_{\text{одн}}$ однородной системы линейных уравнений с матрицей $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$ является линейным пространством, подпространством в K^n.

Теорема 8.67. *Если $r = \mathrm{r}(A) < n$, то $\dim X_{\text{одн}} = n - r$ (т. е. размерность пространства решений равна числу свободных неизвестных). (Если $\mathrm{r}(A) = n$, то система линейных уравнений имеет лишь нулевое решение.)*

Доказательство. Для удобства записи переупорядочим неизвестные, если это необходимо, так, чтобы

$$\underbrace{x_1, \ldots, x_r}_{r \text{ главных неизвестных}} \quad \text{и} \quad \underbrace{x_{r+1}, \ldots, x_n}_{n-r \text{ свободных неизвестных}}.$$

Пусть $E = E_{n-r} \in \mathrm{M}_{n-r}(K)$ — единичная матрица размера $(n-r) \times (n-r)$. Возьмём её строки в качестве наборов значений для свободных неизвестных и дополним их (единственно возможным способом) до решений нашей системы линейных уравнений

$$\alpha_1 = (c_{11}, \ldots, c_{1r}, 1, 0, \ldots, 0),$$
$$\vdots$$
$$\alpha_{n-r} = (c_{(n-r)1}, \ldots, c_{(n-r)r}, 0, 0, \ldots, 1).$$

Эта система $n - r$ строк-решений линейно независима (поскольку строки единичной матрицы, конечно, линейно независимы). Если

$$\beta = (\beta_1, \ldots, \beta_{n-r}, \beta_{n-r+1}, \ldots, \beta_n) \in X_{\text{одн}} -$$

произвольное решение, то

$$\gamma = \beta - \beta_{n-r+1}\alpha_1 - \ldots - \beta_n\alpha_{n-r} = (\gamma_1, \ldots, \gamma_{n-r}, 0, \ldots, 0) \in X_{\text{одн}}.$$

Однако, конечно,

$$(0, \ldots, 0, 0, \ldots, 0) \in X_{\text{одн}},$$

при этом γ и нулевое решение имеют одинаковый набор значений для свободных неизвестных. Так как значения главных неизвестных однозначно определяются по свободным, то $\gamma = 0$, следовательно,

$$\beta = \beta_{n-r+1}\alpha_1 + \ldots + \beta_n\alpha_{n-r}.$$

Итак, мы построили базис $\{\alpha_1, \ldots, \alpha_{n-r}\}$ линейного пространства решений $X_{\text{одн}}$, поэтому $\dim X_{\text{одн}} = n - r$. □

Замечание 8.68. Если вместо строк единичной матрицы E_{n-r} для свободных неизвестных брать строки всевозможных матриц $C \in \mathrm{GL}_{n-r}(K)$ (т. е. $C \in \mathrm{M}_n(K)$, $|C| \neq 0$), то этот алгоритм позволяет построить все базисы в $X_{\text{одн}}$.

Замечание 8.69. Любой базис линейного пространства решений $X_{\text{одн}}$ однородной системы линейных уравнений называется в ряде алгебраических текстов «фундаментальной системой решений однородной системы линейных уравнений».

8.17. Задание любого подпространства в ${}_K V = K^n$ как пространства решений однородной системы линейных уравнений

Пусть K — поле, $u_1, \dots, u_m \in {}_K V = K^n$, $U = \langle u_1, \dots, u_m \rangle$ — подпространство в K^n, являющееся линейной оболочкой строк $u_1, \dots, u_m$, т. е. множеством всех линейных комбинаций строк $u_1, \dots, u_m$. Мы найдём такую матрицу $A \in \mathrm{M}_{s,n}(K)$, что множество решений однородной системы линейных уравнений

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

совпадает с U.

Если U — нулевое подпространство, то в качестве A мы можем взять любую матрицу $n \times n$ с ненулевым определителем (например, $A = E$). Если $U = K^n$ (это эквивалентно тому, что $\dim U = n$), то в качестве A мы можем взять нулевую матрицу из $\mathrm{M}_{s,n}$, $s \geqslant 1$. Если же $1 \leqslant \dim U = \mathrm{r}(u_1, \dots, u_m) < n$, то пусть $u_i = (u_{i1}, u_{i2}, \dots, u_{in})$, $1 \leqslant i \leqslant m$, $u_{ij} \in K$.

Рассмотрим матрицу $B \in \mathrm{M}_{m,n}(K)$, $B = (b_{ij})$, $b_{ij} = u_{ij}$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, и однородную систему линейных уравнений

$$B \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \left.\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\right\} m. \tag{8}$$

Ясно, что $r = \mathrm{r}(B) = \dim U$, поэтому $1 \leqslant r < n$. Размерность s пространства решений $X_{\text{одн}}$ этой системы равна $n - r$, и так как $1 \leqslant r < n$, то $1 \leqslant s < n$.

Пусть строки $v_1, \dots, v_s \in K^n$ образуют фундаментальную систему решений системы (8), $v_i = (v_{i1}, \dots, v_{in})$, $1 \leqslant i \leqslant s$, $v_{ij} \in K$. Пусть $A \in \mathrm{M}_{s,n}(K)$, $A = (a_{ij})$, $a_{ij} = v_{ij}$, $1 \leqslant i \leqslant s$, $1 \leqslant j \leqslant n$. Покажем, что A — искомая матрица.

Действительно, по построению матрицы A любая строка из U (как линейная комбинация строк $u_1, \dots, u_m$) является решением однородной системы уравнений

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{9}$$

т. е. $U \subseteq X_{\text{одн}}$. С другой стороны,

$$\dim X_{\text{одн}} = n - \mathrm{r}(A) = n - s = n - (n - r) = r = \dim U.$$

Следовательно, $U = X_{\text{одн}}$. □

В заключение отметим, что матрица A определена неоднозначно. Например, другая матрица A' может быть получена с помощью другой фундаментальной системы решений системы (8).

Полученное задание линейных подпространств оказывается полезным при решении ряда практических задач. Например, пусть $u_1, \dots, u_m \in \mathbb{R}^n$ — линейно независимые строки, $m < n$. Требуется найти такие строки $u_{m+1}, \dots, u_n$, что $\{u_1, \dots, u_n\}$ — базис линейного пространства $\mathbb{R}^n$. Как и выше, пусть $v_1, \dots, v_s$ — какая-нибудь фундаментальная система решений системы (8) (в нашем случае $\mathrm{r}(B) = m$, $s = n - m$). Положим $u_{m+1} = v_1, \dots, u_n = v_{n-m}$. Покажем, что $\{u_1, \dots, u_n\}$ — базис в $\mathbb{R}^n$. Достаточно показать, что строки $u_1, \ldots, u_n$ линейно независимы над $\mathbb{R}$. Пусть $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ и $\alpha_1 u_1 + \ldots + \alpha_n u_n = 0 \in \mathbb{R}^n$. Тогда для строки

$$z = \alpha_1 u_1 + \ldots + \alpha_m u_m = -\alpha_{m+1} u_{m+1} - \ldots - \alpha_n u_n$$

имеем $z \in U \cap V$, где $V = \langle u_{m+1}, \ldots, u_n \rangle$. Если $z = (z_1, \ldots, z_n)$, $z_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, то по построению подпространств U и V (см. (8), (9)) имеем

$$(z_1, \ldots, z_n) \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = 0,$$

$z_1^2 + \ldots + z_n^2 = 0$, следовательно, $z_1 = \ldots = z_n = 0$, и $z = 0 \in \mathbb{R}^n$. Значит,

$$\alpha_1 u_1 + \ldots + \alpha_m u_m = 0 (\in \mathbb{R}^n) = \alpha_{m+1} u_{m+1} + \ldots + \alpha_n u_n.$$

Но $u_1, \ldots, u_m$ — линейно независимые строки, поэтому $\alpha_1 = \ldots = \alpha_m = 0$. Строки $u_{m+1}, \ldots, u_n$ также линейно независимы, следовательно, $\alpha_{m+1} = \ldots = \alpha_n = 0$. Итак, $\alpha_1 = \ldots = \alpha_n = 0$ и строки $u_1, \ldots, u_n$ линейно независимы.

Таким образом, мы рассмотрели два способа задания линейных подпространств в ${}_K V = K^n$:

1) как множество решений $X_{\text{одн}}$ однородной системы линейных уравнений;

2) как линейную оболочку $\langle u_1, \ldots, u_m \rangle$ строк $u_1, \ldots, u_m \in {}_K V = K^n$.

При этом мы научились переходить от первого задания ко второму (фундаментальная система решений) и от второго задания к первому. Первый способ задания удобен для задания пересечения $U \cap W$ подпространств (надо к первой однородной системе уравнений приписать вторую). Второй способ задания удобен для задания суммы подпространств:

$$\langle u_1, \ldots, u_m \rangle + \langle w_1, \ldots, w_t \rangle = \langle u_1, \ldots, u_m, w_1, \ldots, w_t \rangle.$$

В следующем примере мы увидим комбинацию этих приёмов.

Пример 8.70. Пусть $V_1 = \langle u_1, u_2, u_3 \rangle \subseteq \mathbb{R}^4$ (линейная оболочка строк $u_1 = (1,1,0,0)$, $u_2 = (0,1,1,0)$, $u_3 = (0,0,1,1)$), $V_2 = \langle v_1, v_2, v_3 \rangle \subseteq \mathbb{R}^4$ (линейная оболочка строк $v_1 = (1,0,1,0)$, $v_2 = (0,2,1,1)$, $v_3 = (1,2,1,2)$). Необходимо найти базисы линейных пространств $V_1 + V_2$ и $V_1 \cap V_2$, при этом строки u_1, u_2, u_3, v_1, v_2, v_3 выразить через базис пространства $V_1 + V_2$.

Решение. Запишем строки u_1, u_2, u_3, v_1, v_2, v_3 по столбцам и приведём полученную матрицу к ступенчатому виду с помощью элементарных преобразований строк:

$$\begin{array}{c} \begin{matrix} u_1 & u_2 & u_3 & v_1 & v_2 & v_3 \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 2 & 2 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix} \end{array} \to \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix} \to \begin{array}{c} \begin{matrix} u_1' & u_2' & u_3' & v_1' & v_2' & v_3' \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \end{array}.$$

Поскольку $V_1 + V_2 = \langle u_1, u_2, u_3, v_1, v_2, v_3\rangle$ и элементарные преобразования строк матрицы не меняют линейных соотношений между столбцами, то $\{u_1, u_2, u_3, v_1\}$ — базис в $V_1 + V_2$ (и так как $\dim(V_1 + V_2) = 4$, то $V_1 + V_2 = \mathbb{R}^4$). Из ступенчатого вида мы вычисляем v_2' и v_3' через u_1', u_2', u_3', v_1':

$$v_2' + v_1' = \begin{pmatrix}1\\1\\1\\0\end{pmatrix} = u_1' + u_2' + u_3'.$$

Поэтому $v_2' = u_1' + u_2' + u_3' - v_1'$ и, следовательно, $v_2 = u_1 + u_2 + u_3 - v_1$. Для v_3' мы видим, что $v_3' + v_1' = (2,0,2,0)^* = 2u_1' + 2u_3'$, поэтому $v_3 = 2u_1 + 2u_3 - v_1$. Проведённые вычисления равносильны завершению приведения матрицы к главному ступенчатому виду:

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

Рассмотрим теперь $V_1 \cap V_2$. Для этого найдём однородные системы линейных уравнений, чьи множества решений совпадают с V_1 и V_2 соответственно.

Для V_1:

$$\begin{pmatrix} 1 & 1 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix},$$

система уже имеет ступенчатый вид, x_1, x_2, x_3 — главные неизвестные, x_4 — свободная. Фундаментальная система решений состоит из одной строки $(-1, 1, -1, 1)$. Итак, подпространство V_1 совпадает с пространством решений однородной системы линейных уравнений

$$(-1, 1, -1, 1)\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = 0. \tag{10}$$

Для V_2:

$$\begin{pmatrix} 1 & 0 & 1 & 0\\ 0 & 2 & 1 & 1\\ 1 & 2 & 1 & 2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & 0\\ 0 & 2 & 1 & 1\\ 0 & 2 & 0 & 2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 0 & 1 & 0\\ 0 & 2 & 1 & 1\\ 0 & 1 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 0 & 2 & 1 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} \to$$

$$\to \begin{pmatrix} 1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & -1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix},$$

и мы приходим к ступенчатому виду, при этом x_1, x_2, x_3 — главные неизвестные, а x_4 — свободная. Фундаментальная система решений состоит из одной строки $(-1, -1, 1, 1)$. Значит, однородная система линейных уравнений

$$(-1,-1,1,1)\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=0 \tag{11}$$

задаёт подпространство V_2.

§ 9. Линейные операторы линейного пространства

Пусть V — линейное пространство над полем K. Напомним, что отображение $\mathcal{A}\colon V\to V$ называется *линейным отображением* или *линейным оператором* (на пространстве V), если для всех $x,y\in V$, $\alpha\in K$ имеем

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

Пусть $\dim V=n<\infty$, $e_1,\ldots,e_n$ — базис линейного пространства V. В силу свойства линейности любой линейный оператор однозначно определяется (задаётся) образами базисных элементов $w_j=\mathcal{A}(e_j)\in V$, $j=1,\ldots,n$. Действительно, если

$$x=\sum_{j=1}^n x_je_j,$$

то

$$\mathcal{A}(x)=\mathcal{A}\biggl(\sum_{j=1}^n x_je_j\biggr)=\sum_{j=1}^n x_j\mathcal{A}(e_j)=\sum_{j=1}^n x_jw_j.$$

Если же заданы элементы $w_1,\ldots,w_n\in V$, то для

$$x=x_1e_1+\ldots+x_ne_n=\sum_{j=1}^n x_je_j\in V$$

определим

$$\mathcal{A}(x)=x_1w_1+\ldots+x_nw_n=\sum_{j=1}^n x_jw_j.$$

Полученное отображение $\mathcal{A}\colon V\to V$ является линейным, поскольку для $y=\sum\limits_{j=1}^n y_je_j$ и $\alpha\in K$ имеем:

$$\mathcal{A}(x+y)=\mathcal{A}\biggl(\sum_{j=1}^n x_je_j+\sum_{j=1}^n y_je_j\biggr)=\mathcal{A}\Bigl((x_j+y_j)e_j\Bigr)=$$
$$=\sum_{j=1}^n(x_j+y_j)w_j=\sum_{j=1}^n x_jw_j+\sum_{j=1}^n y_jw_j=\mathcal{A}(x)+\mathcal{A}(y);$$
$$\mathcal{A}(\alpha x)=\mathcal{A}\biggl(\alpha\sum_{j=1}^n x_je_j\biggr)=\mathcal{A}\biggl(\sum_{j=1}^n(\alpha x_j)e_j\biggr)=\sum_{j=1}^n(\alpha x_j)w_j=\alpha\sum_{j=1}^n x_jw_j=\alpha\mathcal{A}(x).$$

Если же

$$w_j=\mathcal{A}(e_j)=a_{1j}e_1+a_{2j}e_2+\ldots+a_{nj}e_n=\sum_{i=1}^n a_{ij}e_i,$$

где $a_{ij} \in K$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$, то линейный оператор $\mathcal{A}\colon V \to V$ полностью определяется и задаётся квадратной **($n \times n$)-матрицей** $A = (a_{ij}) \in \mathrm{M}_n(K)$, $\mathcal{A}(e_j) = \sum\limits_{i=1}^{n} a_{ij} e_i$, называемой матрицей линейного оператора $\mathcal{A}$ **в базисе** $e_1, \ldots, e_n$.

Таким образом, если

$$x = \sum_{j=1}^{n} x_j e_j, \quad X = (x_1, \ldots, x_n) \in \mathrm{M}_{1n}(K),$$

$$w = \mathcal{A}(x) = \sum_{j=1}^{n} x_j \mathcal{A}(e_j) = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{n} a_{ij} e_i \right) = \sum_{i,j=1}^{n} (x_j a_{ij}) e_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} x_j \right) e_i,$$

$$w = \sum_{i=1}^{n} w_i e_i, \quad W = (w_1, \ldots, w_n) \in \mathrm{M}_{1n}(K),$$

$$\hat{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \hat{W} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathrm{M}_{n1}(K),$$

то

$$w_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, n,$$

другими словами,

$$\hat{W} = A\hat{X} \quad (\text{в } \mathrm{M}_{n1}(K)).$$

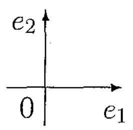
Если $A \in \mathrm{M}_n(K)$, $e_1, \ldots, e_n$ — базис линейного пространства ${}_K V$, то рассмотрим отображение $\mathcal{A}\colon V \to V$: для $x = x_1 e_1 + \ldots + x_n e_n$, $x_i \in K$, положим $w = \mathcal{A}(x) = w_1 e_1 + \ldots + w_n e_n$, где $\hat{W} = A\hat{X}$. Тогда $\mathcal{A}$ — линейный оператор на линейном пространстве V: $\mathcal{A}(x + y) = \mathcal{A}(x) + \mathcal{A}(y)$ для $x, y \in V$, поскольку $A(\hat{X} + \hat{Y}) = A\hat{X} + A\hat{Y}$, и $\mathcal{A}(\alpha x) = \alpha \mathcal{A}(x)$, поскольку $A(\alpha \hat{X}) = \alpha A \hat{X}$ для $\alpha \in K$. Матрица линейного оператора $\mathcal{A}$ в базисе $e_1, \ldots, e_n$ совпадает с A.

Запись $\hat{W} = A\hat{X}$ является матричной записью выражения $w = \mathcal{A}(x)$ с использованием изоморфизма $\varphi\colon V \to \hat{K}^n$, для $x = x_1 e_1 + \ldots + x_n e_n \in V$, $x_i \in K$,

$$\varphi(x) = \hat{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \hat{K}^n.$$

Примеры линейных операторов

1) Пусть $K = \mathbb{R}$, $V = \mathbb{R}^2$ (линейное пространство векторов на действительной плоскости, выходящих из начала координат), $\mathcal{A}$ — поворот векторов на угол φ против часовой стрелки. Тогда $\mathcal{A}$ — линейный оператор на пространстве V, и в стандартном базисе e_1, e_2



имеем

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

2) Пусть $K = \mathbb{R}$, $n \in \mathbb{N}$, V — линейное пространство всех действительных многочленов от переменной t степени $\leqslant n$. Тогда

$$e_0 = 1,\ e_1 = t,\ e_2 = \frac{t^2}{2!}, \ldots,\ e_n = \frac{t^n}{n!} -$$

базис пространства V. Рассмотрим отображение $\mathcal{A}\colon V \to V$, заданное правилом $\mathcal{A}\big(f(t)\big) = f'(t)$ (взятие производной). Из свойств производной функции следует, что $\mathcal{A}$ — линейный оператор на пространстве V. При этом матрица A оператора $\mathcal{A}$ в базисе $e_0, e_1, \ldots, e_n$ имеет следующий вид:

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots\ldots & & 0 \\ 0 & 0 & 1 & & & \vdots \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & \ldots\ldots & & 0 & 1 & 0 \\ 0 & \ldots\ldots & & 0 & 0 & 1 \\ 0 & \ldots\ldots & & 0 & 0 & 0 \end{pmatrix}$$

(жорданова клетка размера $n+1$ с нулём на главной диагонали).

3) Нулевой оператор $\mathcal{O}\colon V \to V$, $\mathcal{O}(x) = 0$ для всех $x \in V$. В любом базисе $\{e_1, \ldots, e_n\}$ пространства V матрица нулевого линейного оператора — нулевая $(n \times n)$-матрица $O_n \in \mathrm{M}_n(K)$.

4) Тождественный оператор $\mathcal{I}_V\colon V \to V$, $\mathcal{I}(x) = x$ для всех $x \in V$. В любом базисе $\{e_1, \ldots, e_n\}$ пространства V матрица тождественного оператора — это единичная $(n \times n)$-матрица $E_n \in \mathrm{M}_n(K)$.

Лемма 9.1. *Пусть $\mathcal{A}$ — линейный оператор на линейном пространстве V, $v \in V$, 0 — нулевой элемент пространства V. Тогда $\mathcal{A}(0) = 0$, $\mathcal{A}(-v) = -\mathcal{A}(v)$.*

Доказательство.

$$\begin{aligned} 0 &= \mathcal{A}(0) + \big(-\mathcal{A}(0)\big) = \mathcal{A}(0+0) + \big(-\mathcal{A}(0)\big) = \\ &= \big(\mathcal{A}(0) + \mathcal{A}(0)\big) + \big(-\mathcal{A}(0)\big) = \mathcal{A}(0) + \big(\mathcal{A}(0) - \mathcal{A}(0)\big) = \mathcal{A}(0) + 0 = \mathcal{A}(0); \\ 0 &= \mathcal{A}(0) = \mathcal{A}\big(v + (-v)\big) = \mathcal{A}(v) + \mathcal{A}(-v). \end{aligned}$$

□

Замечание 9.2. Пусть $\mathcal{A}\colon V \to V$ — линейный оператор на линейном пространстве V и существует обратное отображение $\mathcal{A}^{-1}\colon V \to V$, $\mathcal{A}\mathcal{A}^{-1}(x) = \mathcal{A}^{-1}\mathcal{A}(x) = x$ для всех элементов $x \in V$. Тогда $\mathcal{A}^{-1}$ также линейный оператор на пространстве V: для всех $u, v \in V$

$$\mathcal{A}\big(\mathcal{A}^{-1}(u+v)\big) = u + v = \mathcal{A}\big(\mathcal{A}^{-1}(u)\big) + \mathcal{A}\big(\mathcal{A}^{-1}(v)\big) = \mathcal{A}\big(\mathcal{A}^{-1}(u) + \mathcal{A}^{-1}(v)\big).$$

Так как отображение $\mathcal{A}$ инъективно, то

$$\mathcal{A}^{-1}(u+v) = \mathcal{A}^{-1}(u) + \mathcal{A}^{-1}(v).$$

Аналогично, для $v \in V$ и $\alpha \in K$ из

$$\mathcal{A}\big(\mathcal{A}^{-1}(\alpha v)\big) = \alpha v = \alpha\mathcal{A}\big(\mathcal{A}^{-1}(v)\big) = \mathcal{A}\big(\alpha\mathcal{A}^{-1}(v)\big)$$

следует, что

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

Итак, $\mathcal{A}^{-1}$ — линейный оператор на пространстве V.

Если $\mathcal{A}$ — линейный оператор на линейном пространстве V и для $x \in V$ имеем $\mathcal{A}(x) = 0$ тогда и только тогда, когда $x = 0$, то оператор $\mathcal{A}$ называется *невырожденным*. В противном случае оператор $\mathcal{A}$ называется *вырожденным*. Если $\dim V = n < \infty$, $e_1, \ldots, e_n$ — базис линейного пространства V, то условие $\mathcal{A}(x) = 0$ равносильно условию

$$A\hat{X} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

где A — матрица линейного оператора $\mathcal{A}$ в базисе $e_1, \ldots, e_n$,

$$\hat{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} -$$

столбец координат элемента $x = x_1e_1 + \ldots + x_ne_n$ в том же базисе. Существование ненулевого решения однородной системы линейных уравнений

$$A\hat{X} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

эквивалентно условию $|A| = 0$. Таким образом, оператор $\mathcal{A}$ невырожден в том и только в том случае, когда $|A| \neq 0$.

Лемма 9.3. *Пусть $\mathcal{A}$ — невырожденный линейный оператор на линейном пространстве V, $v_1, \ldots, v_s$ — линейно независимые элементы пространства V. Тогда $\mathcal{A}(v_1), \ldots, \mathcal{A}(v_s)$ также линейно независимые элементы.*

Доказательство. Пусть $\alpha_1, \ldots, \alpha_s \in K$ и

$$\alpha_1\mathcal{A}(v_1) + \ldots + \alpha_s\mathcal{A}(v_s) = 0 \in V.$$

Тогда

$$\mathcal{A}(\alpha_1v_1 + \ldots + \alpha_sv_s) = \alpha_1\mathcal{A}(v_1) + \ldots + \alpha_s\mathcal{A}(v_s) = 0.$$

Так как $\mathcal{A}$ — невырожденный оператор, то $\alpha_1v_1 + \ldots + \alpha_sv_s = 0$, и $\alpha_1 = \ldots = \alpha_s = 0$ (элементы $v_1, \ldots, v_s$ линейно независимы). $\square$

Рассмотрим следующие операции на множестве линейных операторов $\mathrm{End}({}_KV)$ на линейном пространстве V.

1) Сложение операторов. Пусть $\mathcal{A}$ и $\mathcal{B}$ — линейные операторы на пространстве V. Рассмотрим отображение $\mathcal{C}\colon V \to V$, заданное правилом $\mathcal{C}(x) = \mathcal{A}(x) + \mathcal{B}(x)$ для всех $x \in V$ (обозначение: $\mathcal{C} = \mathcal{A} + \mathcal{B}$). Проверим, что $\mathcal{C}$ — линейный оператор на пространстве V. Для $x, y \in V$, $\alpha \in K$ в силу линейности операторов $\mathcal{A}$ и $\mathcal{B}$ и свойств линейного пространства V имеем:

$$\begin{aligned} \mathcal{C}(x+y) &= \mathcal{A}(x+y) + \mathcal{B}(x+y) = \mathcal{A}(x) + \mathcal{A}(y) + \mathcal{B}(x) + \mathcal{B}(y) = \\ &= \big(\mathcal{A}(x) + \mathcal{B}(x)\big) + \big(\mathcal{A}(y) + \mathcal{B}(y)\big) = \mathcal{C}(x) + \mathcal{C}(y), \\ \mathcal{C}(\alpha x) &= \mathcal{A}(\alpha x) + \mathcal{B}(\alpha x) = \alpha\mathcal{A}(x) + \alpha\mathcal{B}(x) = \alpha\big(\mathcal{A}(x) + \mathcal{B}(x)\big) = \alpha\mathcal{C}(x). \end{aligned}$$

2) Умножение линейного оператора на число: для $\alpha \in K$ и линейного оператора $\mathcal{A}$ на линейном пространстве ${}_KV$ над полем K рассмотрим отображение $\alpha\mathcal{A}\colon V \to V$, заданное правилом $(\alpha\mathcal{A})(x) = \alpha\mathcal{A}(x)$ для всех $x \in V$. Проверим, что $\alpha\mathcal{A}$ — линейный оператор. Для $x, y \in V$, $\lambda \in K$

$$(\alpha\mathcal{A})(x+y) = \alpha\mathcal{A}(x+y) = \alpha\big(\mathcal{A}(x) + \mathcal{A}(y)\big) = \alpha\mathcal{A}(x) + \alpha\mathcal{A}(y) = (\alpha\mathcal{A})(x) + (\alpha\mathcal{A})(y),$$

$$(\alpha\mathcal{A})(\lambda x) = \alpha\mathcal{A}(\lambda x) = \alpha\lambda\mathcal{A}(x) = \lambda\big(\alpha\mathcal{A}(x)\big) = \lambda\big((\alpha\mathcal{A})(x)\big).$$

Таким образом, $\alpha\mathcal{A}$ — линейный оператор на пространстве V.

Ряд свойств следует из определений сложения линейных операторов, умножения линейного оператора на число, нулевого оператора $\mathcal{O}$ и из аксиом линейного пространства V:

$$\begin{aligned}
&(\mathcal{A}+\mathcal{B})+\mathcal{C} = \mathcal{A}+(\mathcal{B}+\mathcal{C});\\
&\mathcal{A}+\mathcal{B} = \mathcal{B}+\mathcal{A};\\
&\mathcal{O}+\mathcal{A} = \mathcal{A}+\mathcal{O} = \mathcal{A};\\
&1\cdot\mathcal{A} = \mathcal{A};\\
&(-1)\mathcal{A}+\mathcal{A} = \mathcal{A}+(-1)\mathcal{A} = \mathcal{O};\\
&\alpha(\beta\mathcal{A}) = (\alpha\beta)\mathcal{A};\\
&(\alpha+\beta)\mathcal{A} = \alpha\mathcal{A}+\beta\mathcal{A};\\
&\alpha(\mathcal{A}+\mathcal{B}) = \alpha\mathcal{A}+\alpha\mathcal{B}
\end{aligned}$$

для всех $\alpha, \beta \in K$, $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathrm{End}({}_KV)$. Таким образом, множество всех линейных операторов $\mathrm{End}({}_KV)$ на пространстве ${}_KV$ над полем K образует линейное пространство над полем K.

3) Умножение (композиция) линейных операторов в $\mathrm{End}({}_KV)$. Пусть $\mathcal{A}$ и $\mathcal{B}$ — линейные операторы на линейном пространстве V. Рассмотрим отображение $\mathcal{C}\colon V \to V$, заданное правилом $\mathcal{C}(x) = \mathcal{A}\big(\mathcal{B}(x)\big)$ для всех $x \in V$. Покажем, что $\mathcal{C}$ — линейный оператор на пространстве V (называемый *произведением* (или *композицией*)) линейных операторов $\mathcal{A}$ и $\mathcal{B}$ (обозначение: $\mathcal{C} = \mathcal{AB}$). Для $x, y \in V$, $\alpha \in K$ имеем:

$$\begin{aligned}
&(\mathcal{AB})(x+y) = \mathcal{A}\big(\mathcal{B}(x+y)\big) = \mathcal{A}\big(\mathcal{B}(x)+\mathcal{B}(y)\big) = \mathcal{A}\big(\mathcal{B}(x)\big)+\mathcal{A}\big(\mathcal{B}(y)\big) = (\mathcal{AB})(x)+(\mathcal{AB})(y);\\
&(\mathcal{AB})(\alpha x) = \mathcal{A}\big(\mathcal{B}(\alpha x)\big) = \mathcal{A}\big(\alpha\mathcal{B}(x)\big) = \alpha\mathcal{A}\big(\mathcal{B}(x)\big) = \alpha\big((\mathcal{AB})(x)\big).
\end{aligned}$$

Теорема 9.4. *Пусть $\dim V = n < \infty$, $e_1, \ldots, e_n$ — базис пространства V, $\mathcal{A}$, $\mathcal{B}$ — линейные операторы на пространстве ${}_KV$, $\alpha \in K$, A и B — матрицы линейных операторов $\mathcal{A}$ и $\mathcal{B}$ соответственно в базисе $e_1, \ldots, e_n$. Тогда в базисе $e_1, \ldots, e_n$ оператор $\mathcal{A}+\mathcal{B}$ имеет матрицу $A+B$, $\alpha\mathcal{A}$ — матрицу αA, $\mathcal{AB}$ — матрицу AB.*

Доказательство. Пусть $A = (a_{ij})$, $B = (b_{ij}) \in \mathrm{M}_n(K)$. Тогда

$$(\mathcal{A}+\mathcal{B})(e_j) = \mathcal{A}(e_j)+\mathcal{B}(e_j) = \sum_{i=1}^n a_{ij}e_i + \sum_{i=1}^n b_{ij}e_i = \sum_{i=1}^n (a_{ij}+b_{ij})e_i, \quad 1 \leqslant j \leqslant n.$$

Следовательно, матрица линейного оператора $\mathcal{A}+\mathcal{B}$ в базисе $e_1, \ldots, e_n$ равна $A+B$. В матричной форме: если $y = (\mathcal{A}+\mathcal{B})(x)$, то $y = \mathcal{A}(x)+\mathcal{B}(x)$, $\hat{Y} = A\hat{X}+B\hat{X} = (A+B)\hat{X}$ (и следовательно, оператор $\mathcal{A}+\mathcal{B}$ имеет матрицу $A+B$ в базисе $e_1, \ldots, e_n$).

Аналогично, для $\alpha \in K$ имеем

$$(\alpha\mathcal{A})(e_j) = \alpha\mathcal{A}(e_j) = \alpha\sum_{i=1}^n a_{ij}e_i = \sum_{i=1}^n (\alpha a_{ij})e_i,$$

поэтому матрица линейного оператора $\alpha\mathcal{A}$ в базисе $e_1, \ldots, e_n$ равна αA. В матричной форме: если $y = (\alpha\mathcal{A})(x)$, то $y = \alpha\mathcal{A}(x)$, $\hat{Y} = \alpha A\hat{X}$, и следовательно, оператор $\alpha\mathcal{A}$ имеет матрицу αA в базисе $e_1, \ldots, e_n$.

Если $y = (\mathcal{AB})(x) = \mathcal{A}\big(\mathcal{B}(x)\big)$, $x \in V$, то $\hat{Y} = A(B\hat{X}) = (AB)\hat{X}$ (мы воспользуемся ассоциативностью умножения матриц). Следовательно, матрица линейного оператора $\mathcal{AB}$ в базисе $e_1, \ldots, e_n$ равна AB. $\square$

Операция умножения линейных операторов ассоциативна: для линейных операторов $\mathcal{A}$, $\mathcal{B}$ и $\mathcal{C}$ на пространстве V выполнено равенство

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

Действительно, как для любых отображений,

$$\big(\mathcal{A}(\mathcal{B}\mathcal{C})\big)(x) = \mathcal{A}\big((\mathcal{B}\mathcal{C})(x)\big) = \mathcal{A}\big(\mathcal{B}(\mathcal{C}(x))\big) = (\mathcal{A}\mathcal{B})\big(\mathcal{C}(x)\big)$$

для всех $x \in V$.

Тождественный оператор $\mathcal{I}$ является единичным (нейтральным) элементом относительно операции умножения: для любого линейного оператора $\mathcal{A}$ на линейном пространстве V имеем

$$\mathcal{I}\mathcal{A} = \mathcal{A}\mathcal{I} = \mathcal{A}.$$

Действительно, $(\mathcal{I}\mathcal{A})(x) = \mathcal{I}\big(\mathcal{A}(x)\big) = \mathcal{A}(x)$, $(\mathcal{A}\mathcal{I})(x) = \mathcal{A}\big(\mathcal{I}(x)\big) = \mathcal{A}(x)$ для всех $x \in V$.

Операции сложения и умножения линейных операторов удовлетворяют закону дистрибутивности

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

для любых линейных операторов $\mathcal{A}$, $\mathcal{B}$ и $\mathcal{C}$ на пространстве ${}_KV$:

$$\begin{aligned}\big(\mathcal{A}(\mathcal{B}+\mathcal{C})\big)(x) &= \mathcal{A}\big((\mathcal{B}+\mathcal{C})(x)\big) = \mathcal{A}\big(\mathcal{B}(x)+\mathcal{C}(x)\big) = \\ &= \mathcal{A}\big(\mathcal{B}(x)\big) + \mathcal{A}\big(\mathcal{C}(x)\big) = (\mathcal{A}\mathcal{B})(x) + (\mathcal{A}\mathcal{C})(x) = (\mathcal{A}\mathcal{B} + \mathcal{A}\mathcal{C})(x).\end{aligned}$$

Если $\alpha \in K$, то $\mathcal{A}(\alpha\mathcal{B}) = \alpha(\mathcal{A}\mathcal{B}) = (\alpha\mathcal{A})\mathcal{B}$:

$$\big(\mathcal{A}(\alpha\mathcal{B})\big)(x) = \mathcal{A}\big((\alpha\mathcal{B})(x)\big) = \mathcal{A}\big(\alpha\mathcal{B}(x)\big) = (\alpha\mathcal{A})\big(\mathcal{B}(x)\big) = (\alpha\mathcal{A}\mathcal{B})(x)$$

для всех $x \in V$.

Собирая все свойства операций сложения, умножения на число, умножения операторов, получаем, что множество $\mathrm{End}({}_KV)$ линейных операторов на линейном пространстве V над полем K образует ассоциативную алгебру с единичным элементом над полем K.

Замечание 9.5. Пусть $\dim_K V = n$. Тогда, зафиксировав базис $e_1, \dots, e_n$ пространства V и ставя в соответствие каждому линейному оператору $\mathcal{A}$ на линейном пространстве V его матрицу A в базисе $e_1, \dots, e_n$, $\varphi(\mathcal{A}) = A$, получаем изоморфизм ассоциативных алгебр над полем K: $\mathrm{End}({}_KV) \cong \mathrm{M}_n(K)$ (φ — биекция, изоморфизм линейных пространств, $\varphi(\mathcal{A}\mathcal{B}) = \varphi(\mathcal{A})\varphi(\mathcal{B}) = AB$ для всех $\mathcal{A}, \mathcal{B} \in \mathrm{End}({}_KV)$).

9.1. Изменение матрицы линейного оператора при переходе к новому базису

Пусть $e_1, \dots, e_n$ — базис линейного пространства V,

$$\mathcal{E} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

$e'_1, \dots, e'_n$ — «новый» базис,

$$\mathcal{E}' = \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}$$

$\mathcal{E}' = C^*\mathcal{E}$, где C — матрица перехода ($|C| \neq 0$), $\mathcal{A}$ — линейный оператор на пространстве V, имеющий матрицу A в базисе $e_1, \ldots, e_n$. Тогда матрица оператора $\mathcal{A}$ в базисе $e'_1, \ldots, e'_n$ равна $C^{-1}AC$.

Пусть $y = \mathcal{A}(x)$, $x \in V$, $\hat{X}$ — столбец координат элемента x в базисе $e_1, \ldots, e_n$. Тогда $\hat{Y} = A\hat{X}$. Если $\hat{X}'$ и $\hat{Y}'$ — столбцы координат элементов x и y соответственно в базисе $e'_1, \ldots, e'_n$, то $\hat{X} = C\hat{X}'$, $\hat{Y} = C\hat{Y}'$. Поэтому $C\hat{Y}' = \hat{Y} = A\hat{X} = AC\hat{X}'$. Следовательно, $\hat{Y}' = (C^{-1}AC)\hat{X}'$. Это означает, что матрица линейного оператора $\mathcal{A}$ в базисе $e'_1, \ldots, e'_n$ совпадает с $C^{-1}AC$. Так как для любых матриц $A, B \in \mathrm{M}_n(K)$ имеем $|AB| = |A| \cdot |B|$ (см. **??**), то определитель матрицы линейного оператора $\mathcal{A}$ не зависит от выбора базиса:

$$|C^{-1}AC| = |C^{-1}|\,|A|\,|C| = |C|^{-1}\,|A|\,|C| = |A|.$$

Замечание 9.6. Линейный оператор $\mathcal{A}$ на линейном пространстве V называется *обратимым*, если существует такой линейный оператор $\mathcal{B}$ (обозначение: $\mathcal{B} = \mathcal{A}^{-1}$), что $\mathcal{AB} = \mathcal{BA} = \mathcal{I}$ (тождественный оператор).

Если $\dim V = n < \infty$, $e_1, \ldots, e_n$ — базис пространства V, A — матрица линейного оператора в базисе $e_1, \ldots, e_n$, то условие существования обратного оператора $\mathcal{B}$ в терминах матриц операторов имеет вид $AB = BA = E$ (таким образом, $B = A^{-1}$). Следовательно, для конечномерных линейных пространств понятия обратимого и невырожденного линейного оператора совпадают. Для обратимого линейного оператора $\mathcal{A}$ обратный оператор определён однозначно, его матрица равна A^{-1}.

Ясно, что обратимые линейные операторы (и только они) обладают следующим свойством: если $v_1, \ldots, v_n$ — базис линейного пространства V, то $\mathcal{A}(v_1), \ldots, \mathcal{A}(v_n)$ также базис линейного пространства V.

9.2. Образ и ядро линейного оператора

Пусть $\mathcal{A}$ — линейный оператор на линейном пространстве ${}_K V$. Определим образ

$$\operatorname{Im} \mathcal{A} = \{\mathcal{A}x \mid x \in V\}$$

и ядро

$$\operatorname{Ker} \mathcal{A} = \{x \in V \mid \mathcal{A}x = 0 \in V\}$$

линейного оператора $\mathcal{A}$.

Образ и ядро являются подпространствами линейного пространства V: если $x, y \in \operatorname{Im} \mathcal{A}$, то найдутся такие $z, w \in V$, что $x = \mathcal{A}(z)$, $y = \mathcal{A}(w)$, поэтому

$$x + y = \mathcal{A}(z) + \mathcal{A}(w) = \mathcal{A}(z + w) \in \operatorname{Im} \mathcal{A};$$

для $\alpha \in K$ имеем

$$\alpha x = \alpha\mathcal{A}(z) = \mathcal{A}(\alpha z) \in \operatorname{Im} \mathcal{A};$$

если же $x, y \in \operatorname{Ker} \mathcal{A}$, то

$$\mathcal{A}(x + y) = \mathcal{A}(x) + \mathcal{A}(y) = 0 + 0 = 0 \in V,$$

поэтому $x + y \in \operatorname{Ker} \mathcal{A}$; для $\alpha \in K$:

$$\mathcal{A}(\alpha x) = \alpha\mathcal{A}(x) = \alpha \cdot 0 = 0 \in V,$$

следовательно, $\alpha x \in \operatorname{Ker} \mathcal{A}$.

Рангом $\mathrm{r}(\mathcal{A})$ линейного оператора $\mathcal{A}$ называется число $\dim \operatorname{Im} \mathcal{A}$ (размерность образа), а дефектом — число $\mathrm{d}(\mathcal{A}) = \dim \operatorname{Ker} \mathcal{A}$ (размерность его ядра).

Пусть $e_1, \dots, e_n$ — базис линейного пространства V. Отметим, что линейное пространство $\operatorname{Im}\mathcal{A}$ является линейной оболочкой элементов $\mathcal{A}(e_1), \dots, \mathcal{A}(e_n)$. Таким образом, число $\dim \operatorname{Im}\mathcal{A}$ совпадает с рангом системы элементов $\mathcal{A}(e_1), \dots, \mathcal{A}(e_n)$, следовательно, с рангом (по столбцам) матрицы A линейного оператора $\mathcal{A}$ в базисе $e_1, \dots, e_n$. Итак, $\mathrm{r}(\mathcal{A}) = \dim \operatorname{Im}\mathcal{A} = \mathrm{r}(A)$. Поскольку $\mathrm{r}(C^{-1}AC) = \mathrm{r}(A)$ (см. ??), ранг матрицы линейного оператора в любом базисе совпадает с рангом оператора (размерностью образа).

Теорема 9.7. *Сумма ранга и дефекта линейного оператора совпадает с размерностью линейного пространства:*

$$\mathrm{r}(\mathcal{A}) + \mathrm{d}(\mathcal{A}) = \dim \operatorname{Im}\mathcal{A} + \dim \operatorname{Ker}\mathcal{A} = \dim_K V = n.$$

Доказательство. Пусть $r = \mathrm{r}(\mathcal{A})$. Не ограничивая общности, можно считать, что элементы $\mathcal{A}(e_1), \dots, \mathcal{A}(e_r)$ образуют базис линейного пространства: $\operatorname{Im}\mathcal{A} = \langle \mathcal{A}(e_1), \dots, \mathcal{A}(e_r) \rangle$. Пусть $L = \langle e_1, \dots, e_r \rangle$, $N = \operatorname{Ker}\mathcal{A}$.

Покажем, что $V = L \oplus N$ (это доказывает утверждение теоремы, поскольку $\dim(L \oplus N) = \dim L + \dim N$). Пусть $x \in L \cap N$. Так как $x \in L$, то $x = \alpha_1 e_1 + \ldots + \alpha_r e_r$, $\alpha_i \in K$, $i = 1, \dots, r$. Так как $x \in N$, то $0 = \mathcal{A}(x) = \alpha_1 \mathcal{A}(e_1) + \ldots + \alpha_r \mathcal{A}(e_r)$. Но элементы $\mathcal{A}(e_1), \dots, \mathcal{A}(e_r)$ линейно независимы, поэтому $\alpha_1 = \ldots = \alpha_r = 0$, и следовательно, $x = 0$. Поэтому $L \cap N = \{0\}$.

Покажем теперь, что $L + N = V$. Пусть $x \in V$. Тогда $\mathcal{A}(x) \in \operatorname{Im}\mathcal{A}$, $\mathcal{A}(x) = \beta_1 \mathcal{A}(e_1) + \ldots + \beta_r \mathcal{A}(e_r)$, где $\beta_i \in K$, $i = 1, \dots, r$. Рассмотрим $y = \beta_1 e_1 + \ldots + \beta_r e_r$. Ясно, что $y \in L$. Для элемента $z = x - y$ имеем $\mathcal{A}(z) = \mathcal{A}(x - y) = \mathcal{A}(x) - \mathcal{A}(y) = 0 \in V$, то есть $z \in N$. Итак, $x = y + z$, где $y \in L$, $z \in N$, следовательно, $V = L + N$. Поскольку $L \cap N = \{0\}$, $V = L \oplus N$. □

Замечание 9.8. Можно доказать теорему 9.7 в матричной форме следующим образом. Пусть $e_1, \dots, e_n$ — базис линейного пространства V. Дефект линейного оператора совпадает с размерностью пространства решений однородной системы линейных уравнений

$$A\hat{X} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

(здесь A — матрица оператора $\mathcal{A}$ в базисе $e_1, \dots, e_n$), равной $n - \mathrm{r}(A)$ (см. ??). Это доказывает теорему.

Замечание 9.9. Пусть $\dim V = n < \infty$, $\mathcal{A}$ — невырожденный (т. е. обратимый) линейный оператор на пространстве V. Тогда:

1) $\mathrm{r}(\mathcal{A}) = n$. Действительно, для невырожденного линейного оператора $\mathcal{A}$ имеем $\operatorname{Ker}\mathcal{A} = \{0\}$, и из теоремы ?? следует, что $\mathrm{r}(\mathcal{A}) = n$ (если $\mathrm{r}(\mathcal{A}) = n$, то $\mathrm{r}(A) = n$, матрица A обратима, т. е. $\mathcal{A}$ — обратимый оператор).

2) Если $\mathcal{B}$ — линейный оператор на пространстве V, то $\mathrm{r}(\mathcal{AB}) = \mathrm{r}(\mathcal{BA}) = \mathrm{r}(\mathcal{B})$. Действительно, $\mathrm{r}(\mathcal{A}) = n = \dim V$, в любом базисе пространства V матрица A оператора $\mathcal{A}$ обратима, а для матрицы B линейного оператора $\mathcal{B}$ в этом базисе имеем $\mathrm{r}(B) = \mathrm{r}(\mathcal{B})$. Но мы уже доказали, что в этом случае $\mathrm{r}(AB) = \mathrm{r}(BA) = \mathrm{r}(B)$ (см. ??).

Замечание 9.10. В теореме 9.7 нельзя утверждать, что $\operatorname{Im}\mathcal{A} + \operatorname{Ker}\mathcal{A} = V$. Пусть, например, $K = \mathbb{R}$, $n \in \mathbb{N}$, V — пространство всех многочленов от одной переменной степени $\leqslant n$, $\mathcal{A}$ — линейный оператор взятия производной. Тогда $\operatorname{Im}\mathcal{A}$ — пространство многочленов степени $\leqslant n - 1$, $\operatorname{Im}\mathcal{A} \neq V$, $\operatorname{Ker}\mathcal{A}$ — одномерное подпространство констант (многочленов степени ниже 1), $\operatorname{Ker}\mathcal{A} \subseteq \operatorname{Im}\mathcal{A}$. Поэтому $\operatorname{Im}\mathcal{A} + \operatorname{Ker}\mathcal{A} \neq V$.

Другой пример: $K = \mathbb{R}$, $V = \mathbb{R}^2$, для $(a, b) \in V$ положим $\mathcal{A}(a, b) = (b, 0)$. Тогда $\mathcal{A}$ — линейный оператор на пространстве V, $\operatorname{Im}\mathcal{A} = \operatorname{Ker}\mathcal{A}$.

Рассмотрим теперь класс линейных операторов, для которых $\operatorname{Im}\mathcal{A} \oplus \operatorname{Ker}\mathcal{A} = V$.

Определение 9.11. Пусть U и W — подпространства линейного пространства V, $V = U \oplus W$. Если $v \in V$, то $v = u + w$, $u \in U$, $w \in W$, и элементы u и w определены однозначно (см. ??). Рассмотрим отображение $\mathcal{A}\colon V \to V$, $\mathcal{A}(v) = u$. Ясно, что $\mathcal{A}$ — линейный оператор на пространстве ${}_KV$, он называется **проектором** на подпространство U вдоль подпространства W.

Пример 9.12. $V = \mathbb{R}^2$, e_1, e_2 — стандартный базис векторов на плоскости, $\mathcal{A}$ — проектирование вектора на ось e_1 вдоль оси e_2.

Для проектора $\mathcal{A}$ на подпространство U вдоль подпространства W имеем $U = \operatorname{Im}\mathcal{A}$, $W = \operatorname{Ker}\mathcal{A}$, поэтому $V = \operatorname{Im}\mathcal{A} \oplus \operatorname{Ker}\mathcal{A}$. Обратное, вообще говоря, неверно: пусть $K = \mathbb{R}$, ${}_\mathbb{R}V = {}_\mathbb{R}U \oplus {}_\mathbb{R}W$, для $v = u + w$, где $u \in U$, $w \in W$, положим $\mathcal{A}(v) = \frac{1}{2}u$. Тогда $\mathcal{A}$ — линейный оператор на линейном пространстве V, $V = \operatorname{Im}\mathcal{A} \oplus \operatorname{Ker}\mathcal{A}$, однако линейный оператор $\mathcal{A}$ не является проектором.

Теорема 9.13. *Линейный оператор $\mathcal{A}$ на линейном пространстве V является проектором в том и только в том случае, когда $\mathcal{A}^2 = \mathcal{A}$.*

Доказательство. Ясно, что если $\mathcal{A}$ — проектор, то $\mathcal{A}^2 = \mathcal{A}$: пусть $V = U \oplus V$, тогда $\mathcal{A}^2(u+v) = \mathcal{A}(u) = u = \mathcal{A}(u+v)$. Обратно, допустим, что $\mathcal{A}^2 = \mathcal{A}$. Докажем, что ${}_KV = \operatorname{Im}\mathcal{A} \oplus \operatorname{Ker}\mathcal{A}$. Пусть $v \in \operatorname{Im}\mathcal{A} \cap \operatorname{Ker}\mathcal{A}$. Так как $v \in \operatorname{Ker}\mathcal{A}$, то $\mathcal{A}(v) = 0$. Так как $v \in \operatorname{Im}\mathcal{A}$, то $v = \mathcal{A}(u)$ для некоторого $u \in V$. Но тогда $0 = \mathcal{A}(v) = \mathcal{A}^2(u) = \mathcal{A}(u) = v$. Следовательно, $\operatorname{Im}\mathcal{A} \cap \operatorname{Ker}\mathcal{A} = \{0\}$. Для $v \in V$ рассмотрим представление $v = \mathcal{A}(v) + \big(v - \mathcal{A}(v)\big)$. При этом $\mathcal{A}(v) \in \operatorname{Im}\mathcal{A}$, $\mathcal{A}\big(v - \mathcal{A}(v)\big) = \mathcal{A}(v) - \mathcal{A}^2(v) = 0 \in V$. Поэтому $V = \operatorname{Im}\mathcal{A} + \operatorname{Ker}\mathcal{A}$. Так как $\operatorname{Im}\mathcal{A} \cap \operatorname{Ker}\mathcal{A}$, то $V = \operatorname{Im}\mathcal{A} \oplus \operatorname{Ker}\mathcal{A}$.

Пусть $U = \operatorname{Im}\mathcal{A}$, $V = \operatorname{Ker}\mathcal{A}$. Покажем, что $\mathcal{A}$ — проектор на подпространство U вдоль подпространства W. Пусть $v \in V$, $v = u + w$, $u \in U$, $w \in W$. Так как $U = \operatorname{Im}\mathcal{A}$, то $u = \mathcal{A}(x)$ для некоторого $x \in V$, и $\mathcal{A}(v) = \mathcal{A}(u) + \mathcal{A}(w) = \mathcal{A}(u) = \mathcal{A}^2(x) = \mathcal{A}(x) = u$ ($\mathcal{A}(w) = 0$, поскольку $w \in W = \operatorname{Ker}\mathcal{A}$). Итак, $\mathcal{A}$ — проектор на подпространство U вдоль подпространства W. $\square$

Замечание 9.14 (о линейных отображениях линейных пространств). Пусть K — поле, V и W — линейные пространства над полем K, $\mathcal{A}\colon V \to W$ — линейное отображение ($\mathcal{A}(x+y) = \mathcal{A}(x) + \mathcal{A}(y)$, $\mathcal{A}(\alpha x) = \alpha\mathcal{A}(x)$ для всех $x, y \in V$, $\alpha \in K$), $\dim V = n < \infty$, $\dim W = m < \infty$, $e_1, \ldots, e_n$ — базис линейного пространства V, $f_1, \ldots, f_m$ — базис линейного пространства W. Тогда линейное отображение $\mathcal{A}$ однозначно определяется элементами $\mathcal{A}(e_j)$, $j = 1, \ldots, n$,

$$\mathcal{A}(e_j) = a_{1j}f_1 + \ldots + a_{mj}f_m, \quad 1 \leqslant j \leqslant n \tag{12}$$

$(a_{1j}, \ldots, a_{mj})$ — координаты элемента $\mathcal{A}(e_j)$ в базисе $f_1, \ldots, f_m$, $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ — j-й столбец матрицы $A = (a_{ij}) \in \mathrm{M}_{m,n}(K)$, которая называется матрицей линейного отображения $\mathcal{A}$ по отношению к базисам $e_1, \ldots, e_n$ и $f_1, \ldots, f_m$ (при фиксированных базисах эта матрица определена однозначно). Если $w = \mathcal{A}(x)$, $X = (x_1, \ldots, x_n)$ — строка координат элемента x в базисе $e_1, \ldots, e_n$, $W = (w_1, \ldots, w_m)$ — строка координат элемента w в базисе $f_1, \ldots, f_m$, то $\hat{W} = A\hat{X}$. Если фиксировать любую матрицу $A \in \mathrm{M}_{m,n}(K)$, базисы $e_1, \ldots, e_n$ линейного пространства V и $f_1, \ldots, f_m$ линейного пространства W и рассмотреть линейное отображение $\mathcal{A}\colon V \to W$, заданное на элементах e_i по правилу (12), $\mathcal{A}(x) = x_1\mathcal{A}(e_1) + \ldots + x_n\mathcal{A}(e_n)$ для $x \in V$, $x = x_1e_1 + \ldots + x_ne_n$, то матрица линейного отображения $\mathcal{A}$ относительно базисов $e_1, \ldots, e_n$ и $f_1, \ldots, f_m$ совпадёт с A.

Пусть

$$\mathcal{E} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}, \quad \mathcal{E}' = \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix} —$$

базисы линейного пространства V,

$$\mathcal{F}' = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}, \quad \mathcal{F}' = \begin{pmatrix} f'_1 \\ \vdots \\ f'_m \end{pmatrix} -$$

базисы линейного пространства W, $\mathcal{E}' = C^*\mathcal{E}$ ($C \in \mathrm{GL}_n(K)$, C — матрица перехода от базиса $\mathcal{E}$ к базису $\mathcal{E}'$), $\mathcal{F}' = D^*\mathcal{F}$ ($D \in \mathrm{GL}_m(K)$, D — матрица перехода от базиса $\mathcal{F}$ к базису $\mathcal{F}'$). Тогда для элемента $w = \mathcal{A}(x)$ имеем $\hat{W} = A\hat{X}$, где A — матрица линейного отображения $\mathcal{A}$ относительно базисов $\mathcal{E}$ и $\mathcal{F}$. Пусть $\hat{X}'$ — столбец координат элемента x в базисе $\mathcal{E}'$, $\hat{W}'$ — столбец координат элемента w в базисе $\mathcal{F}'$. Тогда $\hat{X} = C\hat{X}'$, $\hat{W} = D\hat{W}'$ и $D\hat{W}' = AC\hat{X}'$, и следовательно, $\hat{W}' = (D^{-1}AC)\hat{X}'$. Отсюда вытекает, что матрица линейного отображения $\mathcal{A}$ относительно базисов $\mathcal{E}'$ и $\mathcal{F}'$ равна $A' = D^{-1}AC$.

Для линейного отображения $\mathcal{A}\colon V \to W$ ядро $\operatorname{Ker}\mathcal{A} = \{x \in V \mid \mathcal{A}(x) = 0 \in W\}$ является подпространством линейного пространства V, а образ $\operatorname{Im}\mathcal{A} = \{\mathcal{A}(x) \mid x \in V\}$ является подпространством линейного пространства W, ранг $\mathrm{r}(\mathcal{A})$ линейного отображения $\mathcal{A}$ ($\mathrm{r}(\mathcal{A}) = \dim \operatorname{Im}\mathcal{A}\mathrm{r}(A)$ для матрицы A линейного оператора $\mathcal{A}$ относительно любых базисов линейных пространств V и W) и дефект ($\mathrm{d}(\mathcal{A}) = \dim \operatorname{Ker}\mathcal{A}$) связаны соотношением

$$\mathrm{r}(\mathcal{A}) + \mathrm{d}(\mathcal{A}) = \dim \operatorname{Im}\mathcal{A} + \dim \operatorname{Ker}\mathcal{A} = \dim V.$$

Это соотношение особенно прозрачно для матричной характеризации элементов $\operatorname{Ker}\mathcal{A}$: при фиксированных базисах $\mathcal{E}$ в V и $\mathcal{F}$ в W элемент $x \in \operatorname{Ker}\mathcal{A}$ в том и только в том случае, когда

$$A\hat{X} = \left.\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}\right\} m.$$

Но размерность пространства решений этой однородной системы линейных уравнений равна $n - \mathrm{r}(A)$, где $\mathrm{r}(A)$ — ранг матрицы A, $\mathrm{r}(A) = \mathrm{r}(\mathcal{A})$.

9.3. Инвариантные подпространства

Пусть $\mathcal{A}$ — линейный оператор на линейном пространстве V, U — подпространство линейного пространства V. Тогда U называется инвариантным подпространством для линейного оператора $\mathcal{A}$, если $\mathcal{A}(U) = \{\mathcal{A}(u) \mid u \in U\} \subseteq U$.

Примеры 9.15.

1) Ясно, что нулевое подпространство $\{0\}$ и всё пространство V линейного пространства V являются инвариантными подпространствами для любого линейного оператора $\mathcal{A}$ на линейном пространстве V.

2) Для тождественного оператора $\mathcal{I}_V$ (и любого его кратного $\alpha\mathcal{I}_V$, $\alpha \in K$) любое подпространство инвариантно.

3) Пусть $K = \mathbb{R}$, $V = \mathbb{R}^2$ — пространство всех векторов на плоскости, выходящих из одной точки (центра). Тогда:

   а) если $\mathcal{A}$ — центральная симметрия, то в этом случае любое подпространство линейного пространства V инвариантно;

   б) если $\mathcal{A}$ — поворот вокруг центра на угол $\varphi \neq \pi k$, $k \in \mathbb{Z}$, против часовой стрелки, то в этом случае оператор $\mathcal{A}$ не имеет нетривиальных инвариантных подпространств (отличных от $\{0\}$ и V);

в) если $K = \mathbb{R}$, $n \in \mathbb{N}$, ${}_{\mathbb{R}}V = \mathrm{P}_n$ — линейное пространство всех многочленов от одной переменной степени $\leqslant n$, $\mathcal{A}$ — линейный оператор взятия производной, то в этом случае подпространства $\{0\}$, P_1, $\mathrm{P}_2, \ldots$, P_n (и только они) являются инвариантными.

Предложение 9.16. *Пусть $U, W \subseteq V$ — инвариантные подпространства для линейного оператора $\mathcal{A}\colon {}_KV \to {}_KV$. Тогда линейные подпространства $U + W$, $U \cap W$ линейного пространства ${}_KV$ также инвариантны относительно оператора $\mathcal{A}$.*

Доказательство. Пусть $x \in U + W$. Тогда $x = u + w$, где $u \in U$, $w \in W$. Поэтому $\mathcal{A}(x) = \mathcal{A}(u) + \mathcal{A}(w) \in U + W$, так как $\mathcal{A}(u) \in U \subseteq U + W$, $\mathcal{A}(w) \in W \subseteq U + W$ и $U + W$ — линейное пространство.

Пусть теперь $x \in U \cap W$. Поскольку $x \in U$ и $\mathcal{A}(U) \subseteq U$, $\mathcal{A}(x) \in U$. Поскольку $x \in W$ и $\mathcal{A}(W) \subseteq W$, $\mathcal{A}(x) \in W$. Следовательно, $\mathcal{A}(x) \in U \cap W$. □

Лемма 9.17. *Пусть $\dim V < \infty$ и $\mathcal{A}\colon V \to V$ — невырожденный линейный оператор, U — инвариантное подпространство для линейного оператора $\mathcal{A}$. Тогда U — инвариантное подпространство для линейного оператора $\mathcal{A}^{-1}$.*

Доказательство. Если $U = \{0\}$, то утверждение очевидно. Пусть $\dim U = r$, $e_1, \ldots, e_r$ — базис линейного пространства U. Так как оператор $\mathcal{A}$ невырожденный, элементы $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_r)$ линейно независимы. Подпространство U инвариантно, поэтому $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_r) \in U$. Но $\dim V = r$, поэтому $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_r)$ — базис линейного пространства U. Для любого элемента $x \in U$ имеем

$$x = \alpha_1 \mathcal{A}(e_1) + \ldots + \alpha_r \mathcal{A}(e_r), \quad \alpha_i \in K, \quad 1 \leqslant i \leqslant r.$$

Применяя к последнему равенству линейный оператор $\mathcal{A}^{-1}$, имеем

$$\mathcal{A}^{-1}(x) = \alpha_1 e_1 + \ldots + \alpha_r e_r \in U.$$ □

Замечание 9.18. Если $\mathcal{A}\colon V \to V$ — линейный оператор, U и W — нетривиальные инвариантные подпространства для оператора $\mathcal{A}$, $V = U \oplus W$, $e_1, \ldots, e_r$ — базис подпространства U, $e_{r+1}, \ldots, e_n$ — базис подпространства W, то $e_1, \ldots, e_n$ — базис линейного пространства V, и в этом базисе матрица A линейного оператора $\mathcal{A}$ имеет блочный вид

$$\left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right).$$

При этом в верхнем левом углу стоит матрица A_1 размера $r \times r$ линейного оператора $\mathcal{A}_1 = \mathcal{A}|_U : U \to U$ в базисе $e_1, \ldots, e_r$, а в правом нижнем углу — матрица A_2 размера $(n-r)\times(n-r)$ линейного оператора $\mathcal{A}_2 = \mathcal{A}|_W : W \to W$ в базисе $e_{r+1}, \ldots, e_n$.

9.4. Подстановка оператора в многочлен

Пусть ${}_KV$ — линейное пространство над полем K,

$$f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_n x^n \in K[x] -$$

многочлен с коэффициентами из поля K, $\alpha_i \in K$, $i = 0, 1, \ldots, n$. Положим

$$\mathcal{B} = f(\mathcal{A}) = \alpha_0 \mathcal{I} + \alpha_1 \mathcal{A} + \ldots + \alpha_n \mathcal{A}^n \in \mathrm{End}({}_KV),$$

где $\mathcal{I}\colon V \to V$ — тождественный линейный оператор на пространстве V. Ясно, что для линейного оператора $\mathcal{B}\colon V \to V$ имеем $\mathcal{AB} = \mathcal{BA}$. Если $e_1, \ldots, e_m$ — базис линейного пространства ${}_KV$, A — матрица линейного оператора $\mathcal{A}$ в базисе $e_1, \ldots, e_m$, то матрица B линейного оператора $\mathcal{B}$ в этом базисе равна

$$B = f(A) = \alpha_0 E + \alpha_1 A + \ldots + \alpha_n A^n \in \mathrm{M}_m(K),$$

где $E \in \mathrm{M}_m(K)$ — единичная $(m \times m)$-матрица.

Предложение 9.19. *Пусть ${}_KV$ — линейное пространство над полем K, $\mathcal{A}\colon V \to V$ — линейный оператор, $f \in K[x]$, $\mathcal{B} = f(\mathcal{A})$. Тогда подпространства $\operatorname{Im}\mathcal{B}$ и $\operatorname{Ker}\mathcal{B}$ линейного пространства V инвариантны для линейного оператора $\mathcal{A}$. В частности, если $f(x) = x$, то $\mathcal{B} = \mathcal{A}$ и $\operatorname{Im}\mathcal{A}$ и $\operatorname{Ker}\mathcal{A}$ инвариантны для оператора $\mathcal{A}$.*

Доказательство.
а) Пусть $x \in \operatorname{Im}\mathcal{B}$, $x = f(\mathcal{A})(y)$ для некоторого $y \in V$, $f(x) \in K[x]$. Тогда

$$\mathcal{A}(x) = \mathcal{A}\big(f(\mathcal{A})(y)\big) = \big(\mathcal{A}f(\mathcal{A})\big)(y) = (f(\mathcal{A})\mathcal{A})(y) = f(\mathcal{A})\big(\mathcal{A}(y)\big) \in \operatorname{Im} f(\mathcal{A}).$$

б) Пусть теперь $x \in \operatorname{Ker}\mathcal{B}$, $\mathcal{B}(x) = f(\mathcal{A})(x) = 0$. Тогда

$$f(\mathcal{A})\big(\mathcal{A}(x)\big) = (f(\mathcal{A})\mathcal{A})(x) = \big(\mathcal{A}f(\mathcal{A})\big)(x) = \mathcal{A}\big(f(\mathcal{A})(x)\big) = \mathcal{A}(0) = 0,$$

то есть $\mathcal{A}(x) \in \operatorname{Ker}\mathcal{B}$. □

Рассмотрим одномерные инвариантные подпространства. Пусть в этом случае u — базис инвариантного подпространства ${}_KU$. Так как $\mathcal{A}(U) \subseteq U$, то $\mathcal{A}(u) \in U$, следовательно, $\mathcal{A}(u) = \alpha u$, где $\alpha \in K$. Это приводит к следующему определению.

Определение 9.20. Пусть ${}_KV$ — линейное пространство над полем K, $\mathcal{A}\colon {}_KV \to {}_KV$ — линейный оператор, $x \in V$, $x \neq 0$, $\mathcal{A}(x) = \lambda x$, $\lambda \in K$. Тогда λ называется *собственным числом линейного оператора* $\mathcal{A}$, а x — *собственным вектором* (*относительно собственного числа* λ).

Если $e_1, \ldots, e_n$ — базис линейного пространства V, A — матрица линейного оператора $\mathcal{A}$ в этом базисе, $A \in \mathrm{M}_n(K)$, $x \in V$, $x = x_1e_1 + \ldots + x_ne_n$, $x_i \in K$, $X = (x_1, \ldots, x_n)$ — строка координат элемента x, то условие $\mathcal{A}(x) = \lambda x$, $x \in V$, $x \neq 0$, $\lambda \in K$, эквивалентно условию $A\hat{X} = \lambda\hat{X}$, $\hat{X} \in \hat{K}^n$, $\hat{X} \neq \begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$, $\lambda \in K$.

Пусть K — поле, $A \in \mathrm{M}_n(K)$, $0 \neq \hat{X} \in \hat{K}^n = \mathrm{M}_{n,1}(K)$, $\lambda \in K$. Если $A \cdot \hat{X} = \lambda \cdot \hat{X}$, то λ называется *собственным числом* матрицы A, а $\hat{X}$ — *собственным вектором* матрицы A, отвечающим собственному числу λ.

Условие $A\hat{X} = \lambda\hat{X}$ эквивалентно условию

$$(A - \lambda E)\hat{X} = \begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix} \in \hat{K}^n,$$

где $E \in \mathrm{M}_n(K)$ — единичная матрица. При фиксированном λ это условие превращается в однородную систему линейных уравнений относительно неизвестных $x_1, \ldots, x_n$,

$$\hat{X} = \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}.$$

Матрица $A - \lambda E$ этой системы — квадратная матрица размера n. Поэтому наличие ненулевого решения этой системы равносильно тому, что $|A - \lambda E| = 0$. Пусть t — переменная,

$$p(t) = |A - tE| = p_nt^n + p_{n-1}t^{n-1} + \ldots + p_0 \in K[t] -$$

многочлен степени n от переменной t (называемый характеристическим многочленом матрицы A), при этом:

$$p_n = (-1)^n,$$

$$p_{n-1} = (-1)^{n-1} \sum_{i=1}^{n} a_{ii} = (-1)^{n-1} \operatorname{tr} A, \quad p_0 = |A|.$$

Мы показали, что собственные числа и только они являются корнями характеристического многочлена из поля K.

Если $\lambda \in K$ и $p(\lambda) = 0$, то все собственные векторы матрицы A относительно собственного числа λ — это все ненулевые решения системы

$$(A - \lambda E)\hat{X} = (0) \in \hat{K}^n.$$

Отметим, что множество всех собственных векторов матрицы A относительно собственного числа λ не образует линейного подпространства в $\hat{K}^n$, так как все эти векторы ненулевые. Но если к этому множеству добавить нулевой вектор, то получится линейное подпространство всех решений системы

$$(A - \lambda E)\hat{X} = (0).$$

Таким образом, если $p(\lambda) = |A - \lambda E| = 0$, $r = \mathrm{r}(A - \lambda E)$, то $0 \leqslant r < n$, то размерность пространства решений этой системы равна $s = n - r$, поэтому $1 \leqslant s \leqslant n$. Если $\{X_1, \ldots, X_s\}$ — какая-либо фундаментальная система решений системы $(A - \lambda E)\hat{X} = (0)$, то все собственные векторы матрицы A, отвечающие собственному числу λ, — это все нетривиальные линейные комбинации элементов $\hat{X}_1, \ldots, \hat{X}_s$ с коэффициентами из поля K.

Таким образом, задача нахождения собственных чисел и собственных векторов линейного оператора на конечномерном линейном пространстве сводится к задаче нахождения собственных чисел и собственных векторов матрицы.

Если $\dim V = n$, A — матрица линейного оператора $\mathcal{A}$ в базисе $\mathcal{E}$ линейного пространства V, $\mathcal{E}' = C^*\mathcal{E}$ — другой базис, $C \in \mathrm{GL}_n(V)$ — матрица перехода, то матрица A' оператора $\mathcal{A}$ в базисе $\mathcal{E}'$ равна $A' = C^{-1}AC$.

Таким образом, характеристический многочлен матрицы A: $p_A(t) = |A - tE|$, характеристический многочлен матрицы A':

$$p_{A'}(t) = |A' - tE| = |C^{-1}AC - tE| = |C^{-1}(A - tE)C| = |C^{-1}|\,|A - tE|\,|C| = p_A(t).$$

Итак, характеристический многочлен матрицы линейного оператора $\mathcal{A}$ не зависит от выбора базиса линейного пространства V (и мы можем обозначать этот многочлен через $p_{\mathcal{A}}(t)$).

Из теоремы Гамильтона—Кэли для квадратных матриц (см. теорема 7.25) следует теорема Гамильтона—Кэли для линейных операторов на конечномерном линейном пространстве: если $p_{\mathcal{A}}(t)$ — характеристический многочлен линейного оператора $\mathcal{A}\colon V \to V$, то $p_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$ — нулевой оператор на пространстве V.

Пример 9.21.

$$A = \begin{pmatrix} 10 & 3 \\ -5 & 2 \end{pmatrix}, \quad K = \mathbb{R},$$

$$|A - \lambda E| = \begin{vmatrix} 10 - \lambda & 3 \\ -5 & 2 - \lambda \end{vmatrix} = \lambda^2 - 12\lambda + 35.$$

Корни: $\lambda_1 = 7$, $\lambda_2 = 5$, $\lambda_1, \lambda_2 \in \mathbb{R}$ (собственные числа матрицы A).

Собственные векторы для $\lambda_1 = 7$:

$$A - 7E = \begin{pmatrix} 3 & 3 \\ -5 & -5 \end{pmatrix}, \quad \begin{pmatrix} 3 & 3 \\ -5 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения:

$$\left\{ \begin{pmatrix} s \\ -s \end{pmatrix} \,\middle|\, s \in \mathbb{R},\ s \neq 0 \right\}.$$

Собственные векторы для $\lambda_2 = 5$:

$$A - 5E = \begin{pmatrix} 5 & 3 \\ -5 & -3 \end{pmatrix}; \quad \begin{pmatrix} 5 & 3 \\ -5 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения:

$$\left\{ \begin{pmatrix} -3t \\ 5t \end{pmatrix} \middle| \, t \in \mathbb{R}, \ t \neq 0 \right\}.$$

Пример 9.22.

$$K = \mathbb{C}, \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix},$$

$$p(\lambda) = |A - \lambda E| = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda^3.$$

Имеется лишь одно собственное число: $\lambda = 0$. Собственные векторы относительно $\boldsymbol{\lambda = 0}$ **задаются** системой линейных уравнений

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Система уже имеет ступенчатый вид, x_2, x_3 — главные неизвестные, x_1 — **свободная переменная**, множество собственных векторов относительно $\lambda = 0$:

$$\left\{ \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix} \middle| \, s \in \mathbb{C}, \ s \neq 0 \right\}.$$

Пример 9.23. Если

$$A = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix} -$$

диагональная матрица, то $\alpha_1, \ldots, \alpha_n$ — все корни характеристического многочлена матрицы A (и следовательно, собственные числа).

Пример 9.24.

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

$$p(\lambda) = |A - \lambda E| = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

а) $K = \mathbb{R}$: нет действительных корней многочлена $p(\lambda) = \lambda^2 + 1$, поэтому для матрицы A нет действительных собственных чисел (и собственных векторов).

б) $K = \mathbb{C}$: многочлен $p(\lambda)$ имеет корни $\lambda_1 = i \in \mathbb{C}$, $\lambda_2 = -i \in \mathbb{C}$ (собственные числа матрицы A).

Собственные векторы для $\lambda = i$:

$$\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения:

$$\left\{ c \cdot \begin{pmatrix} -i \\ 1 \end{pmatrix} \middle| \, c \in \mathbb{C}, \; c \neq 0 \right\}.$$

Собственные векторы для $\lambda = -i$:

$$\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

ненулевые решения

$$\left\{ c \cdot \begin{pmatrix} i \\ 1 \end{pmatrix} \middle| \, c \in \mathbb{C}, \; c \neq 0 \right\}.$$

Пример 9.25.

$$K = \mathbb{R}, \quad A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

$$p(\lambda) = |A - \lambda E| = -\lambda^3 + 12\lambda + 16.$$

Корни многочлена $p(\lambda)$: $\lambda_1 = -2$, $\lambda_2 = -2$, $\lambda_3 = 4$ (собственные числа).

Собственные векторы для $\lambda = -2$:

$$(A + 2E) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

ненулевые решения

$$\left\{ \begin{pmatrix} s - t \\ s \\ t \end{pmatrix} \middle| \, s, t \in \mathbb{R}, \; s^2 + t^2 \neq 0 \right\}.$$

Собственные векторы для $\lambda = 4$:

$$(A - 4E) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

ненулевые решения

$$\left\{ \begin{pmatrix} s \\ s \\ 2s \end{pmatrix} \middle| \, s \in \mathbb{R}, \; s \neq 0 \right\}.$$

Задача 9.26. Пусть $V = \mathbb{R}^2$ — линейное пространство всех векторов на плоскости, выходящих из одной точки (центра), $\mathcal{A}$ — линейный оператор пространства V поворота на угол φ против часовой стрелки, имеющий в стандартном базисе e_1, e_2 матрицу

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

В зависимости от φ найдите собственные числа и собственные векторы оператора $\mathcal{A}$.

Задача 9.27. Пусть $K = \mathbb{R}$, $n \in \mathbb{N}$, V — линейное пространство всех действительных многочленов от одной переменной степени $\leqslant n$, $\mathcal{A}\colon V \to V$ — линейный оператор взятия производной. Найдите собственные числа и собственные векторы оператора $\mathcal{A}$.

Замечание 9.28. Решение общей задачи нахождения собственных чисел и собственных векторов матрицы связано с нахождением корней многочлена. При степенях многочлена более 4 эта задача часто не имеет алгоритма нахождения точного решения, поэтому приходится искать приближённые значения собственных чисел. Задача приближённого нахождения собственных чисел и собственных векторов матрицы является одной из основных задач численных методов линейной алгебры.

Теорема 9.29. *Пусть $\mathcal{A}\colon {}_KV \to {}_KV$ — линейный оператор на линейном пространстве ${}_KV$ над полем K, $0 \neq v_1, \ldots, v_l \in V$, $\lambda_1, \ldots, \lambda_l \in K$, $\lambda_i \neq \lambda_j$ при $1 \leqslant i \neq j \leqslant l$, $\mathcal{A}(v_i) = \lambda_i v_i$, $i = 1, \ldots, l$. Тогда элементы $v_1, \ldots, v_l$ линейно независимы, т. е. собственные векторы, отвечающие различным собственным значениям, линейно независимы.*

Доказательство. Доказательство проведём индукцией по l. Основание индукции: $l = 1$. $A \cdot v_1 = \lambda_1 \cdot v_1$, $0 \neq v_1$, $\{v_1\}$ — линейно независимая система векторов.

Пусть теперь $l \geqslant 2$ и наше утверждение доказано для всех l', $1 \leqslant l' < l$. Допустим, что

$$\alpha_1 v_1 + \ldots + \alpha_l v_l = 0 \in V, \quad \alpha_i \in K, \quad i = 1, \ldots, l. \tag{13}$$

Применяя оператор $\mathcal{A}$ к обеим частям равенства, получаем, что

$$\alpha_1 \cdot A \cdot v_1 + \ldots + \alpha_l \cdot A \cdot v_l = A \cdot 0 = 0 \in V.$$

и поэтому

$$\alpha_1 \lambda_1 v_1 + \ldots + \alpha_l \lambda_l v_l = 0. \tag{14}$$

Умножая (13) на λ_l, имеем

$$\alpha_1 \lambda_l v_1 + \ldots + \alpha_l \lambda_l v_l = 0. \tag{15}$$

Вычитаем (15) из (14):

$$\alpha_1(\lambda_1 - \lambda_l) v_1 + \ldots + \alpha_{l-1}(\lambda_{l-1} - \lambda_l) v_{l-1} = 0 \in V.$$

Применяя предположение индукции, получаем, что

$$\alpha_1(\lambda_1 - \lambda_l) = \ldots = \alpha_{l-1}(\lambda_{l-1} - \lambda_l) = 0.$$

Поскольку $\lambda_1 \neq \lambda_l, \ldots, \lambda_{l-1} \neq \lambda_l$, отсюда следует, что $\alpha_1 = \ldots = \alpha_{l-1} = 0$. Следовательно, из (13) следует, что $\alpha_l v_l = 0 \in V$. Так как $v_l \neq 0$, то $\alpha_l = 0$. Таким образом, $\alpha_1 = \ldots = \alpha_l = 0$, и поэтому собственные векторы $v_1, \ldots, v_l$ линейно независимы. □

Пусть V — линейное пространство над полем K, $\dim V = n < \infty$, $\mathcal{A}\colon {}_KV \to {}_KV$ — линейный оператор и многочлен $p_{\mathcal{A}}(t)$ имеет n различных корней $\lambda_1, \ldots, \lambda_n \in K$. Тогда существует такой базис $e_1, \ldots, e_n$ линейного пространства V, что матрица A оператора $\mathcal{A}$ в этом базисе диагональна:

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

(в этом случае говорят, что оператор $\mathcal{A}$ *диагонализируем*).

Доказательство. Пусть v_i — собственный вектор оператора $\mathcal{A}$ относительно собственного значения λ_i, $1 \leqslant i \leqslant n$. По теореме **??** элементы $v_1, \ldots, v_n \in V$ линейно независимы. Но $\dim V = n$, поэтому $\{v_1, \ldots, v_n\}$ — базис линейного пространства V. В этом базисе матрица A оператора $\mathcal{A}$ имеет требуемый вид. □

Пусть теперь $K = \mathbb{R}$ или $K = \mathbb{C}$, V — конечномерное линейное пространство над K, $\mathcal{A}\colon {}_KV \to {}_KV$ — линейный оператор. В заключение мы рассмотрим специфику таких линейных операторов. Если $K = \mathbb{C}$, то у оператора $\mathcal{A}$ существуют собственные числа и собственные векторы. Действительно, пусть $p_{\mathcal{A}}(t) \in \mathbb{C}[t]$ — характеристический многочлен оператора $\mathcal{A}$. По основной теореме алгебры комплексных чисел все корни $\lambda_1, \ldots, \lambda_n$ многочлена $p_{\mathcal{A}}(t)$ (считая кратные) являются комплексными числами: $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Для любого λ_i собственные векторы — это ненулевые решения системы $(\mathcal{A} - \lambda_i\mathcal{I})(x) = 0 \in V$ (здесь $\mathcal{I}$ — тождественный оператор).

Теорема 9.30. *Пусть $K = \mathbb{C}$, ${}_{\mathbb{C}}V$ — ненулевое конечномерное линейное пространство над $\mathbb{C}$, $\mathcal{A}$, $\mathcal{B}$ — коммутирующие линейные операторы на ${}_{\mathbb{C}}V$: $\mathcal{AB} = \mathcal{BA}$. Тогда операторы $\mathcal{A}$ и $\mathcal{B}$ имеют общий собственный вектор.*

Доказательство. Пусть $p_{\mathcal{A}}(t)$ — характеристический многочлен оператора $\mathcal{A}$ (так как $V \neq 0$, то $p_{\mathcal{A}}(t)$ — многочлен степени $n \geqslant 1$, где $n = \dim_K V \geqslant 1$). По основной теореме алгебры комплексных чисел существует корень $\lambda \in \mathbb{C}$ многочлена $p_{\mathcal{A}}(t)$: $p_{\mathcal{A}}(\lambda) = 0$. Пусть $U \subseteq V$ — подпространство всех решений $(\mathcal{A} - \lambda\mathcal{I})(x) = 0 \in V$ (это подпространство состоит из 0 и всех собственных векторов оператора $\mathcal{A}$ относительно собственного числа λ). Ясно, что $U \neq \{0\}$.

Покажем, что U — инвариантное подпространство для линейного оператора $\mathcal{B}$. Пусть $x \in U$. Тогда

$$\mathcal{A}\big(\mathcal{B}(x)\big) = \mathcal{B}\big(\mathcal{A}(x)\big) = \mathcal{B}(\lambda x) = \lambda\mathcal{B}(x).$$

Следовательно, $\mathcal{B}(x) \in U$. Рассмотрим линейный оператор $\mathcal{B}|_U\colon {}_{\mathbb{C}}U \to {}_{\mathbb{C}}U$ (ограничение оператора $\mathcal{B}$ на инвариантное подпространство U: $\mathcal{B}|_U(x) = \mathcal{B}(x)$ для всех $x \in U$). Пусть $p_{\mathcal{B}|_U}(t)$ — характеристический многочлен оператора $\mathcal{B}|_U$. Так как $U \neq \{0\}$, то $p_{\mathcal{B}|_U}(t)$ — многочлен степени $\geqslant 1$, и поэтому существует корень $\mu \in \mathbb{C}$, $p_{\mathcal{B}|_U}(\mu) = 0$. Пусть $0 \neq u \in U$ — собственный вектор оператора $\mathcal{B}|_U$ относительно собственного числа μ. Тогда $\mathcal{B}(u) = \mu u$. Но, поскольку $u \in U$, $\mathcal{A}(u) = \lambda u$. Таким образом, u — общий собственный вектор линейных операторов $\mathcal{A}$ и $\mathcal{B}$. □

Замечание 9.31 (о комплексификации). Пусть $K = \mathbb{R}$, V — линейное пространство над $\mathbb{R}$, $\dim V = n < \infty$, $\mathcal{A}\colon V \to V$ — линейный оператор пространства V. Рассмотрим множество $W = \{(x, y) \mid x, y \in V\}$ с операциями: сложения $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$; умножения на комплексные числа $a + bi \in \mathbb{C}$ $(a, b \in \mathbb{R})$, $(a+bi)(x, y) = (ax - by, ay + bx)$. Нетрудно видеть, что W — линейное пространство над $\mathbb{C}$. При этом, обозначая $(x, 0) = x$, $iy = (0, y)$, получаем

$$(x, y) = x + iy, \quad (a + bi)(x + iy) = (ax - by) + i(ay - bx).$$

Если $e_1, \ldots, e_n$ — базис линейного пространства ${}_{\mathbb{R}}V$ над $\mathbb{R}$, то $e_1 = (e_1, 0), \ldots, e_n = (e_n, 0)$ — базис линейного пространства ${}_{\mathbb{C}}W$ над $\mathbb{C}$. Рассмотрим отображение $\tilde{A}\colon {}_{\mathbb{C}}W \to {}_{\mathbb{C}}W$, где для $x, y \in {}_{\mathbb{R}}V$, $x + iy \in W$, положим $\tilde{A}(x + iy) = \mathcal{A}(x) + i\mathcal{A}(y)$. Тогда $\tilde{A}$ — линейный оператор на пространстве ${}_{\mathbb{C}}W$ (над $\mathbb{C}$), называемый комплексификацией оператора $\mathcal{A}$. При этом матрица оператора $\tilde{A}$ в базисе $e_1 = (e_1, 0), \ldots, e_n = (e_n, 0)$ пространства ${}_{\mathbb{C}}W$ (над $\mathbb{C}$) совпадает с матрицей оператора $\mathcal{A}$ в базисе $e_1, \ldots, e_n$ пространства ${}_{\mathbb{R}}V$ (над $\mathbb{R}$). Поэтому характеристические многочлены $p_{\mathcal{A}}(t)$ и $p_{\tilde{A}}(t)$ совпадают.

Теорема 9.32. *Пусть $K = \mathbb{R}$, V — линейное пространство над K, $1 \leqslant \dim V < \infty$, $\mathcal{A}\colon {}_{\mathbb{R}}V \to {}_{\mathbb{R}}V$ — линейный оператор. Тогда для оператора $\mathcal{A}$ существует одномерное или двумерное инвариантное подпространство.*

Доказательство. Пусть $p_{\mathcal{A}}(t)$ — характеристический многочлен оператора $\mathcal{A}$. Если существует $\lambda \in \mathbb{R}$, $p_{\mathcal{A}}(\lambda) = 0$ (действительный корень), то пусть $0 \neq u \in {}_{\mathbb{R}}V$ — собственный вектор оператора $\mathcal{A}$ относительно собственного числа λ, $0 \neq {}_{\mathbb{R}}U = \langle u \rangle$. Тогда U — инвариантное подпространство для оператора $\mathcal{A}$, $\dim_{\mathbb{R}} U = 1$.

Пусть теперь многочлен $p_{\mathcal{A}}(t)$ не имеет действительных корней. По основной теореме алгебры комплексных чисел многочлен $p_{\mathcal{A}}(t)$ имеет комплексный корень $a + bi \in \mathbb{C}$, $a, b \in \mathbb{R}$,

$b \neq 0$: $p_{\mathcal{A}}(a + bi) = 0$. Рассмотрим комплексификацию $\tilde{\mathcal{A}}\colon W \to W$ линейного оператора $\mathcal{A}$ (см. замечание 9.31), при этом $a + bi$ — корень характеристического многочлена оператора $\tilde{\mathcal{A}}$: $p_{\tilde{\mathcal{A}}}(a + bi) = p_{\mathcal{A}}(a + bi) = 0$. Пусть $0 \neq x + iy \in W$ (где $x \in V$, $y \in V$) — собственный вектор оператора $\tilde{\mathcal{A}}$ относительно собственного числа $a + bi$: $\tilde{\mathcal{A}}(x + iy) = \mathcal{A}(x) + i\mathcal{A}(y) = (a + bi)(x + yi)$, при этом элементы x и y одновременно не равны нулевому элементу пространства V. Последнее равенство равносильно тому, что

$$\begin{cases} \mathcal{A}(a) = ax - by, \\ \mathcal{A}(y) = bx + ay. \end{cases}$$

Пусть $U = \langle x, y \rangle \subseteq V$. Тогда ${}_{\mathbb{R}}U$ — инвариантное подпространство для оператора $\mathcal{A}$, при этом $\dim_{\mathbb{R}} U = 2$ (так как элементы x и y одновременно не равны нулевому элементу, то $\dim U \geqslant 1$; по построению $\dim_{\mathbb{R}} U \leqslant 2$, если бы $\dim_{\mathbb{R}} U = 1$, то многочлен $p_{\mathcal{A}}(t)$ имел бы действительный корень). □

Теорема 9.33. *Матрица $A \in \mathrm{M}_n(\mathbb{C})$ нильпотентна (т. е. $A^m = 0 \in \mathrm{M}_n(K)$ для некоторого $m \in \mathbb{N}$) тогда и только тогда, когда собственные числа $\lambda_1, \ldots, \lambda_n$ равны нулю.*

Доказательство. а) Если $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$, то $|A - \lambda E| = (-1)^n \lambda^n$. По теореме Гамильтона—Кэли $A^n = (0) \in \mathrm{M}_n(\mathbb{C})$.

б) Если $A^m = (0) \in \mathrm{M}_n(\mathbb{C})$ и $A\hat{X} = \lambda \hat{X}$, где $\lambda \in \mathbb{C}$, $0 \neq \hat{X} \in \hat{\mathbb{C}}^n$, то $\hat{\mathbb{C}}^n \ni 0 = A^m \hat{X} = \lambda^n \hat{X}$, следовательно, $\lambda^n = 0$ и $\lambda = 0$. □

Замечание 9.34. Одним из фундаментальных результатов об алгебре матриц $\mathrm{M}_n(\mathbb{C})$ над полем комплексных чисел $\mathbb{C}$ (и о строении отдельно взятого линейного оператора конечномерного линейного пространства ${}_{\mathbb{C}}V$) является теорема о *жордановой нормальной форме*:

1) для каждой матрицы $A \in \mathrm{M}_n(\mathbb{C})$ найдётся такая обратимая матрица $C \in \mathrm{GL}_n(\mathbb{C})$, что

$$C^{-1}AC = J_A = \left(\begin{array}{c|c|c|c} J_1 & 0 & \ldots & 0 \\ \hline 0 & J_2 & \ldots & 0 \\ \hline & & \vdots & \\ \hline 0 & 0 & & J_k \end{array}\right) —$$

жорданова матрица (т. е. $J_1, \ldots, J_k$ — жордановы клетки);

2) нормальная жорданова форма J_A матрицы A определена однозначно (с точностью до порядка жордановых клеток).

Эта теорема обычно является одним из центральных результатов курса линейной алгебры.

Конечно, теорема Гамильтона—Кэли над полем $\mathbb{C}$ является следствием теоремы о жордановой нормальной форме. В то же время имеются элегантные доказательства теоремы о жордановой нормальной форме, использующие теорему Гамильтона—Кэли.

Упражнение 9.35. Покажите, что:

1) многочлен Эрмита

$$\mathrm{H}_n(t) = (-1)^n e^{t^2} \frac{d^n}{dt^n} \left(e^{-t^2}\right)$$

является собственным вектором линейного дифференциального оператора

$$\frac{d^2}{dt^2} - 2t\frac{d}{dt}$$

относительно собственного числа $-2n$;

2) многочлен Чебышёва
$$\mathrm{T}_n(t) = \cos(n \arccos t)$$
является собственным вектором линейного дифференциального оператора
$$(t^2-1)\frac{d^2}{dt^2} + t\frac{d}{dt}$$
относительно собственного значения n^2;

3) многочлен Лежандра
$$\mathrm{L}_n(t) = \frac{1}{2^n\, n!}\frac{d^n(t^2-1)^n}{dt^n} \quad (\mathrm{L}_0(t) = 1)$$
является собственным вектором линейного дифференциального оператора
$$(t^2-1)\frac{d^2}{dt^2} + 2t\frac{d}{dt}$$
относительно собственного числа $n(n+1)$.

§ 10. Евклидово пространство

Пусть $K = \mathbb{R}$ — поле действительных чисел, ${}_{\mathbb{R}}V$ — линейное пространство над полем $\mathbb{R}$. Тогда отображение
$$V \times V \to \mathbb{R}, \quad V \times V \ni (x, y) \to (x, y) \in \mathbb{R},$$
называется скалярным произведением на линейном пространстве ${}_{\mathbb{R}}V$, если для всех $x, y, z \in V$ и $\alpha \in K$ выполнены следующие условия:

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

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

3) $(\alpha x, y) = \alpha(x, y)$;

4) $(x, x) \geqslant 0$; $(x, x) = 0$ тогда и только тогда, когда $x = 0 \in V$.

Линейное пространство V над $\mathbb{R}$ с фиксированным скалярным произведением называется евклидовым пространством. Отметим, что в евклидовом пространстве V для $\mathbf{0} \in V$ имеем $(\mathbf{0}, x) = 0 \in \mathbb{R}$ для любого элемента $x \in V$: $(\mathbf{0}, x) = (0 \cdot \mathbf{0}, x) = 0 \cdot (\mathbf{0}, x) = 0 \in \mathbb{R}$.

Из свойств 1)—3) скалярного произведения следует, что $(x, y + z) = (x, y) + (x, z)$, $(x, \alpha y) = \alpha(x, y)$ для всех $x, y, z \in V$, $\alpha \in \mathbb{R}$.

Примеры 10.1.

1) Линейное пространство векторов на действительной плоскости со скалярным произведением векторов.

2) $V = \mathbb{R}^n$, для $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n) \in \hat{\mathbb{R}}^n$ положим
$$(X, Y) = x_1 y_1 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$
(это скалярное произведение будем называть стандартным).

3) $V = C[a, b]$ (линейное пространство всех непрерывных функций на отрезке $[a, b] \subseteq \mathbb{R}$), для $x(t), y(t) \in V$ положим

$$\big(x(t), y(t)\big) = \int\limits_a^b x(t)y(t)\,dt.$$

4) $V = \mathrm{M}_n(\mathbb{R})$, $(A, B) = \mathrm{tr}(AB^*)$ для $A, B \in \mathrm{M}_n(\mathbb{R})$.

Если $U \subseteq V$ — подпространство евклидова пространства V, то U — евклидово пространство (относительно скалярного произведения пространства V).

Определим *длину* (*модуль*) $|x|$ элемента x евклидова пространства V:

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

Если $(x, y) = 0$, то говорят, что элементы $x, y \in V$ *ортогональны* (обозначение: $x \perp y$).

Лемма 10.2 (теорема Пифагора). *Если $x, y \in V$, $(x, y) = 0$, то*

$$|x + y|^2 = |x|^2 + |y|^2.$$

Доказательство.

$$|x + y|^2 = (x + y, x + y) = (x, x) + (x, y) + (y, x) + (y, y) = |x|^2 + |y|^2,$$

поскольку $(y, x) = (x, y) = 0$. □

Лемма 10.3 (неравенство Коши—Буняковского). *Для всех $x, y \in V$*

$$|(x, y)| \leqslant |x| \cdot |y|.$$

Доказательство. Пусть t — переменная, принимающая действительные значения. Тогда $(x - ty, x - ty) \geqslant 0$ для всех $t \in \mathbb{R}$. Поэтому

$$(x - ty, x - ty) = (x, x) - 2t(x, y) + t^2(y, y) = |y|^2t^2 - 2(x, y)t + |x|^2 \geqslant 0.$$

Рассматривая последнее неравенство как неравенство для квадратного трёхчлена относительно переменной t, выполненное для всех $t \in \mathbb{R}$, получаем, что это условие равносильно тому, что

$$0 \geqslant \frac{D}{4} = (x, y)^2 - |x|^2 \cdot |y|^2.$$ □

Следствия 10.4.

1) *Если $V = \mathbb{R}^n$ и для $X = (x_1, \ldots, x_n), Y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ имеем*

$$(X, Y) = \sum_{i=1}^n x_i y_i,$$

то

$$\left|\sum_{i=1}^n x_i y_i\right|^2 \leqslant \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right);$$

2) *если* $V = C[a,b]$ *и для* $x(t), y(t) \in V$ *имеем*

$$\big(x(y), y(t)\big) = \int\limits_a^b x(t)y(t)\,dt,$$

то

$$\left|\int\limits_a^b x(t)y(t)\,dt\right|^2 \leqslant \left(\int\limits_a^b \big(x(t)\big)^2\,dt\right)\left(\int\limits_a^b \big(y(t)\big)^2\,dt\right).$$

Неравенство Коши—Буняковского даёт возможность определить угол φ между элементами x и y евклидова пространства V:

$$\varphi = \arccos\left(\frac{(x,y)}{|x|\,|y|}\right), \quad 0 \leqslant \varphi \leqslant \pi.$$

Лемма 10.5 (неравенство треугольника). *Для всех* $x, y \in V$

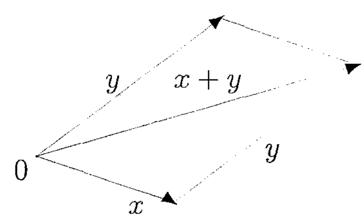
$$|x+y| \leqslant |x| + |y|.$$

Доказательство.

$$|x+y|^2 = (x+y, x+y) = |x|^2 + 2(x,y) + |y|^2 \leqslant |x|^2 + 2|x|\,|y| + |y|^2 = (|x|+|y|)^2$$

(здесь мы воспользовались неравенством Коши—Буняковского). □

Неравенство треугольника особенно наглядно в том случае, когда $V = \mathbb{R}^2$ (евклидово пространство всех векторов на действительной плоскости со стандартным скалярным произведением): $|x+y| \leqslant |x| + |y|$ — это классическое неравенство для длин сторон треугольника:



При этом теорема Пифагора (теорема 10.2) превращается в обычную школьную теорему Пифагора.

Неравенство треугольника в пространстве $C[a,b]$ имеет вид

$$\sqrt{\int\limits_a^b \big(f(t)+g(t)\big)^2\,dt} \leqslant \sqrt{\int\limits_a^b f^2(t)\,dt} + \sqrt{\int\limits_a^b g^2(t)\,dt}.$$

Лемма 10.6. *Попарно ортогональные ненулевые элементы евклидова пространства линейно независимы.*

Доказательство. Пусть $0 \neq v_1, \ldots, 0 \neq v_m \in V$, $(v_i, v_j) = 0$ при $1 \leqslant i \neq j \leqslant m$, $\alpha_1, \ldots, \alpha_m \in K$,

$$\alpha_1 v_1 + \ldots + \alpha_m v_m = 0 \in V.$$

Домножим скалярно последнее равенство на v_i $(1 \leqslant i \leqslant m)$:

$$\alpha_1(v_1, v_i) + \ldots + \alpha_m(v_m, v_i) = \left(\sum_{i=1}^{m} \alpha_i v_i, v_i\right) = (0, v_i) = 0 \in rrr.$$

Так как $(v_i, v_j) = 0$ при $i \neq j$, то $\alpha_i(v_i, v_i) = 0$. Но $v_i \neq 0$, поэтому $(v_i, v_i) > 0$. Следовательно, $\alpha_i = 0$ $(1 \leqslant i \leqslant m)$. □

Определение 10.7. Пусть $v_1, \ldots, v_m$ — попарно ортогональные ненулевые элементы евклидова пространства V. Если $|v_i| = 1$, $i = 1, 2, \ldots, m$, то систему элементов $\{v_1, \ldots, v_m\}$ будем называть ортонормированной.

Упражнение 10.8. $V = C[0, 2\pi]$.

$$(f(t), g(t)) = \int_0^{2\pi} f(t)g(t)\,dt, \quad n \in \mathbb{N},$$

$$e_0 = \frac{1}{\sqrt{2\pi}};\ e_1 = \frac{\cos t}{\sqrt{\pi}};\ e_2 = \frac{\sin t}{\sqrt{\pi}}, \ldots,\ e_{2n-1} = \frac{\cos nt}{\sqrt{\pi}},\ e_{2n} = \frac{\sin nt}{\sqrt{\pi}}.$$

Покажите, что $\{e_0, \ldots, e_{2n}\}$ — ортонормированная система элементов евклидова пространства V.

Теорема 10.9 (неравенство Бесселя). *Пусть $v_1, \ldots, v_n$ — ортонормированная система элементов евклидова пространства V, $x \in V$, $x_i = (x, v_i)$. Тогда:*

$$\sum_{i=1}^{n} |x_i|^2 \leqslant |x|^2;$$

$$\left(x - \sum_{i=1}^{n} x_i v_i, v_j\right) = 0 \ \text{ для } \ i = 1, \ldots, n.$$

Доказательство.

$$0 \leqslant \left(x - \sum_{i=1}^{n} x_i v_i, x - \sum_{i=1}^{n} x_i v_i\right) = |x|^2 - \sum_{i=1}^{n} |x_i|^2;$$

$$\left(x - \sum_{i=1}^{n} x_i v_i, v_j\right) = (x, v_j) - \sum_{i=1}^{n} x_i (v_i, v_j) = x_i - x_j = 0.$$ □

Замечание 10.10. Геометрический смысл неравенства Бесселя состоит в том, что квадрат длины вектора не меньше, чем сумма квадратов его проекций на любые k взаимно ортогональных направлений.

Замечание 10.11 (матрица Грама). Пусть $e_1, \ldots, e_n$ — базис евклидова пространства ${}_{\mathbb{R}}V$. матрица $\Gamma = (g_{ij}) \in \mathrm{M}_n(\mathbb{R})$, $g_{ij} = (e_i, e_j)$, $1 \leqslant i, j \leqslant n$, называется *матрицей Грама базиса* $e_1, \ldots, e_n$. Так как $(v_i, v_j) = (v_j, v_i)$, то $g_{ij} = g_{ji}$ для всех $1 \leqslant i, j \leqslant n$, и $\Gamma^* = \Gamma$ (матрица Γ симметрическая). Если $x, y \in V$, $X = (x_1, \ldots, x_n), Y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ — строки координат элементов x и y в базисе $e_1, \ldots, e_n$, то

$$(x, y) = X\Gamma\hat{Y}.$$

Матрица Грама *положительно определена* в том смысле, что $X\Gamma\hat{X} \geqslant 0$ для всех $X \in \mathbb{R}^n$, если $X\Gamma\hat{X} = 0$, то $X = 0 \in \mathbb{R}^n$. Наоборот, любая симметрическая положительно определённая матрица задаёт скалярное произведение и является матрицей Грама.

Если $\mathcal{E} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$, $\mathcal{E}' = \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}$ — другой базис пространства V, $\mathcal{E}' = C^*\mathcal{E}$ ($C \in \mathrm{GL}_n(\mathbb{R})$ — матрица перехода), то $\hat{X} = C\hat{X}'$, $\hat{Y} = C\hat{Y}'$ ($\hat{X}'$ и $\hat{Y}'$ — столбцы координат элементов x и y в базисе $e'_1, \ldots, e'_n$), и поэтому $(x, y) = X\Gamma\hat{Y} = X'C^*\Gamma C\hat{Y}'$. Следовательно, $\Gamma' = C^*\Gamma C$ — матрица Грама базиса $e'_1, \ldots, e'_n$.

Определение 10.12. Пусть V — евклидово пространство, $\dim V = n < \infty$. Базис $e_1, \ldots, e_n$ линейного пространства V называется *ортогональным*, если $(e_i, e_j) = 0$ для всех $1 \leqslant i \neq j \leqslant n$. Ортогональный базис называется *ортонормированным*, если $|e_i| = \sqrt{(e_i, e_i)} = 1$ для всех $i = 1, \ldots, n$.

Ортогональный базис легко превратить в ортонормированный: если $e_1, \ldots, e_n$ — ортогональный базис евклидова пространства V, то $e'_1 = \frac{1}{e_1}e_1, \ldots, e'_n = \frac{1}{|e_n|}e_n$ — ортонормированный базис пространства V.

Пример 10.13. $V = \mathbb{R}^n$ со стандартным скалярным произведением,

$$e_1 = (1, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots,\ e_n = (0, \ldots, 0, 1).$$

Тогда $e_1, \ldots, e_n$ — ортонормированный базис евклидова пространства V.

Теорема 10.14. *Пусть V — евклидово пространство, $\dim V = n < \infty$. Тогда существует ортонормированный базис пространства V.*

Доказательство (*алгоритм ортогонализации Грама—Шмидта*). Пусть $g_1, \ldots, g_n$ — базис линейного пространства V. Если $n = 1$, то положим $f = g_1$, $\{f_1\}$ — ортогональный базис пространства V.

Пусть $n \geqslant 2$. Тогда положим $f_1 = g_1$, $f_2 = g_2 + \alpha f_1$, $\alpha \in \mathbb{R}$, коэффициент α определяется однозначно из условия $(f_1, f_2) = 0$: $(g_2 + \alpha f_1, f_1) = (g_2, f_1) + \alpha(f_1, f_1) = 0$, следовательно, $\alpha = -\frac{(g_2, f_1)}{(f_1, f_1)}$ (так как f_1 — элемент базиса пространства V, то $f_1 \neq 0 \in V$, $(f_1, f_1) \neq 0$). Имеем: $(f_1, f_2) = 0$; $\langle f_1, f_2 \rangle = \langle g_1, g_2 \rangle$; $\dim\langle g_1, g_2 \rangle = 2$; следовательно, f_1, f_2 — ортогональный базис линейного подпространства $\langle g_1, g_2 \rangle$.

Продолжая процесс, допустим, что мы последовательно нашли такие линейно независимые элементы $f_1, \ldots, f_{k-1} \in V$, $k \leqslant n$, что

$$(f_i, f_j) = 0 \text{ при } 1 \leqslant i, j \leqslant k-1, \quad \langle f_1, \ldots, f_{k-1} \rangle = \langle g_1, \ldots, g_{k-1} \rangle.$$

Положим

$$f_k = g_k + \alpha_1 f_1 + \ldots + \alpha_{k-1} f_{k-1}.$$

Коэффициенты α_i, $1 \leqslant i \leqslant k-1$, однозначно находятся из условий

$$(f_k, f_i) = (g_k, f_i) + \alpha_i(f_i, f_i) = 0, \quad 1 \leqslant i \leqslant k-1.$$

Так как f_i — ненулевые элементы, $1 \leqslant i \leqslant k-1$, то

$$\alpha_i = -\frac{(g_k, f_i)}{(f_i, f_i)}, \quad 1 \leqslant i \leqslant k-1.$$

Итак, по построению $f_1, \ldots, f_k$ — попарно ортогональные элементы, $\langle f_1, \ldots, f_k \rangle = \langle g_1, \ldots, g_k \rangle$, $\dim\langle f_1, \ldots, f_k \rangle = \dim\langle g_1, \ldots, g_k \rangle = k$, и следовательно, $f_1, \ldots, f_k$ — ортогональный базис подпространства $\langle g_1, \ldots, g_k \rangle$. Если $k = n$, то мы построили ортогональный базис всего пространства. Если $k < n$, то повторяем описанную процедуру. Так как $n < \infty$, то за конечное число шагов мы построим ортогональный базис $\{f_1, \ldots, f_n\}$ евклидова пространства V. Положив $e_1 = \frac{1}{|f_1|}f_1, \ldots, e_n = \frac{1}{|f_n|}f_n$, получаем ортонормированный базис в V. □

Замечание 10.15. На любом конечномерном линейном пространстве ${}_{\mathbb{R}}V$ можно задать скалярное произведение (и ${}_{\mathbb{R}}V$ превратится в евклидово пространство): пусть $e_1, \ldots, e_n$ — базис линейного пространства V, для $x = x_1e_1 + \ldots + x_ne_n$, $y = y_1e_1 + \ldots + y_ne_n$, $x_i, y_j \in \mathbb{R}$, $1 \leqslant i, j \leqslant n$, положим $(x, y) = \sum\limits_{i=1}^{n} x_iy_i$. Тогда V — евклидово пространство, при этом $e_1, \ldots, e_n$ — ортонормированный базис евклидова пространства V.

Определение 10.16. Пусть ${}_{\mathbb{R}}U$ и ${}_{\mathbb{R}}V$ — евклидовы пространства. Эти пространства называются изоморфными ($U \cong V$), если существует такая биекция $f\colon U \to V$, что для всех $u, v \in U$, $\alpha \in \mathbb{R}$:

$$f(u+v) = f(u) + f(v);$$
$$f(\alpha u) = \alpha f(u);$$
$$(u,v) = \big(f(u), f(v)\big).$$

Теорема 10.17. *Пусть ${}_{\mathbb{R}}U$ и ${}_{\mathbb{R}}V$ — конечномерные евклидовы пространства. Тогда евклидовы пространства U и V изоморфны в том и только в том случае, когда $\dim_{\mathbb{R}} U = \dim_{\mathbb{R}} V$.*

Доказательство. Если $U \cong V$ как евклидовы пространства, то линейные пространства ${}_{\mathbb{R}}U$ и ${}_{\mathbb{R}}V$ изоморфны. Следовательно, по ?? $\dim_{\mathbb{R}} U = \dim_{\mathbb{R}} V$.

Если ${}_{\mathbb{R}}U$ и ${}_{\mathbb{R}}V$ — евклидовы пространства, $n = \dim_{\mathbb{R}} U = \dim_{\mathbb{R}} V$, то пусть $u_1, \dots, u_n$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}U$, $v_1, \dots, v_n$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$. Определим отображение $f\colon {}_{\mathbb{R}}U \to {}_{\mathbb{R}}V$, положив для любого $x \in U$, $x = x_1u_1 + \ldots + x_nu_n$, $x_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, $f(x) = x_1v_1 + \ldots + x_nv_n$. Ясно, что f — биективное линейное отображение линейных пространств над $\mathbb{R}$. Более того, для

$$x = \sum_{i=1}^{n} x_iu_i,\ \ y = \sum_{j=1}^{n} y_ju_j \in U, \quad x_i, y_j \in \mathbb{R}, \ \ 1 \leqslant i, j \leqslant n,$$

имеем

$$(x,y) = \sum_{i=1}^{n} x_iy_i = \left(\sum_{i=1}^{n} x_iv_i, \sum_{j=1}^{n} y_jv_j\right) = \big(f(x), f(y)\big).$$

Итак, f — изоморфизм евклидовых пространств. □

Замечание 10.18. Пусть $e_1, \dots, e_n$ — ортонормированный базис евклидова пространства V, $x \in V$, $x = x_1e_1 + \ldots + x_ne_n$, $x_i \in \mathbb{R}$. Тогда: $x_i = (x, e_i)$, $1 \leqslant i \leqslant n$, если $y = y_1e_1 + \ldots + y_ne_n$, то $(x,y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n = (X, Y)$, где $X = (x_1, \dots, x_n), Y = (y_1, \dots, y_n) \in \mathbb{R}^n$ и $(X, Y) = x_1y_1 + \ldots + x_ny_n$ — стандартное скалярное произведение строк в $\mathbb{R}^n$.

Замечание 10.19. Если $v_1, \dots, v_m \in V$, $U = \langle v_1, \dots, v_m \rangle \subseteq V$, то можно построить ортонормированный базис в подпространстве U следующим образом: во-первых,найти максимальную линейно независимую подсистему $v_{i_1}, \dots, v_{i_r}$ системы $v_1, \dots, v_m$ (для $V = \mathbb{R}^n$ см. ??), а затем применить процесс ортогонализации к базису $v_{i_1}, \dots, v_{i_r}$ подпространства U. Процесс ортогонализации можно применять только к линейно независимым системам элементов!

Пример 10.20.

1) Пусть $V = \mathbb{R}^4$ — евклидово пространство строк длины 4 со стандартным скалярным произведением,

$$u_1 = (2,1,3,-1),\ u_2 = (7,4,3,-3),\ u_3 = (1,1,-6,0),\ u_4 = (5,7,7,8),$$
$$U = \langle u_1, u_2, u_3, u_4 \rangle.$$

Построим ортонормированный базис евклидова пространства U. Сначала (см. ??) находим максимальную линейно независимую подсистему среди u_1, u_2, u_3, u_4 (базис подпространства U):

$$\begin{pmatrix} 2 & 7 & 1 & 5 \\ 1 & 4 & 1 & 7 \\ 3 & 3 & -6 & 7 \\ -1 & -3 & 0 & 8 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 1 & 7 \\ 2 & 7 & 1 & 5 \\ 3 & 3 & -6 & 7 \\ -1 & -3 & 0 & 8 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 1 & 7 \\ 0 & -1 & -1 & -9 \\ 0 & -9 & -9 & -14 \\ 0 & 1 & 1 & 15 \end{pmatrix} \to \begin{pmatrix} 1 & 4 & 1 & 7 \\ 0 & -1 & -1 & -9 \\ 0 & 0 & 0 & 67 \\ 0 & 0 & 0 & 6 \end{pmatrix};$$

следовательно, u_1, u_2, u_4 — базис подпространства U. Применим процесс ортогонализации к u_1, u_2, u_4:

$$v_1 = u, \quad v_2 = u_2 + \alpha v_1, \quad \alpha = -\frac{(u_2, v_1)}{(u_1, u_1)} = -\frac{30}{15} = -2;$$

таким образом,

$$v_2 = u_2 - 2u_1 = (3, 2, -3, -1).$$

Положим

$$v_3 = u_4 + \beta v_2 + \gamma v_1,$$

при этом

$$\gamma = \frac{(u_4, v_1)}{(v_1, v_1)} = -\frac{30}{15} = -2, \quad \beta = -\frac{(u_4, v_2)}{(v_2, v_2)} = -\frac{0}{(v_2, v_2)} = 0,$$

поэтому

$$v_3 = u_4 - 2v_1 = (1, 5, 1, 10).$$

Нормируя ортогональный базис v_1, v_2, v_3, получаем ортонормированный базис евклидова пространства U:

$$e_1 = \frac{1}{\sqrt{15}}(2, 1, 3, -1), \quad e_2 = \frac{1}{\sqrt{23}}(3, 2, -3, -1), \quad e_3 = \frac{1}{\sqrt{127}}(1, 5, 1, 10).$$

2) Многочлены Лежандра. Пусть $k \in \mathbb{N}$, V — евклидово пространство всех многочленов от переменной t степени $\leqslant k$ со скалярным произведением

$$\big(p(t), q(t)\big) = \int\limits_{-1}^{1} p(t)q(t)\,dt.$$

Применяя процесс ортогонализации к базису $1, t, t^2, \ldots, t^k$ пространства V, получаем ортогональный базис евклидова пространства V, $1, t, t^2 - \frac{1}{3}, t^3 - \frac{3}{5}t, \ldots$, состоящий из (с точностью до числовых множителей) многочленов Лежандра

$$\mathrm{L}_n(t) = \frac{1}{2^n\, n!}\,\frac{d^n}{dt^n}[(t^2-1)^n] = \sum_{m=0}^{n} \frac{(-1)^m\,(m+n)!}{2^m (m!)^2 (n-m)!}(1-t)^m, \quad 0 \leqslant n \leqslant k, \qquad \int\limits_{-1}^{1} \mathrm{L}_n^2(t)\,dt = \frac{2}{2n+1}.$$

Многочлены Лежандра L_n удовлетворяют дифференциальному уравнению

$$(1-t^2)\frac{d^2y}{dt^2} - 2t\frac{dy}{dt} + n(n+1)y = 0.$$

3) Многочлены Чебышёва первого рода

$$\mathrm{T}_n(t) = \frac{(-2)^n\, n!}{(2n)!}\sqrt{1-t^2}\,\frac{d^n}{dt^n}(1-t^2)^{n-\frac{1}{2}} = \cos(n \arccos t), \quad n \geqslant 0,$$
$$\mathrm{T}_0(t) = 1, \quad \mathrm{T}_1(t) = t, \quad \mathrm{T}_2(t) = 2t^2 - 1, \ldots,$$

образуют ортогональную систему относительно скалярного произведения

$$\big(\mathrm{T}_m(t), \mathrm{T}_n(t)\big) = \int\limits_{-1}^{1} \frac{\mathrm{T}_m(t)\mathrm{T}_n(t)}{\sqrt{1-t^2}}\,dt = \begin{cases} 0, & m \neq n, \\ \dfrac{\pi}{2}, & m = n \neq 0, \\ \pi, & m = n = 0. \end{cases}$$

Многочлены Чебышёва второго рода

$$\mathbf{U}_n(t) = \frac{1}{n+1}\frac{d\mathrm{T}_{n+1}(t)}{dt}$$

ортогональны относительно скалярного произведения

$$\int_{-1}^{1} \sqrt{1-t^2}\,\mathrm{U}_m(t)\mathrm{U}_n(t)\,dt = \begin{cases} 0, & m \neq n, \\ \dfrac{\pi}{2}, & m = n. \end{cases}$$

4) Многочлены Эрмита

$$\mathrm{H}_n(t) = (-1)^n e^{t^2}\frac{d^n}{dt^n}e^{-t^2},$$
$$\mathrm{H}_0(t) = 1, \quad \mathrm{H}_1(t) = 2t, \quad \mathrm{H}_2(t) = 4t^2 - 2, \dots,$$

образуют ортогональную систему относительно скалярного произведения

$$\int_{-\infty}^{\infty} e^{-t^2}\mathrm{H}_m(t)\mathrm{H}_n(t)\,dt = \begin{cases} 0, & m \neq n, \\ 2^n\, n!\,\sqrt{\pi}, & m = n. \end{cases}$$

Теорема 10.21 (равенство Парсеваля). *Пусть V — евклидово пространство, $v_1, \dots, v_k \in V$ — ортонормированная система элементов. Тогда следующие условия эквивалентны:*

1) $v_1, \dots, v_k$ *— ортонормированный базис евклидова пространства* V;

2) $x = \sum\limits_{i=1}^{m}(x, v_i)v_i$ *для любого* $x \in V$;

3) $(x, y) = \sum\limits_{i=1}^{m}(x, v_i)(v_i, y)$ *для всех* $x, y \in V$ *(равенство Парсеваля)*;

4) $|x|^2 = \sum\limits_{i=1}^{m}|(x, v_i)|^2$.

Доказательство.

1) $\Longrightarrow$ 2). Если $v_1, \dots, v_k$ — ортонормированный базис пространства V, то $x = x_1v_1 + \ldots + x_kv_k$, $x_i \in \mathbb{R}$. Умножая скалярно последнее равенство на v_j, получаем, что $(x, v_j) = x_j(v_j, v_j) = x_j$, $1 \leqslant j \leqslant n$.

2) $\Longrightarrow$ 3). Если для любого $x \in V$ имеем $x = \sum\limits_{i=1}^{m}(x, v_i)v_i$, то $y = \sum\limits_{j=1}^{m}(y, v_j)v_j$ для $y \in V$, и

$$(x, y) = \sum_{i=1}^{m}(x, v_i)(y, v_i) = \sum_{i=1}^{m}(x, v_i)(v_i, y).$$

3) $\Longrightarrow$ 4). Если в равенстве Парсеваля положить $x = y$, то получим

$$|x|^2 = (x, x) = \sum_{i=1}^{m}|(x, v_i)|^2.$$

4) $\Longrightarrow$ 1). Пусть $|x|^2 = \sum\limits_{i=1}^{m}|(x, v_i)|^2$ для любого $x \in V$. Допустим, что ортонормированная система $v_1, \dots, v_k$ не является базисом пространства V. Тогда существует такой элемент $0 \neq w \in V$,

что $v_1, \ldots, v_k, w$ — линейно независимые элементы. Применяя процесс ортогонализации к элементу $w' = w + \alpha_1 v_1 + \ldots + \alpha_k v_k$, получаем $\alpha_i = -\dfrac{(w, v_i)}{(v_i, v_i)} = -(w, v_i)$, $1 \leqslant i \leqslant k$. Полагая $v_{k+1} = \dfrac{1}{w'} w'$, получаем, что $v_1, \ldots, v_k, v_{k+1}$ — ортонормированная система. Если $x = v_{k+1}$, то

$$|x|^2 = 1 \neq 0 = \sum_{i=1}^{m} |(x, v_i)|^2,$$

тем самым пришли к противоречию. □

Теорема 10.22. *Пусть $m, n \in \mathbb{N}$, $m < n$, $v_1, \ldots, v_m$ — ортонормированная система евклидова пространства V, $\dim V = n < \infty$. Тогда существуют такие элементы $v_{m+1}, \ldots, v_n$, что $v_1, \ldots, v_n$ — ортонормированный базис евклидова пространства V.*

Доказательство. В силу **??** $\{v_1, \ldots, v_m\}$ — линейно независимая система. Пусть $w_{m+1}, \ldots, w_n$ — такие элементы линейного пространства V, что $v_1, \ldots, v_m, w_{m+1}, \ldots, w_n$ — базис пространства V (см. **??**). Применим к этому базису процесс ортогонализации. Так как $v_1, \ldots, v_m$ — ортонормированная система, то после применения процесса ортогонализации мы получим ортогональный базис $v_1, \ldots, v_m, f_{m+1}, \ldots, f_n$ (первые m элементов не будут меняться). Полагая $v_{m+1} = \dfrac{1}{|f_{m+1}|} f_{m+1}, \ldots, v_n = \dfrac{1}{|f_n|} f_n$, получаем ортонормированный базис $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ пространства V. □

Замечание 10.23.

1) Матрица Грама Γ ортонормированного базиса $e_1, \ldots, e_n$ евклидова пространства V единичная, $\Gamma = E \in \mathrm{M}_n(\mathbb{R})$. Если $e'_1, \ldots, e'_n$ — другой базис пространства V (не обязательно ортонормированный), то $\Gamma' = C^* E C = C^* C$, где C — матрица перехода от базиса $e_1, \ldots, e_n$ к базису $e'_1, \ldots, e'_n$ (см. замечание **??**). Поэтому

$$|\Gamma'| = |C^* C| = |C^*|\,|C| = |C|^2 > 0$$

для любого базиса пространства V.

2) Если рассмотреть последовательно подпространства $V_1 = \langle e'_1 \rangle$, $V_2 = \langle e'_1, e'_2 \rangle, \ldots$, $V_{n-1} = \langle e'_1, \ldots, e'_{n-1} \rangle$ пространства $V = V_n = \langle e'_1, \ldots, e'_n \rangle$ и повторить приведённые рассуждения для матриц Грама базисов этих подпространств, воспользовавшись существованием ортонормированных базисов подпространств V_i, $1 \leqslant i \leqslant n-1$, то мы получим, что все угловые миноры матрицы Γ' положительны. Позже (**??**) мы увидим, что это условие для симметрической матрицы A является и достаточным для того, чтобы матрица A задавала некоторое скалярное произведенис, $A = \Gamma$ (это условие эквивалентно тому, что матрица A положительно определена).

3) Если $\mathcal{E} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$ и $\mathcal{E}' = \begin{pmatrix} e'_1 \\ \vdots \\ e'_n \end{pmatrix}$ — ортогональные базисы евклидова пространства V, $\mathcal{E}' = C^* \mathcal{E}$, $C \in \mathrm{GL}_n(\mathbb{R})$ — матрица перехода, то матрица Грама в обоих базисах единичная, поэтому $C^* C = E$ (т. е. матрица C ортогональная, $C^* = C^{-1}$). Обратно, если фиксирован ортонормированный базис $\mathcal{E}$ и матрица перехода от базиса $\mathcal{E}$ к базису $\mathcal{E}'$ ортогональная, то $\mathcal{E}'$ — ортонормированный базис ($\Gamma' = C^* E C = E$).

Предложение 10.24. *Пусть $v_1, \ldots, v_m$ — элементы евклидова пространства V, $\Gamma(v_1, \ldots, v_m) = (a_{ij}) \in \mathrm{M}_m(\mathbb{R})$, где $a_{ij} = (v_i, v_j)$, $1 \leqslant i, j \leqslant m$. Тогда если элементы $v_1, \ldots, v_m$ линейно независимы, то $|\Gamma(v_1, \ldots, v_m)| > 0$. Если же элементы $v_1, \ldots, v_m$ линейно зависимы, то $|\Gamma(v_1, \ldots, v_m)| = 0$. Итак, $|\Gamma(v_1, \ldots, v_m)| = 0$ тогда и только тогда, когда $\{v_1, \ldots, v_m\}$ — линейно зависимая система.*

Доказательство.

1) Первая часть утверждения уже доказана выше (достаточно рассмотреть евклидово пространство $V' = \langle v_1, \ldots, v_m \rangle$ с базисом $e_1 = v_1, \ldots, e_m = v_m$).

2) Допустим теперь, что элементы $v_1, \ldots, v_m$ линейно зависимы. Тогда

$$\alpha_1 v_1 + \ldots + \alpha_m v_m = 0 \in V, \quad (\alpha_1, \ldots, \alpha_m) \neq 0 \in \mathbb{R}^m.$$

Умножая последовательно последнее равенство слева скалярно на $v_1, \ldots, v_m$, получаем, что $(\alpha_1, \ldots, \alpha_m)$ — ненулевое решение однородной квадратной системы линейных уравнений

$$\Gamma(v_1, \ldots, v_m)\hat{X} = 0 \in \mathbb{R}^m,$$

где $X = (x_1, \ldots, x_m)$. Следовательно, $|\Gamma(v_1, \ldots, v_m)| = 0$. □

Следствие 10.25. *При $k = 2$ для $x, y \in V$ (V — евклидово пространство) имеем*

$$\begin{vmatrix} (x,x) & (x,y) \\ (y,x) & (y,y) \end{vmatrix} = (x,x)(y,y) - (x,y)(y,x) = |x|^2|y|^2 - (x,y)^2 \geqslant 0$$

(неравенство Коши—Буняковского).

Замечание 10.26.

1) Если в процессе ортогонализации (без нормировки ортогонального базиса) мы перешли от линейно независимых элементов $g_1, \ldots, g_m$ к элементам $f_1, \ldots, f_m$ (ортогональный базис подпространства $\langle g_1, \ldots, g_m \rangle$), то $\Gamma(g_1, \ldots, g_m) = \Gamma(f_1, \ldots, f_m) = \prod\limits_{i=1}^{m} (f_i, f_i)$.

2) $\Gamma(g_1, \ldots, g_m) \leqslant \prod\limits_{i=1}^{m} (g_i, g_i)$. Действительно, в процессе ортогонализации

$$g_j = f_j + \alpha_{j1} f_1 + \ldots + \alpha_{j,j-1} f_{j-1}, \quad j = 2, \ldots, m.$$

Поэтому

$$(g_j, g_j) = (f_j, f_j) + \sum_{i=1}^{j-1} \alpha_{ii}^2 (f_i, f_i) \geqslant (f_j, f_j).$$

3) Если V — евклидово пространство, $v_1, \ldots, v_m \in V$, то квадрат m-мерного объёма параллелепипеда, построенного на элементах $v_1, \ldots, v_m$ равен $\Gamma(v_1, \ldots, v_m)$.

4) Неравенство Адамара. Пусть $A \in \mathrm{M}_n(\mathbb{R})$. Тогда

$$|A|^2 \leqslant \prod_{i=1}^{n} (a_{i1}^2 + \ldots + a_{in}^2).$$

Пусть $V = \mathbb{R}^n$ — евклидово пространство со стандартным скалярным произведением, $A_1, \ldots, A_n \in \mathbb{R}^n$ — строки матрицы A. Тогда $\Gamma(A_1, \ldots, A_n) = |AA^*|$. Применяя пункт 2), получаем

$$\Gamma(A_1, \ldots, A_n) \leqslant \prod_{i=1}^{n} (A_i, A_i) = \prod_{i=1}^{n} (a_{i1}^2 + a_{i2}^2 + \ldots + a_{in}^2).$$

Но $|AA^*| = |A| \cdot |A^*| = |A|^2$.

Геометрический смысл неравенства Адамара состоит в том, что объём параллелепипеда не превосходит произведения длин его рёбер; он равен этому произведению в том и только в том случае, когда рёбра взаимно ортогональны.

10.1. Ортогональное дополнение

Определение 10.27. Пусть V_1, V_2 — подпространства евклидова пространства ${}_{\mathbb{R}}V$. Подпространства V_1 и V_2 называются **ортогональными** (обозначение: $V_1 \perp V_2$), если $(v_1, v_2) = 0$ для всех $v_1 \in V_1$, $v_2 \in V_2$.

Предложение 10.28. *Пусть V_1, V_2 — подпространства евклидова пространства ${}_{\mathbb{R}}V$. Тогда:*

1) *если $e_1, \dots, e_m$ — базис пространства V_1, $f_1, \dots, f_n$ — базис пространства V_2, то условие $V_1 \perp V_2$ равносильно тому, что $(e_i, f_j) = 0$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$;*

2) *если $V_1 \perp V_2$, то $V_1 \cap V_2 = \{0\}$.*

Доказательство.

1) Если $V_1 \perp V_2$, то $(e_i, f_j) = 0$ для всех i, j. Допустим, что $(e_i, f_j) = 0$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$. Тогда для $v_1 \in V_1$, $v_2 \in V_2$ имеем $v_1 = x_1 e_1 + \ldots + x_m e_m$, $v_2 = y_1 f_1 + \ldots + y_n f_n$, $x_i, y_j \in \mathbb{R}$,

$$(v_1, v_2) = \sum_{i,j} x_i y_j (e_i, f_j) = 0.$$

2) Пусть $x \in V_1 \cap V_2$. Так как $V_1 \perp V_2$, то $(x, x) = 0$, следовательно, $x = 0$. □

Определение 10.29. Пусть U — подпространство евклидова пространства V ($\dim V = n < \infty$), $0 \neq U \subseteq V$, $U \neq V$. Рассмотрим подмножество

$$U^\perp = \{v \in V \mid (u, v) = 0 \text{ для всех } u \in U\} \subseteq V.$$

Заметим, что $U^\perp$ — подпространство линейного пространства V: если $(u, v_1) = 0$ и $(u, v_2) = 0$ для всех $u \in U$, то $(u, v_1 + v_2) = (u, v_1) + (u, v_2) = 0$; имеем $(u, \alpha v_1) = \alpha(u, v_1) = 0$ для всех $u \in U$, $\alpha \in \mathbb{R}$. Более того, $V = U \oplus U^\perp$. Действительно, $U^\perp \perp U$, и по предложению ?? $U^\perp \cap U = \{0\}$. Пусть $v_1, \dots, v_m$ — ортонормированный базис евклидова пространства V. Для любого $x \in V$ имеем $x = x_1 v_1 + \ldots + x_n v_n$. Пусть $U = \langle u_1, \dots, u_m \rangle$, $u_i = a_{i_1} v_1 + \ldots + a_{i_n} v_m$, $1 \leqslant i \leqslant m$. Тогда условие $(u, x) = 0$ для всех $u \in U$, определяющее пространство $U^\perp$, эквивалентно однородной системе

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \hat{\mathbb{R}}^m, \quad A = (a_{ij}) \in \mathrm{M}_{m,n}(\mathbb{R}), \quad x = x_1 v_1 + \ldots + x_n v_n. \tag{16}$$

При этом $\mathrm{r}(A) = \dim U = m$, и $\dim U^\perp$ совпадает с размерностью пространства решений однородной системы линейных уравнений, равной $n - \mathrm{r}(A)$. Следовательно, $n = \dim V = \dim U + \dim U^\perp$, и поэтому $V = U \oplus U^\perp$.

Замечание 10.30.

1) Подпространство $U^\perp$ евклидова пространства V называется ортогональным дополнением к подпространству U. Отметим, что, так как $V = U \oplus U^\perp$, то $\dim V = \dim U + \dim U^\perp$, следовательно, $\dim U^\perp = \dim V - \dim U$. Ортогональное дополнение находится достаточно просто: если $U = \langle u_1, \dots, u_m \rangle \subseteq V$ — линейное подпространство, являющееся линейной оболочкой конечного числа элементов, то в качестве базиса ортогонального дополнения $U^\perp$ достаточно взять фундаментальную систему решений однородной системы линейных уравнений (16).

2) Если $e_1, \dots, e_s$ — ортонормированный базис подпространства U, $e_{s+1}, \dots, e_n$ — ортонормированный базис ортогонального дополнения $U^\perp$, то $e_1, \dots, e_s, e_{s+1}, \dots, e_n$ — ортонормированный базис пространства V.

Лемма 10.31. $(U^\perp)^\perp = U$.

Доказательство. Так как $(u, v) = 0$ для всех $u \in U$, $v \in U^{\perp}$, то $U \subseteq (U^{\perp})^{\perp}$. Но $\dim U^{\perp} = \dim V - \dim U$, $\dim(U^{\perp})^{\perp} = \dim V - \dim(U^{\perp}) = \dim V - (\dim V - \dim U) = \dim U$. По лемме ?? $U = (U^{\perp})^{\perp}$. □

Замечание 10.32. Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $U \subseteq V$ — подпространство евклидова пространства ${}_{\mathbb{R}}V$, $0 \neq U \neq V$. Так как ${}_{\mathbb{R}}V = U \oplus U^{\perp}$, то элемент $v \in {}_{\mathbb{R}}V$ может быть однозначно представлен в виде $v = u + w$, где $u \in U$, $w \in U^{\perp}$.

Элемент u называется *ортогональной проекцией элемента v на подпространство U*. Если $v \notin U$, то $w \neq 0$. При этом, по теореме ??, если $z \in U$, $z \neq u$, то $v - u = w \in U^{\perp}$, $u - z \in U$, и поэтому

$$|v - z|^2 = |(v - u) + (u - z)|^2 = |v - u|^2 + |u - z|^2 > |v - u|^2 = |w|^2,$$

следовательно, $|v - z| > |w|$. Таким образом, $|w|$ — кратчайшее расстояние от v до U.

Пусть $u_1, \dots, u_n$ — базис линейного подпространства U, $v \in V$, $v = u + w$, $u \in U$, $w \in U^{\perp}$. Тогда $u = b_1u_1 + \dots + b_nu_n$, $b_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$. Коэффициенты b_i находим из условия $w = v - u \in U^{\perp}$. Это условие равносильно тому, что $(u, u_i) = (v, u_i)$, $1 \leqslant i \leqslant n$, то есть $b_1, \dots, b_n$ — единственное решение системы линейных уравнений

$$\Gamma(u_1, \dots, u_n) \cdot \hat{B} = \begin{pmatrix} (v, u_1) \\ \vdots \\ (v, u_n) \end{pmatrix},$$

где $\Gamma(u_1, \dots, u_n) = \big((u_i, u_j)\big) \in \mathrm{M}_n(K)$ — симметрическая матрица Грама базиса $u_1, \dots, u_n$.

Если $u_1, \dots, u_n$ — ортонормированный базис подпространства U, то $\Gamma(u_1, \dots, u_n)$ — единичная $(n \times n)$-матрица, и поэтому $b_i = (v, u_i)$, $1 \leqslant i \leqslant n$.

10.2. Линейные функционалы на евклидовых пространствах

Лемма 10.33. *Пусть V — евклидово пространство, $u, v \in V$, $(x, u) = (x, v)$ для всех $x \in V$. Тогда $u = v$.*

Доказательство. Так как $(x, u) = (x, v)$ для всех $x \in V$, то $(x, u - v) = 0$ для всех $x \in V$. Положим $x = u - v$. Тогда $(u - v, u - v) = 0$, следовательно, $u - v = 0$, $u = v$. □

Пусть V — линейное пространство над полем K. Линейное отображение $f\colon {}_KV \to {}_KK$ называется *линейным функционалом* на линейном пространстве ${}_KV$. Если $\dim V = n < \infty$, $e_1, \dots, e_n$ — базис пространства V, то для любого $x \in V$ имеем $x = x_1e_1 + \dots + x_ne_n$, $x_i \in K$, $f(x) = x_1f(e_1) + \dots + x_nf(e_n)$. Обозначая $a_i = f(e_i) \in K$, получаем, что $f(x) = x_1a_1 + \dots + x_na_n$ — линейная форма от переменных $x_1, \dots, x_n$.

Если V — евклидово пространство, $f\colon V \to \mathbb{R}$ — линейный функционал на пространстве ${}_{\mathbb{R}}V$, $e_1, \dots, e_n$ — ортонормированный базис пространства ${}_{\mathbb{R}}V$, то для всех $x = \sum\limits_{i=1}^{n} x_ie_i \in V$ $f(x) = \sum\limits_{i=1}^{n} x_ia_i = (x, a)$, где $a = \sum\limits_{i=1}^{n} a_ie_i \in {}_{\mathbb{R}}V$. Элемент $a \in V$ определён однозначно: если $a, b \in V$, $f(x) = (x, a) = (x, b)$ для всех $x \in V$, то по лемме ?? $a = b$.

10.3. Сопряжённый оператор

Пусть $\mathcal{A}$ — фиксированный линейный оператор на евклидовом пространстве ${}_{\mathbb{R}}V$, $\dim V = n < \infty$. Для каждого элемента $y \in V$ рассмотрим отображение $f_y\colon V \to \mathbb{R}$, $f_y(x) = (\mathcal{A}(x), y)$.

Покажем, что для каждого $y \in V$ отображение f_y — линейный функционал на пространстве V: для $u, v \in V$ имеем

$$f_y(u+v) = (\mathcal{A}(u+v), y) = (\mathcal{A}(u) + \mathcal{A}(v), y) = (\mathcal{A}(u), y) + (\mathcal{A}(v), y) = f_y(u) + f_y(v),$$

аналогично, для $\alpha \in \mathbb{R}$, $x \in V$

$$f_y(\alpha x) = (\mathcal{A}(\alpha x), y) = (\alpha\mathcal{A}(x), y) = \alpha(\mathcal{A}(x), y) = \alpha f_y(x).$$

Функционал f_y на евклидовом пространстве имеет однозначное представление в виде

$$f_y(x)\big(= (\mathcal{A}(x), y)\big) = \big(x, z(y)\big) \tag{17}$$

для всех $x \in V$, где элемент $z(y)$ зависит только от $y \in V$.

Обозначим через $\mathcal{A}^*$ отображение

$$\mathcal{A}^*\colon V \to V, \quad \mathcal{A}^*(y) = z(y), \quad y \in V,$$

элемент $z(y)$ однозначно определён условием (17). Покажем, что $\mathcal{A}^*$ — линейный оператор на пространстве V: для всех $x, y, v \in V$, $\alpha \in \mathbb{R}$

$$\begin{aligned}\big(x, \mathcal{A}^*(u+v)\big) = (\mathcal{A}(x), u+v) = (\mathcal{A}(x), u) + (\mathcal{A}(x), v) = \\ = \big(x, \mathcal{A}^*(u)\big) + \big(x, \mathcal{A}^*(v)\big) = \big(x, \mathcal{A}^*(u) + \mathcal{A}^*(v)\big),\end{aligned}$$

и следовательно, по лемме **??** $\mathcal{A}^*(u+v) = \mathcal{A}^*(u) + \mathcal{A}^*(v)$ для всех $u, v \in V$;

$$\big(x, \mathcal{A}^*(\alpha v)\big) = (\mathcal{A}(x), \alpha v) = \alpha(\mathcal{A}(x), v) = \alpha\big(x, \mathcal{A}^*(v)\big) = \big(x, \alpha\mathcal{A}^*(v)\big),$$

и по лемме **??** $\mathcal{A}^*(\alpha v) = \alpha\mathcal{A}^*(v)$ для всех $\alpha \in \mathbb{R}$, $v \in V$.

Линейный оператор $\mathcal{A}^*$ однозначно определён условием $(\mathcal{A}(x), y) = \big(x, \mathcal{A}^*(y)\big)$ для всех $x, y \in V$. Действительно, если $\mathcal{B}, \mathcal{C}$ — такие линейные операторы на пространстве V, что $\big(x, \mathcal{B}(y)\big) = \big(x, \mathcal{C}(y)\big)$ для всех $x, y \in V$, то по лемме **??** $\mathcal{B}(y) = \mathcal{C}(y)$ для всех $y \in V$. Следовательно, $\mathcal{B} = \mathcal{C}$.

Определение 10.34. Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $\dim_{\mathbb{R}} V < \infty$, $\mathcal{A}\colon V \to V$ — линейный оператор, $\mathcal{A}^*$ — построенный линейный оператор на пространстве V, однозначно определённый условием

$$(\mathcal{A}(x), y) = \big(x, \mathcal{A}^*(y)\big)$$

для всех $x, y \in V$. Линейный оператор $\mathcal{A}^*$ называется *сопряжённым оператором* к линейному оператору $\mathcal{A}$.

Пример 10.35. Если $\mathcal{I}\colon V \to V$ — тождественный оператор, то $\mathcal{I}^* = \mathcal{I}$: $(\mathcal{I}(x), y) = \big(x, \mathcal{I}(y)\big)$ для всех $x, y \in V$.

Лемма 10.36. *Пусть $e_1, \ldots, e_n$ — ортонормированный базис евклидова пространства V, $\mathcal{A}\colon V \to V$ — линейный оператор, A — матрица оператора в базисе $e_1, \ldots, e_n$, $\mathcal{A}^*$ — сопряжённый оператор. Тогда матрица оператора $\mathcal{A}^*$ в базисе $e_1, \ldots, e_n$ равна транспонированной матрице A^*.*

Доказательство. Пусть $A = (a_{ij})$, $a_{ij} \in \mathbb{R}$, $\mathcal{A}_1$ — линейный оператор, имеющий матрицу A^* в базисе $e_1, \ldots, e_n$. Тогда для всех i, j, $1 \leqslant i, j \leqslant n$,

$$(\mathcal{A}(e_i), e_j) = (a_{1i}e_1 + a_{2i}e_2 + \ldots + a_{ni}e_n, e_j) = a_{ji},$$
$$\big(e_i, \mathcal{A}_1(e_j)\big) = (e_i, a_{j1}e_1 + a_{j2}e_2 + \ldots + a_{jn}e_n) = a_{ji},$$

и следовательно, для $x = \sum\limits_{i=1}^{n} x_i e_i$, $y = \sum\limits_{j=1}^{n} y_j e_j$, $x_i, y_j \in \mathbb{R}$,

$$\big(x, \mathcal{A}^*(y)\big) = (\mathcal{A}(x), y) = \sum_{i,j=1}^{n} x_i y_j (\mathcal{A}(e_i), e_j) = \sum_{i,j=1}^{n} x_i y_j \big(e_i, \mathcal{A}_1(y)\big) = \big(x, \mathcal{A}_1(y)\big).$$

По лемме **??** $\mathcal{A}_1 = \mathcal{A}^*$. $\square$

Замечание 10.37. Если базис не является ортонормированным, то матрица оператора $\mathcal{A}^*$ может быть отлична от транспонированной матрицы A^*, где A — матрица оператора $\mathcal{A}$ в этом базисе. Например, пусть $V = \mathbb{R}^3$ — евклидово пространство строк со стандартным скалярным произведением, $\mathcal{A}\colon V \to V$ — линейный оператор, имеющий матрицу

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

в базисе $v_1 = (1,2,1)$, $v_2 = (1,1,2)$, $v_3 = (1,1,0)$ пространства V. Покажите, что матрица сопряжённого оператора $\mathcal{A}^*$ в базисе v_1, v_2, v_3 отлична от A^*.

Теорема 10.38 (свойства сопряжённого оператора). *Пусть V — конечномерное евклидово пространство, $\mathcal{A}$, $\mathcal{B}$ — линейные операторы на линейном пространстве V, $\alpha \in \mathbb{R}$. Тогда:*

1) $(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$;

2) $(\alpha\mathcal{A})^* = \alpha\mathcal{A}^*$;

3) $\mathcal{A}^{**} = \mathcal{A}$;

4) $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$.

Доказательство.
1) Для всех $x, y \in V$:

$$\big(x, (\mathcal{A}+\mathcal{B})^*(y)\big) = \big((\mathcal{A}+\mathcal{B})(x), y\big) = (\mathcal{A}(x) + \mathcal{B}(x), y) = \\ = (\mathcal{A}(x), y) + (\mathcal{B}(x), y) = \big(x, \mathcal{A}^*(y)\big) + \big(x, \mathcal{B}^*(y)\big) = \big(x, \mathcal{A}^*(y) + \mathcal{B}^*(y)\big).$$

По лемме **??** $(\mathcal{A}+\mathcal{B})^*(y) = \mathcal{A}^*(y) + \mathcal{B}^*(y) = (\mathcal{A}^* + \mathcal{B}^*)(y)$ для всех $y \in V$, следовательно, $(\mathcal{A}+\mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$.

2) Для всех $\alpha \in \mathbb{R}$, $x, y \in V$ имеем:

$$\big(x, (\alpha\mathcal{A})^*(y)\big) = \big((\alpha\mathcal{A})(x), y\big) = (\alpha\mathcal{A}(x), y) = \alpha(\mathcal{A}(x), y) = \alpha\big(x, \mathcal{A}^*(y)\big) = \big(x, \alpha\mathcal{A}^*(y)\big) = \big(x, (\alpha\mathcal{A}^*)(y)\big).$$

По лемме **??** $(\alpha\mathcal{A})^*(y) = (\alpha\mathcal{A}^*)(y)$ для всех $y \in V$. Следовательно, $(\alpha\mathcal{A})^* = \alpha\mathcal{A}^*$.

3) Для всех $x, y \in V$ имеем:

$$\big(x, \mathcal{A}^{**}(y)\big) = (\mathcal{A}^*(x), y) = \big(y, \mathcal{A}^*(x)\big) = (\mathcal{A}(y), x) = \big(x, \mathcal{A}(y)\big),$$

следовательно, по лемме **??** $\mathcal{A}^{**} = \mathcal{A}$.

4) Для всех $x, y \in V$ имеем:

$$\big(x, (\mathcal{A}\mathcal{B})^*(y)\big) = \big((\mathcal{A}\mathcal{B})(x), y\big) = \big(\mathcal{A}(\mathcal{B}(x)), y\big) = \big(\mathcal{B}(x), \mathcal{A}^*(y)\big) = \big(x, \mathcal{B}^*(\mathcal{A}^*(y))\big) = \big(x, (\mathcal{B}^*\mathcal{A}^*)(y)\big),$$

и по лемме **??** $(\mathcal{A}\mathcal{B})^*(y) = (\mathcal{B}^*\mathcal{A}^*)(y)$ для всех $y \in V$, следовательно, $(\mathcal{A}\mathcal{B})^* = \mathcal{B}^*\mathcal{A}^*$. □

Следствие 10.39. *Пусть $A, B \in \mathrm{M}_n(\mathbb{R})$. Тогда $(AB)^* = B^*A^*$.*

Следствие 10.40. *Пусть $\mathcal{A}$ — обратимый оператор на конечномерном евклидовом пространстве V. Тогда $(\mathcal{A}^{-1})^* = (\mathcal{A}^*)^{-1}$.*

Доказательство. $\mathcal{A}\mathcal{A}^{-1} = \mathcal{I} = \mathcal{A}^{-1}\mathcal{A}$. Следовательно, $\mathcal{I} = \mathcal{I}^* = (\mathcal{A}\mathcal{A}^{-1})^* = (\mathcal{A}^{-1})^*\mathcal{A}^*$, $\mathcal{I} = \mathcal{I}^* = \mathcal{A}^*(\mathcal{A}^{-1})^*$, поэтому $(\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^*$. □

Лемма 10.41. *Пусть $\mathcal{A}$ — линейный оператор на конечномерном евклидовом пространстве V. Тогда линейные операторы $\mathcal{A}$ и $\mathcal{A}^*$ имеют один и тот же характеристический многочлен.*

Доказательство. Характеристический многочлен оператора совпадает с характеристическим многочленом матрицы этого оператора в любом базисе. Пусть $e_1, \dots, e_n$ — ортонормированный базис евклидова пространства V, A — матрица оператора $\mathcal{A}$ в этом базисе. По лемме **??** A^* — матрица оператора $\mathcal{A}^*$ в базисе $e_1, \dots, e_n$. Следовательно,

$$p_{\mathcal{A}}(t) = |A - tE| = |(A - tE)^*| = |A^* - tE| = p_{\mathcal{A}^*}(t). \qquad \square$$

Лемма 10.42. *Пусть U — инвариантное подпространство для линейного оператора $\mathcal{A}$ на евклидовом пространстве V ($\dim V < \infty$, $0 \neq U \neq V$). Тогда $U^\perp$ — инвариантное подпространство для линейного оператора $\mathcal{A}^*$.*

Доказательство. Для всех $x \in U^\perp$, $y \in U$ имеем $(\mathcal{A}^*(x), y) = \big(x, \mathcal{A}(x)\big) = 0$, следовательно, $\mathcal{A}^*(x) \in U^\perp$ для всех $x \in U^\perp$. $\square$

Определение 10.43. Линейный оператор $\mathcal{A}$ на конечномерном евклидовом пространстве V называется *самосопряжённым*, если $\mathcal{A} = \mathcal{A}^*$.

Лемма 10.44.

1) *$\mathcal{O}$ (нулевой оператор), $\mathcal{I}$ (тождественный оператор) — самосопряжённые операторы.*

2) *Множество всех самосопряжённых операторов образует линейное пространство над полем $\mathbb{R}$.*

Доказательство.
1) Для всех $x, y \in V$ имеем:

$$\big(\mathcal{O}(x), y\big) = 0 = \big(x, \mathcal{O}(y)\big); \quad \big(\mathcal{I}(x), y\big) = (x, y) = \big(x, \mathcal{I}(y)\big).$$

2) Если $\mathcal{A} = \mathcal{A}^*$, $\mathcal{B} = c\mathcal{B}^*$, $\alpha \in \mathbb{R}$, то:

$$(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^* = \mathcal{A} + \mathcal{B}, \quad (\alpha\mathcal{A})^* = \alpha\mathcal{A}^* = \alpha\mathcal{A}. \qquad \square$$

Предложение 10.45. *Если $\mathcal{A}$, $\mathcal{B}$ — самосопряжённые линейные операторы, то $\mathcal{AB}$ — самосопряжённый линейный оператор в том и только в том случае, когда $\mathcal{AB} = \mathcal{BA}$.*

Доказательство. Так как $(\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{BA}$, то $(\mathcal{AB})^* = \mathcal{AB}$ тогда и только тогда, когда $\mathcal{AB} = \mathcal{BA}$. $\square$

Замечание 10.46. По лемме **??** матрица A самосопряжённого оператора $\mathcal{A}$ в ортонормированном базисе симметрическая: $A^* = A$. Если фиксировать ортонормированный базис $e_1, \dots, e_n$ евклидова пространства V, то любой симметрической матрице $A \in \mathrm{GL}_n(\mathbb{R})$ однозначно соответствует самосопряжённый оператор $\mathcal{A}$, имеющий матрицу A в этом базисе.

Лемма 10.47. *Пусть $\mathcal{A}$ — самосопряжённый линейный оператор на конечномерном евклидовом пространстве V, U — инвариантное подпространство для оператора $\mathcal{A}$, $0 \neq U \subseteq V$, $U \neq V$. Тогда $U^\perp$ также инвариантное подпространство.*

Доказательство. По лемме **??** $U^\perp$ — инвариантное подпространство для линейного оператора $\mathcal{A}^* = \mathcal{A}$. $\square$

Лемма 10.48. *Собственные векторы самосопряжённого оператора евклидова пространства относительно различных собственных значений ортогональны.*

Лемма 10.41. *Пусть $\mathcal{A}$ — линейный оператор на конечномерном евклидовом пространстве V. Тогда линейные операторы $\mathcal{A}$ и $\mathcal{A}^*$ имеют один и тот же характеристический многочлен.*

Доказательство. Характеристический многочлен оператора совпадает с характеристическим многочленом матрицы этого оператора в любом базисе. Пусть $e_1, \dots, e_n$ — ортонормированный базис евклидова пространства V, A — матрица оператора $\mathcal{A}$ в этом базисе. По лемме **??** A^* — матрица оператора $\mathcal{A}^*$ в базисе $e_1, \dots, e_n$. Следовательно,

$$p_{\mathcal{A}}(t) = |A - tE| = |(A - tE)^*| = |A^* - tE| = p_{\mathcal{A}^*}(t). \qquad \square$$

Лемма 10.42. *Пусть U — инвариантное подпространство для линейного оператора $\mathcal{A}$ на евклидовом пространстве V ($\dim V < \infty$, $0 \neq U \neq V$). Тогда $U^\perp$ — инвариантное подпространство для линейного оператора $\mathcal{A}^*$.*

Доказательство. Для всех $x \in U^\perp$, $y \in U$ имеем $(\mathcal{A}^*(x), y) = \big(x, \mathcal{A}(x)\big) = 0$, следовательно, $\mathcal{A}^*(x) \in U^\perp$ для всех $x \in U^\perp$. $\square$

Определение 10.43. Линейный оператор $\mathcal{A}$ на конечномерном евклидовом пространстве V называется *самосопряжённым*, если $\mathcal{A} = \mathcal{A}^*$.

Лемма 10.44.

1) *$\mathcal{O}$ (нулевой оператор), $\mathcal{I}$ (тождественный оператор) — самосопряжённые операторы.*

2) *Множество всех самосопряжённых операторов образует линейное пространство над полем $\mathbb{R}$.*

Доказательство.
1) Для всех $x, y \in V$ имеем:

$$\big(\mathcal{O}(x), y\big) = 0 = \big(x, \mathcal{O}(y)\big); \quad \big(\mathcal{I}(x), y\big) = (x, y) = \big(x, \mathcal{I}(y)\big).$$

2) Если $\mathcal{A} = \mathcal{A}^*$, $\mathcal{B} = c\mathcal{B}^*$, $\alpha \in \mathbb{R}$, то:

$$(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^* = \mathcal{A} + \mathcal{B}, \quad (\alpha\mathcal{A})^* = \alpha\mathcal{A}^* = \alpha\mathcal{A}. \qquad \square$$

Предложение 10.45. *Если $\mathcal{A}$, $\mathcal{B}$ — самосопряжённые линейные операторы, то $\mathcal{AB}$ — самосопряжённый линейный оператор в том и только в том случае, когда $\mathcal{AB} = \mathcal{BA}$.*

Доказательство. Так как $(\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{BA}$, то $(\mathcal{AB})^* = \mathcal{AB}$ тогда и только тогда, когда $\mathcal{AB} = \mathcal{BA}$. $\square$

Замечание 10.46. По лемме **??** матрица A самосопряжённого оператора $\mathcal{A}$ в ортонормированном базисе симметрическая: $A^* = A$. Если фиксировать ортонормированный базис $e_1, \dots, e_n$ евклидова пространства V, то любой симметрической матрице $A \in \mathrm{GL}_n(\mathbb{R})$ однозначно соответствует самосопряжённый оператор $\mathcal{A}$, имеющий матрицу A в этом базисе.

Лемма 10.47. *Пусть $\mathcal{A}$ — самосопряжённый линейный оператор на конечномерном евклидовом пространстве V, U — инвариантное подпространство для оператора $\mathcal{A}$, $0 \neq U \subseteq V$, $U \neq V$. Тогда $U^\perp$ также инвариантное подпространство.*

Доказательство. По лемме **??** $U^\perp$ — инвариантное подпространство для линейного оператора $\mathcal{A}^* = \mathcal{A}$. $\square$

Лемма 10.48. *Собственные векторы самосопряжённого оператора евклидова пространства относительно различных собственных значений ортогональны.*

Лемма 10.41. *Пусть $\mathcal{A}$ — линейный оператор на конечномерном евклидовом пространстве V. Тогда линейные операторы $\mathcal{A}$ и $\mathcal{A}^*$ имеют один и тот же характеристический многочлен.*

Доказательство. Характеристический многочлен оператора совпадает с характеристическим многочленом матрицы этого оператора в любом базисе. Пусть $e_1, \dots, e_n$ — ортонормированный базис евклидова пространства V, A — матрица оператора $\mathcal{A}$ в этом базисе. По лемме **??** A^* — матрица оператора $\mathcal{A}^*$ в базисе $e_1, \dots, e_n$. Следовательно,

$$p_{\mathcal{A}}(t) = |A - tE| = |(A - tE)^*| = |A^* - tE| = p_{\mathcal{A}^*}(t). \qquad \square$$

Лемма 10.42. *Пусть U — инвариантное подпространство для линейного оператора $\mathcal{A}$ на евклидовом пространстве V ($\dim V < \infty$, $0 \neq U \neq V$). Тогда $U^\perp$ — инвариантное подпространство для линейного оператора $\mathcal{A}^*$.*

Доказательство. Для всех $x \in U^\perp$, $y \in U$ имеем $(\mathcal{A}^*(x), y) = \big(x, \mathcal{A}(x)\big) = 0$, следовательно, $\mathcal{A}^*(x) \in U^\perp$ для всех $x \in U^\perp$. $\square$

Определение 10.43. Линейный оператор $\mathcal{A}$ на конечномерном евклидовом пространстве V называется *самосопряжённым*, если $\mathcal{A} = \mathcal{A}^*$.

Лемма 10.44.

1) *$\mathcal{O}$ (нулевой оператор), $\mathcal{I}$ (тождественный оператор) — самосопряжённые операторы.*

2) *Множество всех самосопряжённых операторов образует линейное пространство над полем $\mathbb{R}$.*

Доказательство.
1) Для всех $x, y \in V$ имеем:

$$\big(\mathcal{O}(x), y\big) = 0 = \big(x, \mathcal{O}(y)\big); \quad \big(\mathcal{I}(x), y\big) = (x, y) = \big(x, \mathcal{I}(y)\big).$$

2) Если $\mathcal{A} = \mathcal{A}^*$, $\mathcal{B} = c\mathcal{B}^*$, $\alpha \in \mathbb{R}$, то:

$$(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^* = \mathcal{A} + \mathcal{B}, \quad (\alpha\mathcal{A})^* = \alpha\mathcal{A}^* = \alpha\mathcal{A}. \qquad \square$$

Предложение 10.45. *Если $\mathcal{A}$, $\mathcal{B}$ — самосопряжённые линейные операторы, то $\mathcal{AB}$ — самосопряжённый линейный оператор в том и только в том случае, когда $\mathcal{AB} = \mathcal{BA}$.*

Доказательство. Так как $(\mathcal{AB})^* = \mathcal{B}^*\mathcal{A}^* = \mathcal{BA}$, то $(\mathcal{AB})^* = \mathcal{AB}$ тогда и только тогда, когда $\mathcal{AB} = \mathcal{BA}$. $\square$

Замечание 10.46. По лемме **??** матрица A самосопряжённого оператора $\mathcal{A}$ в ортонормированном базисе симметрическая: $A^* = A$. Если фиксировать ортонормированный базис $e_1, \dots, e_n$ евклидова пространства V, то любой симметрической матрице $A \in \mathrm{GL}_n(\mathbb{R})$ однозначно соответствует самосопряжённый оператор $\mathcal{A}$, имеющий матрицу A в этом базисе.

Лемма 10.47. *Пусть $\mathcal{A}$ — самосопряжённый линейный оператор на конечномерном евклидовом пространстве V, U — инвариантное подпространство для оператора $\mathcal{A}$, $0 \neq U \subseteq V$, $U \neq V$. Тогда $U^\perp$ также инвариантное подпространство.*

Доказательство. По лемме **??** $U^\perp$ — инвариантное подпространство для линейного оператора $\mathcal{A}^* = \mathcal{A}$. $\square$

Лемма 10.48. *Собственные векторы самосопряжённого оператора евклидова пространства относительно различных собственных значений ортогональны.*

Доказательство. Пусть $\mathcal{A} = \mathcal{A}^*$, $0 \neq u, v \in V$, $\mathcal{A}(u) = \lambda_1 u$, $\mathcal{A}(v) = \lambda_2 v$, $\lambda_1 \neq \lambda_2$. Тогда

$$\lambda_1(u,v) = (\lambda_1 u, v) = (\mathcal{A}(u), v) = \big(u, \mathcal{A}^*(v)\big) = \big(u, \mathcal{A}(v)\big) = (u, \lambda_2 v) = \lambda_2(u,v);$$

следовательно, $(\lambda_1 - \lambda_2)(u,v) = 0$, $\lambda_1 - \lambda_2 \neq 0$, и поэтому $(u,v) = 0$. □

Теорема 10.49. *Все собственные значения самосопряжённого оператора $\mathcal{A}$ действительны.*

Доказательство. Допустим противное: пусть $\lambda = \alpha + i\beta$ — комплексный корень характеристического многочлена оператора $\mathcal{A}$, $\beta \neq 0$. Тогда (см. **??**) существуют такие линейно независимые элементы $u, v \in V$, что

$$\begin{cases} \mathcal{A}(u) = \alpha u - \beta v, \\ \mathcal{A}(v) = \beta u + \alpha v. \end{cases}$$

Рассматривая скалярное произведение первого равенства на v, а второго — на u, получаем, что

$$\begin{aligned} (\mathcal{A}(u), v) &= \alpha(u,v) - \beta(v,v), \\ \big(u, \mathcal{A}(v)\big) &= \beta(u,u) + \alpha(u,v). \end{aligned}$$

Так как $\mathcal{A}^* = \mathcal{A}$, то $(\mathcal{A}(u), v) = \big(u, \mathcal{A}(v)\big)$, и имеем $0 = \big(u, \mathcal{A}(v)\big) - (\mathcal{A}(u), v) = \beta\big((u,u) + (v,v)\big)$. Но $(u,u) + (v,v) > 0$, следовательно, $\beta = 0$. Полученное противоречие завершает доказательство теоремы. □

Теорема 10.50 (основная теорема о самосопряжённых операторах). *Пусть $\mathcal{A}$ — самосопряжённый оператор на евклидовом пространстве V, $\dim V = n < \infty$. Тогда существует ортонормированный базис пространства V, состоящий из собственных векторов линейного оператора $\mathcal{A}$, в этом базисе матрица A оператора $\mathcal{A}$ диагональна:*

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

где $\lambda_1, \ldots, \lambda_n$ — корни характеристического многочлена оператора $\mathcal{A}$, $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Доказательство. Пусть $p_{\mathcal{A}}(t)$ — характеристический многочлен оператора $\mathcal{A}$. По теореме **??** существует $\lambda_1 \in \mathbb{R}$, $p_{\mathcal{A}}(\lambda_1) = 0$. Пусть $0 \neq v_1$ — собственный вектор оператора $\mathcal{A}$ относительно собственного числа λ_1: $\mathcal{A}(v_1) = \lambda_1 v_1$. Положим $e_1 = \dfrac{1}{|v_1|} v_1$, $U = \langle e_1 \rangle$.

Доказательство проведём индукцией по n. Если $n = 1$, то $U = V$, и мы завершили доказательство теоремы. Если же $n > 1$, то по лемме **??** $U^\perp$ — инвариантное подпространство для оператора $\mathcal{A}$. При этом $1 \leqslant \dim U^\perp < n$. Рассматривая самосопряжённый оператор $\mathcal{A}|_{U^\perp} : U^\perp \to U^\perp$ и применяя предположение индукции, получаем, что существует такой ортонормированный базис $e_2, \ldots, e_n$ подпространства $U^\perp$, что матрица $\tilde{A}$ оператора $\mathcal{A}|_{U^\perp}$ в этом базисе имеет вид

$$\tilde{A} = \begin{pmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \lambda_i \in \mathbb{R}, \quad 2 \leqslant i \leqslant n.$$

Но тогда в ортонормированном базисе $e_1, \ldots, e_n$ евклидова пространства V матрица A оператора $\mathcal{A}$ равна

$$A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \lambda_i \in \mathbb{R},$$

$e_1, \ldots, e_n$ — собственный векторы оператора $\mathcal{A}$, $\lambda_1, \ldots, \lambda_n$ — все корни (считая их кратности) характеристического многочлена $p_{\mathcal{A}}(t)$. □

Замечание 10.51. Для практических задач часто удобнее поступать следующим образом: пусть $\lambda_1, \ldots, \lambda_k$ — все различные корни характеристического многочлена $p_{\mathcal{A}}(t)$, U_i, $1 \leqslant i \leqslant k$, — линейное пространство всех элементов x, удовлетворяющих уравнению $(\mathcal{A} - \lambda_i\mathcal{I})(x) = 0 \in V$ (подпространство U_i состоит из 0 и всех собственных векторов оператора $\mathcal{A}$ относительно собственного числа λ_i). Тогда для построения ортонормированного базиса пространства V, состоящего из собственных векторов оператора $\mathcal{A}$, достаточно последовательно построить ортонормированные базисы подпространств U_i и объединить эти базисы. Построение ортонормированного базиса подпространства U_i сводится к применению процесса ортогонализации к фундаментальной системе решений однородной системы линейных уравнений $(\mathcal{A} - \lambda_i\mathcal{I})(x) = 0$.

Пример 10.52. Пусть ${}_{\mathbb{R}}V = \mathbb{R}^3$; рассмотрим стандартное скалярное произведение; $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$ — стандартный ортонормированный базис; $\mathcal{A}\colon V \to V$ — линейный оператор, имеющий матрицу

$$A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$$

в базисе e_1, e_2, e_3. Так как $A^* = A$ (матрица A симметрическая), то $\mathcal{A}$ — самосопряжённый оператор. Необходимо найти канонический вид оператора $\mathcal{A}$ и ортонормированный базис пространства ${}_{\mathbb{R}}V$, состоящий из собственных векторов оператора $\mathcal{A}$. Вначале находим корни характеристического многочлена $P_{\mathcal{A}}(\lambda) = |A - \lambda E|$, $\lambda_1 = 10$, $\lambda_2 = 1$, $\lambda_3 = 1$. Находим собственные векторы для $\lambda = 10$:

$$(A - 10E)\hat{X} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

в ступенчатом виде:

$$\begin{array}{|ccc|c} 2 & -5 & -4 & 0 \\ 0 & 1 & 1 & 0 \end{array}, \tag{18}$$

фундаментальная система решений состоит из одного элемента $(1, -2, 2)$, нормируем его и полагаем $f_1 = \frac{1}{3}(1, -2, 2)$. Находим собственные векторы для $\lambda_2 = \lambda_3 = 1$:

$$(A - E)\hat{X} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

в ступенчатом виде

$$\begin{array}{|ccc|c} 1 & 2 & -2 & 0 \end{array}, \tag{19}$$

фундаментальная система решений: $g_2 = (2, 0, 1)$, $g_3 = (-2, 1, 0)$. Применяем к g_2, g_3 процесс ортогонализации: $h_2 = g_2$, $h_3 = g_3 + \alpha g_2$, $\alpha = -\frac{(g_3, g_2)}{(g_2, g_2)} = -\frac{-4}{5} = \frac{4}{5}$, $h_3 = \left(-\frac{2}{5}, 1, \frac{4}{5}\right)$. Нормируем элементы h_2 и h_3: $f_2 = \frac{1}{\sqrt{5}}(2, 0, 1)$, $f_3 = \frac{1}{3\sqrt{5}}(-2, 5, 4)$. Таким образом, канонический вид самосопряжённого оператора $\mathcal{A}$: матрица оператора $\mathcal{A}$ в ортонормированном базисе из собственных векторов f_1, f_2, f_3 равна

$$\begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

В рассмотренном нами частном примере легко было заметить, что $V = U_1 \oplus U_1^{\perp}$, где U_1 — подпространство, задаваемое системой (18), поэтому, так как имеется лишь два различных корня характеристического многочлена, то $U_1^\perp = U_2$ (подпространство, отвечающее собственному значению $\lambda_2 = 1$), и можно было, не находя фундаментальную систему решений g_2, g_3 системы (19), сразу применить процесс ортогонализации к строкам системы (18).

Канонический вид матрицы самосопряжённого оператора определён однозначно (с точностью до перестановки вдоль диагонали корней характеристического многочлена). Однако ортонормированный базис из собственных векторов определён неоднозначно. В рассмотренном примере для подпространства $U_2 = U_1^\perp$ существует бесконечно много ортонормированных базисов, состоящих из собственных векторов оператора $\mathcal{A}$ ($\dim U_2 = 2$).

Определение 10.53. Пусть $\mathcal{A}\colon {}_{\mathbb{R}}V \to {}_{\mathbb{R}}V$ — линейный оператор на конечномерном евклидовом пространстве ${}_{\mathbb{R}}V$. Оператор $\mathcal{A}$ называется ортогональным, если

$$\big(\mathcal{A}(x), \mathcal{A}(y)\big) = (x, y) \quad \text{для всех} \quad x, y \in V.$$

Лемма 10.54. *Оператор $\mathcal{A}\colon {}_{\mathbb{R}}V \to {}_{\mathbb{R}}V$ является ортогональным в том и только в том случае, когда $|\mathcal{A}(v)| = |v|$ для всех $v \in {}_{\mathbb{R}}V$ (то есть оператор $\mathcal{A}$ сохраняет длины элементов пространства ${}_{\mathbb{R}}V$).*

Доказательство.

1) Пусть $\mathcal{A}$ — ортогональный оператор. Тогда $\big(\mathcal{A}(x), \mathcal{A}(y)\big)$ для всех $x, y \in V$. Положим $v = x = y$:

$$|\mathcal{A}(v)|^2 = \big(\mathcal{A}(v), \mathcal{A}(v)\big) = (v, v) = |v|^2,$$

поэтому $|\mathcal{A}(v)| = |v|$.

2) Если $|\mathcal{A}(v)| = |v|$ для всех $v \in {}_{\mathbb{R}}V$, то

$$\big(\mathcal{A}(v), \mathcal{A}(v)\big) = |\mathcal{A}(v)|^2 = |v|^2 = (v, v).$$

Положим $v = x + y$:

$$\begin{aligned}\big(\mathcal{A}(x), \mathcal{A}(x)\big) + 2\big(\mathcal{A}(x), \mathcal{A}(y)\big) + \big(\mathcal{A}(y), \mathcal{A}(y)\big) &= \\ = \big(\mathcal{A}(x+y), \mathcal{A}(x+y)\big) = (x+y, x+y) &= (x, x) + 2(x, y) + (y, y),\end{aligned}$$

следовательно, $\big(\mathcal{A}(x), \mathcal{A}(y)\big) = (x, y)$ для всех $x, y \in V$. □

Лемма 10.55.

1) *$\mathcal{I}$ (тождественный оператор) — ортогональный оператор;*

2) *если $\mathcal{A}$ — ортогональный оператор, то он обратим, при этом $\mathcal{A}^{-1} = \mathcal{A}^*$ (это условие равносильно ортогональности), $\mathcal{A}^{-1}$ также ортогональный оператор;*

3) *если $\mathcal{A}$, $\mathcal{B}$ — ортогональные операторы, то $\mathcal{AB}$ также ортогональный оператор.*

Таким образом (учитывая, ассоциативность композиции отображений), множество всех ортогональных операторов на конечномерном евклидовом пространстве образует группу (обозначение: $\mathrm{O}({}_{\mathbb{R}}V)$).

Доказательство.

1) $\big(\mathcal{I}(x), \mathcal{I}(y)\big) = (x, y)$ для всех $x, y \in V$;

2) $(x, y) = \big(\mathcal{A}(x), \mathcal{A}(y)\big) = \Big(x, \mathcal{A}^*\big(\mathcal{A}(y)\big)\Big) = \big(x, (\mathcal{A}^*\mathcal{A})(y)\big)$ для всех $x, y \in V$. По лемме ?? $y = (\mathcal{A}^*\mathcal{A})(y)$ для всех $y \in V$, следовательно, $\mathcal{A}^*\mathcal{A} = \mathcal{I}$, $\mathcal{A}^{-1} = \mathcal{A}^*$. Если оператор $\mathcal{A}$ обратим и $\mathcal{A}^{-1} = \mathcal{A}^*$, то $\mathcal{A}^*\mathcal{A} = \mathcal{I}$ и

$$\big(\mathcal{A}(x), \mathcal{A}(y)\big) = \Big(\mathcal{A}^*\big(\mathcal{A}(y)\big)\Big) = \big(x, (\mathcal{A}^*\mathcal{A})(y)\big) = (x, y)$$

для всех $x, y \in V$, и тогда $\mathcal{A}$ — ортогональный оператор.

Если $\mathcal{A}$ — ортогональный оператор, $\mathcal{A}^{-1} = \mathcal{A}^*$, то для оператора $\mathcal{A}^{-1}$ имеем $(\mathcal{A}^{-1})^{-1} = (\mathcal{A}^*)^{-1} = (\mathcal{A}^{-1})^*$ (см. следствие ??), следовательно, $\mathcal{A}^{-1}$ — ортогональный оператор.

3) Пусть $\mathcal{A}$, $\mathcal{B}$ — ортогональные операторы. Тогда

$$((\mathcal{AB})(x), (\mathcal{AB})(y)) = \Big(\mathcal{A}\big(\mathcal{B}(x)\big), \mathcal{A}\big(\mathcal{B}(y)\big)\Big) = \big(\mathcal{B}(x), \mathcal{B}(y)\big) = (x, y)$$

для всех $x, y \in V$, следовательно, $\mathcal{AB}$ — ортогональный оператор. □

Лемма 10.56. *Если λ — собственное значение ортогонального оператора $\mathcal{A}$ евклидова пространства ${}_{\mathbb{R}}V$, $\lambda \in \mathbb{R}$, то $\lambda = 1$ или $\lambda = -1$.*

Доказательство. Пусть $0 \neq x \in V$ — собственный вектор оператора $\mathcal{A}$ относительно собственного числа λ, $\mathcal{A}(x) = \lambda x$. Тогда $x \neq 0$, $(x, x) > 0$, и, поскольку $\mathcal{A}$ — ортогональный оператор,

$$(x, y) = \big(\mathcal{A}(x), \mathcal{A}(x)\big) = (\lambda x, \lambda x) = \lambda^2 (x, x),$$

и $\lambda^2 = 1$, следовательно, $\lambda = +1$ или $\lambda = -1$. □

Лемма 10.57. *Для линейного оператора $\mathcal{A}$ на конечномерном евклидовом пространстве ${}_{\mathbb{R}}V$ следующие условия равносильны:*

а) *$\mathcal{A}$ — ортогональный оператор;*

б) *$\mathcal{A}$ переводит любой ортонормированный базис в ортонормированный базис.*

Доказательство.

1) Пусть $\mathcal{A}$ — ортогональный оператор, $e_1, \ldots, e_n$ — ортонормированный базис пространства ${}_{\mathbb{R}}V$. Так как $\mathcal{A}$ — обратимый оператор, то по ?? $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_n)$ — базис пространства ${}_{\mathbb{R}}V$. При этом $\big(\mathcal{A}(e_i), \mathcal{A}(e_j)\big) = (e_i, e_j) = \delta_{ij}$ для всех $1 \leqslant i, j \leqslant n$. Следовательно, $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_n)$ — ортонормированный базис пространства ${}_{\mathbb{R}}V$.

2) Пусть теперь $e_1, \ldots, e_n$ — ортонормированный базис и $\mathcal{A}$ — такой линейный оператор, что $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_n)$ — ортонормированный базис пространства ${}_{\mathbb{R}}V$. Если $x \in V$, $x = x_1 e_1 + \ldots + x_n e_n$, $y = y_1 e_1 + \ldots + y_n e_n$, $x_i, y_j \in \mathbb{R}$, то

$$\big(\mathcal{A}(x), \mathcal{A}(y)\big) = \left(\sum_{i=1}^{n} x_i \mathcal{A}(e_i), \sum_{j=1}^{n} y_j \mathcal{A}(e_j)\right) =$$

$$= \sum_{i,j=1}^{n} x_i y_j \big(\mathcal{A}(e_i), \mathcal{A}(e_j)\big) = \sum_{i=1}^{n} x_i y_i = \left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = (x, y)$$

для всех $x, y \in V$. Итак, $\mathcal{A}$ — ортогональный оператор. □

Лемма 10.58. *Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $e_1, \ldots, e_n$ — ортонормированный базис пространства ${}_{\mathbb{R}}V$, $\mathcal{A}$ — ортогональный оператор на пространстве ${}_{\mathbb{R}}V$. Тогда матрица A оператора $\mathcal{A}$ в ортонормированном базисе $e_1, \ldots, e_n$ ортогональна: $AA^* = E = A^*A$, т. е. $A^* = A^{-1}$. Обратно, любой оператор на евклидовом пространстве ${}_{\mathbb{R}}V$, имеющий ортогональную матрицу в ортонормированном базисе, ортогонален.*

Доказательство.

1) Для ортонормированного базиса $e_1, \ldots, e_n$ и ортогонального оператора $\mathcal{A}$ элементы $\mathcal{A}(e_1), \ldots, \mathcal{A}(e_n)$ образуют ортогональный базис пространства V (по лемме ??), следовательно, столбцы матрицы A образуют ортонормированный базис пространства $\hat{\mathbb{R}}^n$ (со стандартным скалярным произведением), поэтому $A^*A = E$, и следовательно, $A^* = A^{-1}$.

2) Если $e_1, \dots, e_n$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$, A — ортогональная матрица, $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{R})$, $AA^* = E = A^*A$, $\mathcal{A}$ — линейный оператор, матрица которого в базисе $e_1, \dots, e_n$ равна A, то

$$(\mathcal{A}(e_i), \mathcal{A}(e_j)) = (\hat{A}_i, \hat{A}_j) = \delta_{ij} = (e_i, e_j),$$

где $\hat{A}_i$, $\hat{A}_j$ — i-й и j-й столбцы матрицы A, $(\hat{A}_i, \hat{A}_j)$ — стандартное покомпонентное скалярное произведение столбцов (так как $A^*A = E$, то $(\hat{A}_i, \hat{A}_j) = \delta_{ij}$). Следовательно, $\mathcal{A}(e_1), \dots, \mathcal{A}(e_n)$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$, и по лемме ?? $\mathcal{A}$ — ортогональный оператор. □

Замечание 10.59. Ортогональные матрицы размера $n \times n$ образуют группу $\mathrm{O}_n(\mathbb{R})$ (ортогональная группа), подгруппу линейной группы $\mathrm{GL}_n(\mathbb{R})$. Определитель любой ортогональной матрицы равен $+1$ или -1: $AA^* = E$, $1 = |AA^*| = |A|\,|A^*| = |A|^2$. Группа $\mathrm{O}_n(\mathbb{R})$ содержит нормальную подгруппу всех ортогональных матриц с определителем $+1$ (специальная ортогональная группа $\mathrm{SO}_n(\mathbb{R})$).

Любая ортогональная матрица $A \in \mathrm{O}_n(\mathbb{R})$ обладает следующими свойствами: столбцы матрицы образуют ортонормированный базис в $\mathbb{R}^n$ (со стандартным скалярным произведением); строки матрицы образуют ортонормированный базис в $\mathbb{R}^n$ (со стандартным скалярным произведением). Действительно, первое свойство следует из равенства $A^*A = E$, а второе — из равенства $AA^* = E$. Каждое из этих свойств эквивалентно ортогональности матрицы A.

Замечание 10.60 (QR-разложение матрицы и QR-алгоритм). Пусть $A \in \mathrm{GL}_n(\mathbb{R})$, $\hat{A}_1, \dots, \hat{A}_n$ — столбцы матрицы A. Применяя процесс ортогонализации к линейно независимой системе столбцов $\hat{A}_1, \dots, \hat{A}_n$ в евклидовом пространстве $\hat{\mathbb{R}}^n$ со стандартным скалярным произведением, мы переходим к ортонормированной системе столбцов $\hat{Q}_1, \dots, \hat{Q}_n$. Пусть Q — матрица со столбцами $\hat{Q}_1, \dots, \hat{Q}_n$. Тогда Q — ортогональная матрица. Каждый шаг процесса ортогонализации для столбцов эквивалентен умножению матрицы, состоящей из рассматриваемых столбцов, справа на верхнетреугольную матрицы. Поэтому $A = QR$, где Q — ортогональная, а R — верхнетреугольная матрицы. Такое разложение матрицы называется *QR-разложением*. При этом можно сделать так, что диагональные элементы матрицы R будут положительны. Тогда матрицы Q и R в QR-разложении обратимой матрицы A, $|A| \neq 0$, определены однозначно. Действительно, пусть $A = Q_1R_1 = Q_2R_2$. Так как матрица A обратима, то $0 \neq |A| = |Q_1|\,|R_1|$, и поэтому матрица R_1 также обратима. Тогда $Q_2^{-1}Q_1 = R_2R_1^{-1}$. Матрица $R_2R_1^{-1}$ является произведением верхнетреугольных матрицы, следовательно, $R_2R_1^{-1}$ и $(R_2R_1^{-1}) = R_1R_2^{-1}$ — верхнетреугольные матрицы. Но $R_2R_1^{-1} = Q_2^{-1}Q_1$ — ортогональная матрица. Следовательно, $(R_2R_1^{-1})^* = (R_2R_1^{-1})^{-1} = R_1R_2^{-1}$ — нижнетреугольная матрица. Поэтому $R_2R_1^{-1}$ — диагональная ортогональная матрица, и все её диагональные элементы равны ± 1. Так как $R_2 = (R_2R_1^{-1})R_1$ и все диагональные элементы матриц R_1 и R_2 положительны, то $R_2R_1^{-1} = E$, $R_2 = R_1$, поэтому $Q_1 = Q_2$.

QR-разложение матриц активно используется в вычислительной линейной алгебре. Если известно QR-разложение матрицы A и решается система линейных уравнений $A\hat{X} = \hat{B} \in \hat{\mathbb{R}}^n$, то эта система эквивалентна системе $R\hat{X} = Q^*\hat{B}$, которая решается очень быстро (матрица R верхнетреугольная).

Один из эффективных алгоритмов приближённого нахождения собственных чисел матрицы — QR-алгоритм, состоящий в следующем. Пусть $A \in \mathrm{GL}_n(\mathbb{R})$, и все собственные числа матрицы A различны по модулю. Положим $A_0 = A$, и пусть $A_0 = Q_0R_0$ — QR-разложение матрицы A_0. Положим $A_1 = R_0Q_0$, рассмотрим QR-разложение матрицы $A_1 = Q_1R_1$ и положим $A_2 = R_1Q_1$, и так далее, $A_k = Q_kR_k$ — QR-разложение матрицы A_k, $A_{k+1} = R_kQ_k, \dots$ Можно показать, что при $k \to \infty$ матрицы A_k сходятся к верхнетреугольной матрице, на диагонали которой стоят собственные числа матрицы A.

Отметим, что существуют более эффективные, чем обычный процесс ортогонализации, методы построения QR-разложений матриц.

Упражнение 10.61. Получите QR-разложения:

$$\begin{pmatrix} 3 & 0 & 1 \\ 0 & 1 & 2 \\ 4 & 0 & 3 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} & 0 & -\frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & \frac{3}{5} \end{pmatrix} \begin{pmatrix} 5 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix};$$

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

Если $A \in \mathrm{M}_{m,n}(\mathbb{R})$, $\mathrm{r}(A) = r$, то существуют такая матрица $Q \in \mathrm{M}_{m,n}(\mathbb{R})$ с ортогональными столбцами и верхнетреугольная матрица $R \in \mathrm{M}_n(\mathbb{R})$, что $A = QR$ (снова достаточно провести процесс ортогонализации столбцов матрицы A). Такое разложение называется QR-разложением прямоугольной матрицы A (с линейно независимыми столбцами).

10.4. Метод наименьших квадратов

Пусть (в результате последовательных измерений) требуется найти величину y как линейную функцию величин $x_1, \ldots, x_n$: $y = b_1x_1 + \ldots + b_nx_n$, $b_i \in \mathbb{R}$. Для этого производятся m измерений: $x_{i1}, \ldots, x_{in}, y_i$ — результаты i-го измерения ($1 \leqslant i \leqslant m$) величин $x_1, \ldots, x_n, y$ соответственно. Обозначим $\hat{X}_i = \begin{pmatrix} x_{i1} \\ \vdots \\ x_{im} \end{pmatrix} \in \mathbb{R}^m$, $1 \leqslant i \leqslant n$, $a_{ij} = x_{ji}$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, $A = (a_{ij}) \in \mathrm{M}_{m,n}(\mathbb{R})$ (столбцы матрицы A: $A_i = \hat{X}_i$, $1 \leqslant i \leqslant n$). Коэффициенты b_i, $1 \leqslant i \leqslant n$, удовлетворяют системе линейных уравнений

$$A\hat{B} = \hat{Y}, \quad \text{где} \quad \hat{B} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad \hat{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}. \tag{20}$$

Обычно число измерений бывает больше числа неизвестных: $m > n$. Принимая во внимание погрешности измерений, мы ищем не точное, а приближённое решение системы (20) (система (20) может быть и, как правило, бывает несовместной). Точнее, ищется такой набор чисел $b_1, \ldots, b_n \in \mathbb{R}$, при котором квадратичное уклонение

$$\rho = \sum_{l=1}^{m} \left(\sum_{i=1}^{n} x_{il} b_i - y_l \right)^2$$

имеет наименьшее значение (если система (20) совместна, то $\rho = 0$, и это значение ρ достигается на любом решении системы (20)).

Задача наименьших квадратов впервые была сформулирована Гауссом при решении практической задачи повышения точности задания координат межевых вех (по заказу германского правительства).

Пусть $\hat{\mathbb{R}}^n$ — евклидово пространство столбцов длины n со стандартным скалярным произведением. Тогда $\rho = |b_1\hat{X}_1 + \ldots + b_n\hat{X}_n - \hat{Y}|$. Пусть $U = \langle \hat{X}_1, \ldots, \hat{X}_n \rangle \subseteq {}_{\mathbb{R}}\hat{\mathbb{R}}^m$. Если $\hat{Y} \notin U$, то по **??** минимальное значение ρ достигается только на ортогональной проекции элемента $\hat{Y}$ на подпространство U. Допустим, что столбцы $\hat{X}_1, \ldots, \hat{X}_n$ линейно независимы. Тогда $(b_1, \ldots, b_n)$ — *единственное* решение системы линейных уравнений

$$\Gamma(\hat{X}_1, \ldots, \hat{X}_n)\hat{B} = \begin{pmatrix} (\hat{Y}, \hat{X}_1) \\ \vdots \\ (\hat{Y}, \hat{X}_n) \end{pmatrix}.$$

Так как в нашем случае

$$\Gamma(\hat{X}_1, \ldots, \hat{X}_n) = A^*A, \qquad \begin{pmatrix} (\hat{Y}, \hat{X}_1) \\ \vdots \\ (\hat{Y}, \hat{X}_n) \end{pmatrix} = A^*\hat{Y},$$

то система (20) может быть записана в виде

$$A^*A\hat{B} = A^*\hat{Y}. \tag{21}$$

Эта система называется нормальной системой для системы (20).

Таким образом, решение системы (20) по методу наименьших квадратов задаётся единственным точным решением системы (21), $\hat{B} = (A^*A)^{-1}A^*\hat{Y}$.

Пример 10.62. Необходимо найти решение (методом наименьших квадратов) системы

$$\begin{cases} x - 2y + 4z = -1, \\ x - y + z = 1, \\ x + y + z = 7, \\ x + 2y + 4z = 3. \end{cases}$$

Эта система является несовместной

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \in \mathrm{M}_{4,3}(\mathbb{R}), \quad \mathrm{r}(A) = 3.$$

Находим нормальную систему

$$A^*A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^*\begin{pmatrix} -1 \\ 1 \\ 7 \\ 3 \end{pmatrix}: \qquad \begin{pmatrix} 4 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 10 \\ 14 \\ 16 \end{pmatrix}.$$

Решая эту систему, получаем единственное решение $x = 5$, $y = 1{,}4$, $z = -1$. Наилучшее отклонение

$$\left| A\begin{pmatrix} 5 \\ 1{,}4 \\ -1 \end{pmatrix} - \begin{pmatrix} -1 \\ 1 \\ 7 \\ 3 \end{pmatrix} \right| = \frac{8}{\sqrt{10}}.$$

Пример 10.63. В евклидовом пространстве непрерывных функций на $(0, 2\pi)$ со скалярным произведением

$$(f, g) = \int\limits_0^{2\pi} f(t)g(t)\,dt$$

для функции $h(t)$ требуется найти такую функцию $\omega(t)$, принадлежащую подпространству с ортонормированным базисом

$$e_0 = \frac{1}{\sqrt{2\pi}},\ e_1 = \frac{\cos t}{\sqrt{\pi}},\ e_2 = \frac{\sin t}{\sqrt{\pi}}, \ldots,\ e_{2n-1} = \frac{\cos nt}{\sqrt{\pi}},\ e_{2n} = \frac{\sin nt}{\sqrt{\pi}},$$

что величина

$$|h(t) - \omega(t)| = \sqrt{\int\limits_0^{2\pi} |h(t) - \omega(t)|^2\, dt}$$

минимальна. Тогда в соответствии с **??**

$$\omega(t) = \frac{a_0}{2} + \sum_{k=1}^{n} \big(a_k \cos(kt) + b_k \sin(kt)\big),$$

где

$$a_0 = \frac{1}{\pi}\int\limits_0^{2\pi} f(t)\, dt, \quad a_k = \frac{1}{\pi}\int\limits_0^{2\pi} f(t)\cos(kt)\, dt, \quad b_k = \frac{1}{\pi}\int\limits_0^{2\pi} f(t)\sin(kt)\, dt\ —$$

коэффициенты Фурье функции $h(t)$.

Замечание 10.64. В общем случае решения системы (20) (необязательно с линейно независимыми столбцами матрицы A) — это в точности решения системы (21). Если столбцы матрицы A линейно зависимы, то нормальное уравнение (21) имеет бесконечно много решений.

Упражнение 10.65. Найдите решение по методу наименьших квадратов системы

$$\begin{cases} 3x + 2y - z = 2, \\ x - 4y + 3z = -2, \\ x + 10y - 7z = 1. \end{cases}$$

Для матрицы

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{pmatrix}$$

имеем

$$A^*A = \begin{pmatrix} 11 & 12 & -7 \\ 12 & 120 & -84 \\ -7 & -84 & 59 \end{pmatrix}, \quad A^* \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 22 \\ -15 \end{pmatrix}.$$

Решение системы

$$A^*A = \begin{pmatrix} 5 \\ 22 \\ -15 \end{pmatrix}:$$

$(x, y, z)_t = \frac{1}{7}\left(2 - t, \frac{13}{12} + 5t, t\right)$, $t \in \mathbb{R}$. Все эти решения (x, y, z) дают одно и то же минимальное квадратичное отклонение:

$$\left| A \begin{pmatrix} x \\ y \\ z \end{pmatrix}_t - \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \right| = \frac{5}{\sqrt{6}}.$$

Замечание 10.66. Если столбцы матрицы A линейно независимы, то система линейных уравнений (20) $(A^*A)\hat{B} = A^*\hat{Y}$ имеет единственное решение. Это решение может быть эффективно найдено, например, с использованием LU-разложения положительно определённой симметрической матрицы A^*A. В практических вычислениях для уменьшения накопления ошибок стараются избежать прямого вычисления матрицы A^*A. Для этого можно воспользоваться QR-разложением матрицы A: если $A \in \mathrm{M}_{m,n}(\mathbb{R})$, $\mathrm{r}(A) = r$, $A = QR$ — QR-разложение прямоугольной матрицы A, то система (21) записывается в виде

$$R^*R\hat{B} = R^*Q^*QR\hat{B} = (QR)^*(QR)\hat{B} = (QR)^*\hat{B} = R^*Q^*\hat{B} \tag{22}$$

(ясно, что R^*R — разложение Холеского матрицы A^*A). Так как $R \in \mathrm{GL}_r(\mathbb{R})$, то система (22) эквивалентна системе

$$R\hat{B} = Q^*\hat{Y} \tag{23}$$

с обратимой верхнетреугольной матрицей R. В таком виде система решается эффективно (формула для коэффициентов $b_1, \ldots, b_r$: $\hat{B} = R^{-1}Q^*\hat{Y}$).

Пример 10.67 (использования QR-разложения для решения нормальной системы линейных уравнений в методе наименьших квадратов). Требуется найти решение (методом наименьших квадратов) для системы

$$\begin{cases} x + 3y + 5z = -2, \\ x + y + 0 \cdot z = 3, \\ x + y + 2z = -1, \\ x + 3y + 3z = 2. \end{cases}$$

В этом случае

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{pmatrix} \in \mathrm{M}_{4,3}(\mathbb{R}), \quad \mathrm{r}(A) = 3.$$

Матрица A имеет следующее QR-разложение:

$$Q = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}, \quad R = \begin{pmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}.$$

Поэтому система (23) имеет вид

$$\begin{pmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = Q^*\begin{pmatrix} -2 \\ 3 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -4 \end{pmatrix}.$$

Решая эту систему, получаем $x = \frac{1}{2}$, $y = \frac{5}{2}$, $z = -2$.

Лемма 10.68. *Пусть $_{\mathbb{R}}V$ — евклидово пространство, $\mathcal{A}\colon V \to V$ — ортогональный оператор, $0 \neq U \subseteq V$, $U \neq V$, U — инвариантное подпространство для оператора $\mathcal{A}$. Тогда $U^{\perp}$ также инвариантное подпространство для оператора $\mathcal{A}$.*

Доказательство. Так как $\mathcal{A}(U) \subseteq U$, то по лемме **??** $\mathcal{A}^*(U^{\perp}) \subseteq U^{\perp}$. Но $\mathcal{A}$ — ортогональный оператор, поэтому $\mathcal{A}^{-1} = \mathcal{A}^*$ и $\mathcal{A}^{-1}(U^{\perp}) = \mathcal{A}^*(U^{\perp}) \subseteq U^{\perp}$. Следовательно, по лемме **??** $\mathcal{A}(U^{\perp}) = (\mathcal{A}^{-1})^{-1}(U^{\perp}) \subseteq U^{\perp}$. □

Лемма 10.69 (ортогональные операторы в одномерном случае). *Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $\dim_{\mathbb{R}}V = 1$, $\mathcal{A}$ — ортогональный оператор на пространстве V. Тогда либо $\mathcal{A} = \mathcal{I}$ (тождественный оператор), либо $\mathcal{A} = -\mathcal{I}$ (центральная симметрия).*

Доказательство. Пусть v — базис одномерного пространства ${}_{\mathbb{R}}V$. Тогда $\mathcal{A}(v) = \lambda v$, $\lambda \in \mathbb{R}$, λ — собственное число ортогонального оператора $\mathcal{A}$. Следовательно, или $\lambda = 1$, и тогда для любого $x \in V$, $x = \alpha v$, $\alpha \in \mathbb{R}$, имеем

$$\mathcal{A}(x) = \mathcal{A}(\alpha v) = \alpha\mathcal{A}(v) = \alpha v = x,$$

и $\mathcal{A} = \mathcal{I}$, или $\lambda = -1$, и тогда для $x = \alpha v$, $\alpha \in \mathbb{R}$, имеем

$$\mathcal{A}(x) = \mathcal{A}(\alpha v) = \alpha\mathcal{A}(v) = \alpha(-v) = -x,$$

и $\mathcal{A} = -\mathcal{I}$. ∎

Лемма 10.70 (о структуре ортогонального оператора на двумерном евклидовом пространстве). *Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $\dim_{\mathbb{R}}V = 2$, $\mathcal{A}\colon V \to V$ — ортогональный оператор. Тогда:*

1) *если оператор $\mathcal{A}$ имеет действительное собственное значение, то $\mathcal{A} = \mathcal{I}$ (тождественный оператор), или $\mathcal{A} = -\mathcal{I}$ (центральная симметрия), или существует ортонормированный базис v_1, v_2 пространства V, в котором матрица A оператора $\mathcal{A}$ равна*

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

(осевая симметрия);

2) *если оператор $\mathcal{A}$ не имеет действительных собственных чисел, то существует ортонормированный базис пространства V, в котором матрица оператора $\mathcal{A}$ равна*

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \quad \varphi \neq \pi k, \quad k \in \mathbb{Z}.$$

Доказательство. Пусть e_1, e_2 — ортонормированный базис,

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

матрица оператора $\mathcal{A}$ в базисе e_1, e_2. Оператор $\mathcal{A}$ ортогональный, поэтому $A^*A = E \in \mathrm{M}_2(\mathbb{R})$, что эквивалентно системе

$$\begin{cases} a_{11}^2 + a_{21}^2 = 1, \\ a_{12}^2 + a_{22}^2 = 1, & a_{ij} \in \mathbb{R}, \quad 1 \leqslant i, j \leqslant n \\ a_{11}a_{12} + a_{21}a_{22} = 0, \end{cases}$$

(эти условия означают, что столбцы матрицы A образуют ортонормированный базис евклидова пространства ${}_{\mathbb{R}}\hat{\mathbb{R}}^2$ со стандартным скалярным произведением). Так как $a_{ij} \in \mathbb{R}$, $a_{11}^2 + a_{21}^2 = 1$, $a_{12}^2 + a_{22}^2 = 1$, то можно положить

$$a_{11} = \cos\varphi, \quad a_{21} = \sin\varphi, \quad a_{12} = \cos\psi, \quad a_{22} = \sin\psi,$$

где $0 \leqslant \varphi, \psi < 2\pi$. Тогда

$$0 = a_{11}a_{12} + a_{21}a_{22} = \cos\varphi\cos\psi + \sin\varphi\sin\psi = \cos(\psi - \varphi) = 0.$$

Следовательно, по модулю 2π, $\psi - \varphi = \dfrac{\pi}{2}$ или $\psi - \varphi = \dfrac{3\pi}{2}$.

1) Если $\psi = \varphi + \dfrac{2\pi}{2}$, то

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

в этом случае матрицы A в ортонормированном базисе не только ортогональная, но и симметрическая, по теореме ?? собственные числа λ_1, λ_2 матрицы A действительны, а так как A — ортогональная матрица, то $\lambda_1, \lambda_2 \in \{1, -1\}$. По теореме ?? существует ортонормированный базис v_1, v_2 евклидова пространства ${}_{\mathbb{R}}V$, в котором матрица оператора $\mathcal{A}$ имеет вид

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

если $\lambda_1 = \lambda_2 = 1$, $\mathcal{A} = \mathcal{I}$ — тождественный оператор;

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

если $\lambda_1 = \lambda_2 = -1$, $\mathcal{A} = \mathcal{I}$ — центральная симметрия; или

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

если $\lambda_1 = 1$, $\lambda_2 = -1$, $\mathcal{A}$ — осевая симметрия,

$$\mathcal{A}(x_1v_1 + x_2v_2) = x_1v_1 - x_2v_2, \quad x_1, x_2 \in \mathbb{R}.$$

2) Если же $\psi = \varphi = \dfrac{3\pi}{2}$, то

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

$\mathcal{A}$ — оператор поворота на угол φ против часовой стрелки на плоскости. Характеристический многочлен оператора $\mathcal{A}$:

$$p_A(t) = (\cos\varphi - t)^2 + \sin^2 t = t^2 - 2(\cos\varphi)t + 1.$$

Если $\varphi = 0$, то многочлен $p_A(t)$ имеет корни $\lambda_1 = \lambda_2 = 1$,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$\mathcal{A} = \mathcal{I}$ — тождественный оператор. Если $\varphi = \pi$, то

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

$\mathcal{A} = -\mathcal{I}$ — центральная симметрия. Если $\varphi \neq \pi k$, $k \in \mathbb{Z}$, то многочлен $p_A(t)$ не имеет действительных корней. □

Замечание 10.71. Ортонормированный базис, существование которого утверждается в теореме ??, может быть определён неоднозначно. Например, тождественный оператор $\mathcal{I}$ имеет единичную матрицу в любом базисе.

Следствие 10.72. *Композиция двух поворотов плоскости относительно общего центра — это поворот относительно этого центра; композиция поворота и осевой симметрии относительно оси, проходящей через центр поворота — осевая симметрия; композиция двух осевых симметрий относительно пересекающихся осей — поворот относительно точки пересечения осей.*

Доказательство. Из леммы ?? следует, что все ортогональные преобразования векторов на плоскости исчерпываются поворотами (включая тождественный оператор и центральную симметрию (поворот на угол π)), определители матриц этих операторов в ортонормированном базисе равны 1, и осевыми симметриями (определители матриц осевых симметрий в ортонормированном базисе равны -1). Определитель матрицы композиции линейных операторов равен произведению определителей. □

Пример 10.73. Пусть e_1, e_2 — ортонормированный базис евклидова пространства $\mathbb{R}^2$, $\mathcal{A}$ — осевая симметрия относительно оси e_2, $\mathcal{B}$ — осевая симметрия относительно оси e_1. Тогда $\mathcal{AB}$ — центральная симметрия (поворот на угол π):

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad AB = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Теорема 10.74 (основная теорема о структуре ортогонального оператора). *Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $\dim V = n < \infty$, $\mathcal{A}\colon V \to V$ — ортогональный оператор. Тогда существует ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$, в котором матрица A оператора $\mathcal{A}$ имеет вид*

$$A = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & 0 & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & \boxed{\begin{matrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{matrix}} & & \\ & & 0 & & & & & \ddots & \\ & & & & & & & & \boxed{\begin{matrix} \cos\varphi_m & -\sin\varphi_m \\ \sin\varphi_m & \cos\varphi_m \end{matrix}} \end{pmatrix}. \qquad \textbf{(24)}$$

матрица A имеет блочный вид, по диагонали стоят k единиц, $0 \leqslant k \leqslant n$, затем l элементов -1, $0 \leqslant l \leqslant n$, затем m блоков

$$\begin{pmatrix} \cos\varphi_i & -\sin\varphi_i \\ \sin\varphi_i & \cos\varphi_i \end{pmatrix}, \quad 0 \leqslant \varphi < 2\pi, \quad \varphi \neq 0, \pi,$$

$0 \leqslant m \leqslant \dfrac{n}{2}$, *при этом $k + l + 2m = n = \dim_{\mathbb{R}} V$.*

Доказательство проведём индукцией по n. Основание индукции: $n = 1, 2$, в этих случаях утверждение теоремы следует из лемм 10.69 и 10.70. Пусть $n > 2$ и утверждение теоремы выполнено для всех $n' < n$. Пусть характеристический многочлен $p_{\mathcal{A}}(t)$ имеет действительный корень λ. Тогда по лемме ?? $\lambda = 1$ или $\lambda = -1$. Пусть $0 \neq v$ — собственный вектор оператора $\mathcal{A}$ относительно собственного числа λ, $\mathcal{A}(v) = \lambda v$, $e_1 = \dfrac{1}{|v|}v$, $V_1 = \langle e_1 \rangle$. Тогда V_1 — инвариантное подпространство для оператора $\mathcal{A}$. По лемме ?? $\mathcal{A}(V_1^\perp) \subseteq V_1^\perp$, при этом $\dim V_1^\perp = n - 1$, $\mathcal{A}|_{V_1^\perp}$ — ортогональный оператор на евклидовом пространстве $V_1^\perp$. Применяя предположение индукции, получаем ортонормированный базис $e_2, \ldots, e_n$ подпространства $V_1^\perp$, в котором матрица оператора $\mathcal{A}|_{V_1^\perp}$ имеет требуемый вид. Так как $V = V_1 \oplus V_1^\perp$, то $e_1, \ldots, e_n$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$. Изменяя, если необходимо, порядок элементов базиса $e_1, \ldots, e_n$, получаем ортонормированный базис, в котором матрица A оператора $\mathcal{A}$ имеет требуемый вид (24).

Пусть теперь характеристический многочлен $p_{\mathcal{A}}(t)$ не имеет действительных корней. Тогда по лемме ?? существует инвариантное подпространство $U \subseteq V$, $\dim U = 2$. По теореме ?? существует такой ортонормированный базис e_1, e_2 пространства U, что матрица оператора $\mathcal{A}|_U$ в этом базисе равна

$$\begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix}, \quad 0 < \varphi < 2\pi, \quad \varphi_1 \neq \pi.$$

По лемме ?? $\mathcal{A}(U^{\perp}) \subseteq U^{\perp}$. Так как $\dim(U^{\perp}) = n - 2$, то, применяя предположение индукции, получаем, что существует такой ортонормированный базис $e_3, \ldots, e_n$, что матрица ортогонального оператора $\mathcal{A}|_{U^{\perp}}$ в этом базисе имеет требуемый вид. Так как $V = U \oplus U^{\perp}$, то $e_1, \ldots, e_n$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$, в котором матрица A оператора $\mathcal{A}$ имеет требуемый вид 24. □

Следствие 10.75. *Теорема* 10.74 *позволяет дать геометрическую классификацию ортогональных преобразований трёхмерного евклидова пространства. Например,*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} -$$

симметрия относительно прямой,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} -$$

симметрия относительно плоскости,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} -$$

поворот на угол φ вокруг прямой.

В частности, справедлива следующая

Теорема Эйлера. *В трёхмерном евклидовом пространстве любой ортогональный линейный оператор $\mathcal{A}$, не меняющий ориентацию (т. е. определитель матрицы этого оператора в любом базисе равен $+1$, $\mathcal{A} \in \mathrm{SO}(3)$), является вращением относительно некоторой оси.*

Пример 10.76. Пусть ${}_{\mathbb{R}}V = \mathbb{R}^3$ — евклидово пространство строк со стандартным скалярным произведением, $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ — стандартный ортонормированный базис, линейный оператор $\mathcal{A}$ в ортонормированном базисе e_1, e_2, e_3 имеет матрицу

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

Тогда A — ортогональная матрица, $\mathcal{A}$ — ортогональный оператор. Требуется найти ортонормированный базис, в котором матрица оператора $\mathcal{A}$ имеет канонический вид.

Вначале находим корни характеристического многочлена $p_A(t) = |A - tE|$: $\lambda_1 = 1$, $\lambda_{2,3} = \frac{1}{2}(1 \pm \sqrt{3}i)$. Поэтому канонический вид равен

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}, \quad \varphi = \frac{\pi}{3} \quad \text{или} \quad \varphi = \frac{5\pi}{3}.$$

Для $\lambda_1 = 1$ находим собственные векторы оператора $\mathcal{A}$:

$$(A-E)\hat{X} = \begin{pmatrix}0\\0\\0\end{pmatrix}. \quad \frac{1}{3}\begin{pmatrix}-1 & -1 & 2\\ 2 & -1 & -1\\ -1 & 2 & -1\end{pmatrix}\hat{X} = \begin{pmatrix}0\\0\\0\end{pmatrix},$$

главный ступенчатый вид

$$\begin{array}{|ccc} 1 & 0 & -1\\ \hline 0 & \multicolumn{1}{|c}{1} & -1 \\ \hline \end{array}.$$

Фундаментальная система решений состоит из одного элемента: $x = (1,1,1)$. Нормируем элемент x: $e_1 = \dfrac{1}{\sqrt{3}}(1,1,1)$.

Для $\lambda_2 = \dfrac{1}{2}\left(\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i\right)$ находим собственные векторы для комплексификации оператора $\mathcal{A}$:

$$(A-\lambda_2 E)\hat{X} = \begin{pmatrix} \frac{1}{6}-\frac{\sqrt{3}}{2}i & -\frac{1}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{6}-\frac{\sqrt{3}}{2}i & -\frac{1}{3}\\ -\frac{1}{3} & \frac{2}{3} & \frac{1}{6}-\frac{\sqrt{3}}{2}i \end{pmatrix}\hat{X} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Главный ступенчатый вид:

$$\begin{array}{|ccc} 1 & 0 & \frac{1}{2}+\frac{\sqrt{3}}{2}i\\ \hline 0 & \multicolumn{1}{|c}{1} & \frac{1}{2}-\frac{\sqrt{3}}{2}i \\ \hline \end{array}.$$

Фундаментальная система решений состоит из одного элемента

$$z = \left(-\frac{1}{2}-\frac{\sqrt{3}}{2}i,\ -\frac{1}{2}+\frac{\sqrt{3}}{2}i,\ 1\right) = \left(-\frac{1}{2}, -\frac{1}{2}, 1\right) + i\left(-\frac{\sqrt{3}}{2}i, \frac{\sqrt{3}}{2}i, 1\right).$$

Нормируем элементы $u = \left(-\dfrac{1}{2}, -\dfrac{1}{2}, 1\right)$ и $v = i\left(-\dfrac{\sqrt{3}}{2}, \dfrac{\sqrt{3}}{2}, 0\right)$: $e_2 = \dfrac{1}{\sqrt{6}}(-1,-1,2)$, $e_3 = \dfrac{1}{\sqrt{2}}(-1,1,0)$ (нет необходимости применять процесс ортогонализации к найденным элементам u и v, они ортогональны).

Итак, в ортонормированном базисе e_1, e_2, e_3 евклидова пространства $\mathbb{R}^3$ матрица A оператора $\mathcal{A}$ имеет вид

$$A = \begin{pmatrix} 1 & 0 & 0\\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2}\\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\varphi & -\sin\varphi\\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}, \quad \varphi = \frac{5\pi}{3}.$$

Теорема 10.77 (полярное разложение). *Пусть V — евклидово пространство, $\dim V < \infty$, $\mathcal{A}\colon V \to V$ — обратимый линейный оператор. Тогда $\mathcal{A} = \mathcal{B}\mathcal{C}$, где $\mathcal{B}$ — ортогональный оператор, $\mathcal{C}$ — самосопряжённый оператор с положительными собственными значениями (операторы $\mathcal{B}$ и $\mathcal{C}$ определены однозначно). Аналогично, $\mathcal{A} = \mathcal{C}'\mathcal{B}'$, где $\mathcal{B}'$ — ортогональный оператор, $\mathcal{C}'$ — самосопряжённый оператор с положительными собственными значениями.*

Доказательство. Рассмотрим линейный оператор $\mathcal{D} = \mathcal{A}^*\mathcal{A}$. Отметим, что $\mathcal{D}$ — самосопряжённый оператор: $\mathcal{D}^* = (\mathcal{A}^*\mathcal{A})^* = \mathcal{A}^*\mathcal{A}^{**} = \mathcal{A}^*\mathcal{A} = \mathcal{D}$. По теореме ?? все собственные числа оператора $\mathcal{D}$ вещественны. Пусть $0 \neq v \in V$, $\mathcal{A}(v) = \lambda v$, $\lambda \in \mathbb{R}$. тогда $(v, v) > 0$,

$$\lambda(v,v) = (\lambda v, v) = (\mathcal{D}(v), v) = \big((\mathcal{A}^*\mathcal{A})(v), v\big) = \big(\mathcal{A}^*(\mathcal{A}(v)), v\big) = \big(\mathcal{A}(v), \mathcal{A}(v)\big) > 0$$

(так как оператор $\mathcal{A}$ обратим, то $\operatorname{Ker}\mathcal{A} = \{0\}$ и $\mathcal{A}(v) \neq 0$ для $v \neq 0$), следовательно. $\lambda = \dfrac{\big(\mathcal{A}(v), \mathcal{A}(v)\big)}{(v,v)} > 0$.

Пусть по теореме ?? $e_1, \ldots, e_n$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$, в котором матрица оператора $\mathcal{D}$ имеет вид

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

$\lambda_i \in \mathbb{R}$, $\lambda_i > 0$, $1 \leqslant i \leqslant n$, — корни характеристического многочлена оператора $\mathcal{D}$. Рассмотрим самосопряжённый оператор $\mathcal{C}$, имеющий матрицу

$$C = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}$$

в базисе $e_1, \ldots, e_n$. Тогда $\mathcal{C}^2 = \mathcal{D}$. Положим $\mathcal{B} = \mathcal{A}\mathcal{C}^{-1}$ и покажем, что $\mathcal{B}$ — ортогональный оператор. Действительно,

$$\mathcal{B}^*\mathcal{B} = (\mathcal{A}\mathcal{C}^{-1})^*\mathcal{A}\mathcal{C}^{-1} = (\mathcal{C}^{-1})^*\mathcal{A}^*\mathcal{A}\mathcal{C}^{-1} = (\mathcal{C}^*)^{-1}\mathcal{C}^2\mathcal{C}^{-1} = \mathcal{C}^{-1}\mathcal{C} = \mathcal{I}.$$

Итак, $\mathcal{A} = \mathcal{B}\mathcal{C}$ — искомое разложение оператора $\mathcal{A}$.

Пусть $\mathcal{F}$ — такой самосопряжённый линейный оператор с положительными собственными значениями и $\mathcal{H}$ — такой ортогональный оператор, что $\mathcal{A} = \mathcal{H}\mathcal{F}$. Тогда $\mathcal{A}^*\mathcal{A} = \mathcal{F}^*\mathcal{H}^*\mathcal{H}\mathcal{F}$, и так как $\mathcal{H}$ — ортогональный оператор, то $\mathcal{H}^*\mathcal{H} = \mathcal{I}$ (тождественный оператор), $\mathcal{D} = \mathcal{A}^*\mathcal{A} = \mathcal{F}^*\mathcal{F} = \mathcal{F}^2$ ($\mathcal{F}$ — самосопряжённый оператор). Покажем, что $\mathcal{F} = \mathcal{C}$ (и тогда $\mathcal{H} = \mathcal{A}\mathcal{C}^{-1} = \mathcal{B}$). По теореме ?? существует ортонормированный базис $v_1, \ldots, v_n$, в котором матрица F оператора $\mathcal{F}$ равна

$$F = \begin{pmatrix} \alpha_1 & & 0 \\ & \ddots & \\ 0 & & \alpha_n \end{pmatrix}$$

$\alpha_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$. Так как все собственные значения оператора $\mathcal{F}$ положительны, то $\alpha_i > 0$, $1 \leqslant i \leqslant n$. При этом $\mathcal{F}^2 = \mathcal{D}$, поэтому $\alpha_1^2, \ldots, \alpha_n^2$ — все собственные значения оператора $\mathcal{D}$, v_i, $1 \leqslant i \leqslant n$, — собственный вектор оператора $\mathcal{D}$ относительно собственного значения α_i^2. Пусть для i, $1 \leqslant i \leqslant n$, $e_i = \sum\limits_{j=1}^{n} \gamma_{ij} v_j$. Так как по ?? собственные векторы, соответствующие различным собственным значениям (для оператора $\mathcal{D}$), линейно независимы, то $e_i = \sum\limits_{1 \leqslant j \leqslant n} \gamma_{ij} v_j$, $\mathcal{D}(v_j) = (\sqrt{\lambda_i})^2 v_j$. Таким образом,

$$\mathcal{F}(e_j) = \sum_{1 \leqslant j \leqslant n} \gamma_{ij}\mathcal{F}(v_j) = \sqrt{\lambda_i} \sum_{1 \leqslant j \leqslant n} \gamma_{ij} v_j = \sqrt{\lambda_i} e_i, \quad \mathcal{F}(v_j) = \sqrt{\lambda_i} v_j.$$

Таким образом, матрица оператора $\mathcal{F}$ в базисе $e_1, \ldots, e_n$ равна

$$\begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}.$$

Следовательно, $\mathcal{F} = \mathcal{C}$.

Пусть $\mathcal{D}' = \mathcal{A}\mathcal{A}^*$, $\mathcal{C}'$ — такой самосопряжённый линейный оператор с положительными собственными значениями (построенный аналогично построению оператора $\mathcal{C}$), что $(\mathcal{C}')^2 = \mathcal{D}'$. Тогда $\mathcal{B}' = (\mathcal{C}')^{-1}\mathcal{A}$ — ортогональный оператор:

$$\mathcal{B}'(\mathcal{B}')^* = (\mathcal{C}')^{-1}\mathcal{A}\big((\mathcal{C}')^{-1}\mathcal{A}\big)^* = (\mathcal{C}')^{-1}\mathcal{A}\mathcal{A}^*\big((\mathcal{C}')^{-1}\big)^* = (\mathcal{C}')^{-1}\mathcal{D}'\big((\mathcal{C}')^*\big)^{-1} = (\mathcal{C}')^{-1}(\mathcal{C}')^2(\mathcal{C}^{-1}) = \mathcal{I}.$$

Следовательно, $\mathcal{A} = \mathcal{C}'\mathcal{B}'$ — искомое разложение оператора $\mathcal{A}$.

Единственность представления $\mathcal{A} = \mathcal{C}'\mathcal{B}'$ доказывается аналогично единственности представления $\mathcal{A} = \mathcal{B}\mathcal{C}$. $\square$

Замечание 10.78. Полярное разложение имеет важные приложения в теории упругости при вычислении деформаций и напряжений.

Упражнение 10.79. Найдите полярные разложения:

$$\begin{pmatrix} \sqrt{3} & 0 \\ 2 & \sqrt{3} \end{pmatrix} = \begin{pmatrix} \dfrac{\sqrt{3}}{2} & -\dfrac{1}{2} \\ \dfrac{1}{2} & \dfrac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} \dfrac{5}{2} & \dfrac{\sqrt{3}}{2} \\ \dfrac{\sqrt{3}}{2} & \dfrac{3}{2} \end{pmatrix};$$

$$\begin{pmatrix} 2 & 2 & -4 & -4 \\ -1 & 1 & 2 & -2 \\ 4 & 4 & 2 & 2 \\ -2 & 2 & -1 & 1 \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 1 & -2 & -2 \\ -1 & 1 & 2 & -2 \\ 2 & 2 & 1 & 1 \\ -2 & 2 & -1 & 1 \end{pmatrix} \cdot \frac{\sqrt{10}}{2} \begin{pmatrix} 3 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

Замечание 10.80. Если $\mathcal{A}$ — оператор (не обязательно обратимый) на евклидовом пространстве $_{\mathbb{R}}V$, $\dim V < \infty$, то $\mathcal{A} = \mathcal{B}\mathcal{C}$, где $\mathcal{B}$ — ортогональный оператор, $\mathcal{C}$ — самосопряжённый оператор с неотрицательными собственными значениями, $\mathcal{C}^2 = \mathcal{A}^*\mathcal{A}$, оператор $\mathcal{C}$ определён однозначно.

Упражнение 10.81. Пусть $\mathcal{A}$, $\mathcal{F}$ — такие операторы на евклидовом пространстве V, $\dim V < \infty$, что $|\mathcal{A}(x)| = |\mathcal{F}(x)|$ для всех $x \in V$. Тогда существует такой ортогональный оператор $\mathcal{H}$, что $\mathcal{A} = \mathcal{H}\mathcal{F}$.

Указание. Из $|\mathcal{A}(x)| = |\mathcal{F}(x)|$ для всех $x \in V$ следует, что

$$\big(u, (\mathcal{A}^*\mathcal{A})(v)\big) = \big(\mathcal{A}(u), \mathcal{A}(v)\big) = \big(\mathcal{F}(u), \mathcal{F}(v)\big) = \big(u, (\mathcal{F}^*\mathcal{F})(v)\big)$$

для всех $u, v \in V$. Следовательно, по лемме **??** $\mathcal{A}^*\mathcal{A} = \mathcal{F}^*\mathcal{F}$. По замечанию **??** $\mathcal{A} = \mathcal{B}\mathcal{C}$, $\mathcal{F} = \mathcal{B}'\mathcal{C}$ ($\mathcal{C}^2 = \mathcal{A}^*\mathcal{A} = \mathcal{F}^*\mathcal{F}$), где $\mathcal{B}$, $\mathcal{B}'$ — ортогональные операторы. Тогда $\mathcal{H} = \mathcal{B}(\mathcal{B}')^{-1}$ — ортогональный оператор, $\mathcal{A} = \mathcal{H}\mathcal{F}$.

Напомним следующее утверждение для матриц.

Предложение 10.82. *Пусть 2 — обратимый элемент поля K, $n \in \mathbb{N}$, $A \in \mathrm{M}_n(K)$. Тогда $A = B + C$, где $B^* = B$ (симметрическая матрица), $C^* = -C$ (кососимметрическая матрица). Матрицы B и C определены однозначно.*

Доказательство. Пусть $B = \frac{1}{2}(A + A^*)$, $C = \frac{1}{2}(A - A^*)$. Тогда $A = B + C$, $B^* = B$, $C^* = -C$. Пусть $A = B + C$, $B^* = B$, $C^* = -C$. Тогда $A^* = B - C$, и $B = \frac{1}{2}(A + A^*)$, $C = \frac{1}{2}(A - A^*)$. $\square$

Следствие 10.83. *Пусть $\mathcal{A}$ — оператор на евклидовом пространстве $_{\mathbb{R}}V$, $\dim V < \infty$. Тогда $\mathcal{A} = \mathcal{B} + \mathcal{C}$, где $\mathcal{B}$ — самосопряжённый оператор ($\mathcal{B}^* = \mathcal{B}$), $\mathcal{C}^* = -\mathcal{C}$. Операторы $\mathcal{B}$ и $\mathcal{C}$ определены однозначно.*

Доказательство. Утверждение следует из предложения 10.82 для матриц A, B и C операторов $\mathcal{A}$, $\mathcal{B}$ и $\mathcal{C}$ в ортонормированном базисе. □

Замечание 10.84. Теорема и замечание дают другое, отличное от ??, доказательство того факта, что все собственные значения самосопряжённого оператора действительны.

§ 11. Билинейные формы

Определение 11.1. Пусть ${}_KV$ — линейное пространство над полем K. отображение $\mathcal{A}\colon V \times V \to K$ называется билинейной формой на линейном пространстве ${}_KV$, если для всех $x, y, z \in V$, $\alpha \in K$

$$\begin{aligned} \mathcal{A}(x+y, z) &= \mathcal{A}(x, z) + \mathcal{A}(y, z), \\ \mathcal{A}(z, x+y) &= \mathcal{A}(z, x) + \mathcal{A}(z, y), \\ \mathcal{A}(\alpha x, y) &= \alpha\mathcal{A}(x, y) = \mathcal{A}(x, \alpha y), \end{aligned}$$

то есть форма $\mathcal{A}$ линейна по каждому аргументу.

Если $e_1, \ldots, e_n$ — базис линейного пространства ${}_KV$, $\mathcal{A}$ — билинейная форма на ${}_KV$, $x, y \in V$, $x = \sum\limits_{i=1}^{n} x_i e_i$, $y = \sum\limits_{j=1}^{n} y_j e_j$, $x_i, y_j \in K$, $1 \leqslant i, j \leqslant n$, то

$$\mathcal{A}(x, y) = \sum_{i,j=1}^{n} a_{ij} x_i y_j,$$

где $a_{ij} = \mathcal{A}(e_i, e_j) \in K$, $\mathcal{A}(x, y) = XA\hat{Y}$, $A = (a_{ij}) \in \mathrm{M}_n(K)$, $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n)$. Матрица $A = (a_{ij}) \in \mathrm{M}_n(K)$ называется *матрицей билинейной формы* $\mathcal{A}$ в базисе $e_1, \ldots, e_n$.

С другой стороны, если $e_1, \ldots, e_n$ — базис линейного пространства ${}_KV$, $A \in \mathrm{M}_n(K)$, то отображение $\mathcal{A}\colon V \times V \to K$, определённое по правилу $\mathcal{A}(x, y) = XA\hat{Y}$, где $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n)$ — строки координат элементов x и y пространства ${}_KV$ в базисе $e_1, \ldots, e_n$, является билинейной формой на пространстве ${}_KV$, имеющей матрицу A в базисе $e_1, \ldots, e_n$.

Пример 11.2.

1) Пусть $K = \mathbb{R}$, ${}_{\mathbb{R}}V$ — евклидово пространство над $\mathbb{R}$, $\mathcal{A}(x, y) = (x, y)$ — скалярное произведение для всех $x, y \in {}_{\mathbb{R}}V$. Тогда $\mathcal{A}$ — билинейная форма на пространстве ${}_{\mathbb{R}}V$. Если $e_1, \ldots, e_n$ — базис пространства ${}_{\mathbb{R}}V$, то матрица A билинейной формы $\mathcal{A}$ в базисе $e_1, \ldots, e_n$ равна матрице Грама $\Gamma(e_1, \ldots, e_n)$ базиса $e_1, \ldots, e_n$. Если $e_1, \ldots, e_n$ — ортонормированный базис, то A — единичная $(n \times n)$-матрица.

2) Если ${}_KV$ — линейное пространство над полем K, $L_1, L_2\colon {}_KV \to {}_KV$ — линейные отображения (линейные формы), то отображение $\mathcal{A}\colon V \times V \to K$, где $\mathcal{A}(x, y) = L_1(x)L_2(y)$ для всех $x, y \in {}_KV$, является билинейной формой на линейном пространстве ${}_KV$. В частности, если $K = \mathbb{R}$, $V = C[a, b]$, то отображение $\mathcal{A}\colon V \times V \to \mathbb{R}$, заданное для $f, g \in V$ формулой

$$\mathcal{A}(f, g) = \int_a^b\!\!\int_a^b f(x)g(t)\,dx\,dt = \int_a^b f(x)\,dx \int_a^b g(t)\,dt,$$

является билинейной формой на пространстве ${}_{\mathbb{R}}V$.

3) Пусть e_1, e_2 — базис линейного пространства ${}_KV$ над полем K и для $x = x_1e_1 + x_2e_2$, $y = y_1e_1 + y_2e_2 \in {}_KV$

$$\mathcal{A}(x, y) = x_1y_1 + 2x_1y_2 - x_2y_1 + 3x_2y_2.$$

Тогда $\mathcal{A}\colon V\times V\to K$ — билинейная форма на пространстве ${}_KV$, имеющая матрицу $\begin{pmatrix}1 & 2\\ -1 & 3\end{pmatrix}$:

$$\mathcal{A}(x,y)=\begin{pmatrix}x_1 & x_2\end{pmatrix}\begin{pmatrix}1 & 2\\ -1 & 3\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}.$$

4) Если $\mathcal{A}$ — линейный оператор на евклидовом пространстве ${}_{\mathbb{R}}V$, то $f_1,f_2\colon V\times V\to\mathbb{R}$,

$$f_1(x,y)=\big(x,\mathcal{A}(y)\big),\quad f_2(x,y)=\big(\mathcal{A}(x),y\big)\ —$$

билинейные формы на линейном пространстве V.

Пусть

$$\mathcal{E}=\begin{pmatrix}e_1\\ \vdots\\ e_n\end{pmatrix},\quad \mathcal{E}'=\begin{pmatrix}e_1'\\ \vdots\\ e_n'\end{pmatrix}\ —$$

базисы линейного пространства ${}_KV$, $\mathcal{E}'=C^*\mathcal{E}$ ($C\in\mathrm{GL}_n(K)$ — матрица перехода), $\mathcal{A}$ — билинейная форма на линейном пространстве ${}_KV$, A — матрица билинейной формы $\mathcal{A}$ в базисе $\mathcal{E}$. Тогда для $x,y\in{}_KV$, $x=\sum\limits_{i=1}^{n}x_ie_i$, $y=\sum\limits_{j=1}^{n}y_je_j$, $X=(x_1,\ldots,x_n)$, $Y=(y_1,\ldots,y_n)\in K^n$, $\hat{X}=C\hat{X}'$, $\hat{Y}=C\hat{Y}'$, где X' и Y' — строки координат элементов x и y в базисе $\mathcal{E}'$, имеем

$$\mathcal{A}(x,y)=XA\hat{Y}=X'C^*ACY'.$$

Следовательно, матрица A' билинейной формы $\mathcal{A}$ в базисе $e_1',\ldots,e_n'$ равна C^*AC.

Определение 11.3. Билинейная форма $\mathcal{A}$ на пространстве V называется ***симметрической***, если $\mathcal{A}(x,y)=\mathcal{A}(y,x)$ для всех $x,y\in{}_KV$, и *кососимметрической*, если $\mathcal{A}(x.y)=-\mathcal{A}(y.x)$ для всех $x,y\in{}_KV$.

Пример 11.4. Пусть e_1, e_2 — базис линейного пространства ${}_KV$, $x=x_1e_1+x_2e_2$, $y=y_1e_1+{}$ $+y_2e_2\in V$,

$$\mathcal{A}_1(x,y)=x_1y_2+x_2y_1,\quad \mathcal{A}_2(x,y)=x_1y_2-x_2y_1.$$

Тогда $\mathcal{A}_1$ — симметрическая билинейная форма на ${}_KV$, $\mathcal{A}_2$ — кососимметрическая билинейная форма на ${}_KV$.

Определение 11.5. Пусть ${}_KV$ — линейное пространство над полем K, $\mathcal{A}\colon V\times V\to K$ — симметрическая билинейная форма. Отображение $\mathcal{Q}\colon V\to K$, $\mathcal{Q}(x)=\mathcal{A}(x,x)$ для всех $x\in{}_KV$, называется *квадратичной формой* на пространстве V, ассоциированной с билинейной формой $\mathcal{A}$.

Пример 11.6.

1) Пусть ${}_KV=K$, $\mathcal{Q}\colon{}_KV\to K$, $\mathcal{Q}(x)=x^2$ для всех $x\in{}_KV$. Тогда $\mathcal{Q}$ — квадратичная форма на пространстве V.

2) Пусть e_1, e_2 — базис линейного пространства ${}_KV$ над полем K, $\mathcal{Q}(x)=x_1^2-x_1x_2$ для всех $x=x_1e_1+x_2e_2\in V$. Тогда $\mathcal{Q}$ — квадратичная форма на линейном пространстве ${}_KV$.

3) Пусть $K=\mathbb{R}$, $n\in\mathbb{N}$, ${}_KV=\mathbb{R}^n$, f — действительнозначная функция, $f=f(x_1,\ldots,x_n)$, $x_i\in\mathbb{R}$, определённая, непрерывная и имеющая непрерывные производные первого и второго порядка в некоторой окрестности точки $(x_1^0,\ldots,x_n^0)$, $x_i^0\in\mathbb{R}$, $1\leqslant i\leqslant n$. Тогда второй дифференциал $d^2f\colon{}_{\mathbb{R}}V\to{}_{\mathbb{R}}V$ в точке $(x_1^0,\ldots,x_n^0)$,

$$d^2f(dx_1,\ldots,dx_n)=\sum_{i,j=1}^{n}\frac{\partial^2f}{\partial x_i\partial x_j}(x_1^0,\ldots,x_n^0)\,dx_i\,dx_j$$

является квадратичной формой на линейном пространстве ${}_{\mathbb{R}}V$.

4) Пусть ${}_{\mathbb{R}}V = C[0, 1]$,

$$\mathcal{Q}(f) = \int\limits_0^1 \big(f(t)\big)^2\,dt.$$

Покажите, что $\mathcal{Q}$ — квадратичная форма на линейном пространстве ${}_K V$.

Лемма 11.7. *Пусть 2 — обратимый элемент поля K, (т. е. $\operatorname{char} K \neq 2$) $\mathcal{Q}$ —* **квадратичная форма** *на линейном пространстве ${}_K V$ над K. Тогда существует единственная* **симметрическая билинейная** *форма $\mathcal{A}$ на пространстве V, такая что $\mathcal{Q}(x) = \mathcal{A}(x, x)$ для всех $x \in V$.*

Доказательство. Пусть $\mathcal{A}$ — симметрическая билинейная форма, $\mathcal{Q}(v) = \mathcal{A}(v, v)$ для всех $v \in V$. Положим $v = x + y$, $x, y \in V$. Тогда

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

Следовательно,

$$\mathcal{A}(x, y) = \frac{1}{2}\big(\mathcal{Q}(x + y) - \mathcal{Q}(x) - \mathcal{Q}(y)\big)$$

для всех $x, y \in V$. □

Определение 11.8. *Матрицей $Q \in \mathrm{M}_n(K)$ квадратичной формы $\mathcal{Q}$* в базисе $e_1, \ldots, e_n$ линейного пространства V над полем K называется матрица симметрической билинейной формы $\mathcal{A}$ на пространстве V, для которой $\mathcal{Q}(x) = \mathcal{A}(x, x)$ для всех $x \in V$.

Замечание 11.9. Из определения матрицы Q квадратичной формы $\mathcal{Q}$ на пространстве ${}_K V$ в базисе $e_1, \ldots, e_n$ следует, что матрица Q симметрическая: $Q = Q^*$. Если $Q = (q_{ij}) \in \mathrm{M}_n(K)$, $x \in {}_K V$, то

$$\mathcal{Q}(x) = XQ\hat{X} = \sum_{i,j=1}^{n} q_{ij}x_i x_j = \sum_{i=1}^{n} q_{ii}x_i^2 + \sum_{\substack{i,j=1\\ i<j}}^{n} 2q_{ij}x_i x_j,$$

где $X = (x_1, \ldots, x_n) \in {}_K V^n$ — строка координат элемента x в базисе $e_1, \ldots, e_n$.

Пример 11.10. $K = \mathbb{R}$, $V = \mathbb{R}^2$, e_1, e_2 — базис пространства V, $\mathcal{Q}(x) = x_1^2 + 2x_1x_2 + 3x_2^2$ для $x = x_1e_1 + x_2e_2 \in V$. Тогда $Q = \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}$ — матрица квадратичной формы $\mathcal{Q}$ в базисе e_1, e_2.

Пусть $\mathcal{Q}\colon {}_K V \to K$ — квадратичная форма на линейном пространстве ${}_K V$ над полем K, $\dim V = n < \infty$, $\mathcal{E}$, $\mathcal{E}'$ — базисы пространства V, $C \in \mathrm{GL}_n(K)$ — матрица перехода от $\mathcal{E}$ к $\mathcal{E}'$: $\mathcal{E}' = C^*\mathcal{E}$, Q — матрица квадратичной формы $\mathcal{Q}$ в базисе $\mathcal{E}$, Q' — матрица формы $\mathcal{Q}$ в базисе $\mathcal{E}'$. Тогда $Q' = C^*QC$. Так как $C \in \mathrm{GL}_n(K)$, по **??** $\mathrm{r}(Q') = \mathrm{r}(Q)$. Таким образом, ранг матрицы квадратичной формы не зависит от выбора базиса. Это число называется *рангом квадратичной формы $\mathcal{Q}$*.

Теорема 11.11. *Пусть ${}_K V$ — линейное пространство над полем K, 2 — обратимый элемент поля K, $\dim V = n < \infty$, $\mathcal{Q}\colon V \to K$ — квадратичная форма на линейном пространстве ${}_K V$. Тогда существует такой базис $e_1, \ldots, e_n$ линейного пространства ${}_K V$, что матрица $Q = (q_{ij}) \in \mathrm{M}_n(K)$ квадратичной формы $\mathcal{Q}$ в этом базисе диагональна: для $x = x_1e_1 + \ldots + x_ne_n$, $x_i \in K$, $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} q_{ii}x_i^2$, при этом число ненулевых элементов q_{ii}, $1 \leqslant i \leqslant n$, равно $\mathrm{r}(\mathcal{Q})$.*

Доказательство (*метод Лагранжа выделения полных квадратов*). Применим индукцию по числу m переменных x_i, входящих в запись $\mathcal{Q}(x)$ в некотором базисе. Пусть $e_1, \dots, e_n$ — исходный базис.

Если $m = 1$, то $\mathcal{Q}(x) = qx_i^2$, где $1 \leqslant i \leqslant n$, $q \in K$, и утверждение теоремы очевидно.

Допустим, что утверждение теоремы доказано для всех $m' < m$ и

$$\mathcal{Q}(x) = q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2 + \ldots + q_{mm}x_m^2$$

(не ограничивая общности, с точностью до перемены мест базисных элементов $e_1, \dots, e_n$, можно считать, что все коэффициенты при x_ix_j, где i или j равны $m+1, \dots, n$, равны нулю).

Рассмотрим вначале случай, когда $q_{mm} \neq 0$. Соберём все члены, содержащие переменную x_m:

$$\begin{aligned}2q_{1m}x_1x_m + 2q_{2m}x_2x_m + \ldots + 2q_{m-1,m}x_{m-1}x_m + q_{mm}x_m^2 = \\ = \frac{1}{q_{mm}}(q_{1m}x_1 + q_{2m}x_2 + \ldots + q_{mm}x_m)^2 - \frac{1}{q_{mm}}(q_{1m}x_1 + q_{2m}x_2 + \ldots + q_{m-1,m}x_{m-1})^2\end{aligned}$$

(выделение полного квадрата). Обозначим

$$y_1 = x_1, \dots, \; y_{m-1} = x_{m-1}, \; y_m = q_{1m}x_1 + \ldots + q_{mm}x_m;$$

если $m < n$, то

$$y_{m+1} = x_{m+1}, \dots, \; y_n = x_n.$$

Тогда

$$\hat{Y} = D\hat{X}, \quad D = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & \mathbf{0} & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ q_{1m} & q_{2m} & \ldots\ldots & q_{mm} & 0 & \ldots & 0 \\ & & & & 1 & & \\ & \mathbf{0} & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

так как $|D| = q_{mm} \neq 0$, то $D \in \mathrm{GL}_n(K)$, $C = (D^{-1})^*$ — матрица перехода от базиса $\mathcal{E} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$ в базису $\mathcal{E}' = \begin{pmatrix} e_1' \\ \vdots \\ e_n' \end{pmatrix}$, в котором $\mathcal{Q}(y) = \dfrac{1}{q_{mm}}y_m^2 + \mathcal{P}(y)$, где квадратичная форма $\mathcal{P}$ зависит только от $y_1, \dots, y_{m-1}$. По предположению индукции существует базис $e_1'', \dots, e_{m-1}''$ подпространства $\langle e_1', \dots, e_{m-1}' \rangle$, в котором форма $\mathcal{P}(z)$ имеет требуемый вид: $\mathcal{P}(z) = \sum\limits_i^n p_{ii}z_i^2$. Следовательно, в базисе $e_1'', \dots, e_{m-1}'', e_m', e_{m+1}, \dots, e_n$ форма $\mathcal{Q}(x) = \sum\limits_{i=1}^n \tilde{q}_{ii}\tilde{x}_i^2$, где $(\tilde{x}_1, \dots, \tilde{x}_n)$ — строка координат элемента x в этом базисе, $\tilde{q}_{ii}$, $1 \leqslant i \leqslant n$.

Пусть теперь $q_{mm} = 0$. Тогда найдётся такой индекс i, $1 \leqslant i < m$, что $q_{im} \neq 0$ (иначе форма $\mathcal{A}(x, x)$ в исходном базисе $e_1, \dots, e_n$ зависит менее чем от m переменных). Положим $x_i = y_i + y_m$, $x_m = y_i - y_m$, $x_j' = y_j'$ при $j \neq i, m$. Это обратимая замена переменных: $y_i = \dfrac{1}{2}(x_i + x_m)$, $y_m = \dfrac{1}{2}(x_i - x_m)$. Этой замене переменных соответствует замена базиса $\mathcal{E} \to \mathcal{E}'$, в новом базисе $\mathcal{E}'$

коэффициент при y_m^2 равен $-2q_{im} \neq 0$, и мы переходим к рассмотренному случаю с ненулевым коэффициентом при x_m^2.

Ранг матрицы квадратичной формы не зависит от выбора базиса, в диагональной форме ранг равен числу ненулевых диагональных элементов. □

Пример 11.12. Приведение квадратичной формы к каноническому виду с помощью выделения полных квадратов. $V = \mathbb{R}^4$. В исходном базисе e_1, e_2, e_3, e_4

$$\mathcal{Q}(x) = \underline{x_1^2} - \underline{2x_1x_2} + \underline{2x_1x_3} - \underline{2x_1x_4} + x_2^2 + 2x_2x_3 - 4x_2x_4 + x_3^2 - 2x_4 =$$
$$= (x_1 - x_2 + x_3 - x_4)^2 - x_2^2 - x_3^2 - x_4^2 + 2x_2x_3 + 2x_3x_4 - 2x_2x_4.$$

Обратимая замена: $y_1 = x_1 - x_2 + x_3 - x_4$, $y_2 = x_2$, $y_3 = x_3$, $y_4 = x_4$.

$$\mathcal{Q} = y_1^2 + \underline{4y_2y_3 - 6y_2y_4 - 3y_4^2 + 2y_3y_4}.$$

Обратимая замена: $w_1 = y_1$, $y_2 = \dfrac{1}{2}(w_2 - w_3)$, $y_3 = \dfrac{1}{2}(w_2 + w_3)$, $y_4 = w_4$ (создание квадрата, если его нет).

$$\mathcal{Q} = w_1^2 + \underline{w_2^2} - w_3^2 - 3(\underline{w_2} - w_3)w_4 - 3w_4^2 + (\underline{w_2} + w_3)w_4,$$

$w_2^2 - 2w_2w_4 = (w_2 - w_4)^2 - w_4^2$ (выделение полного квадрата). Обратимая замена: $f_1 = w_1$, $f_2 = w_2 - w_4$, $f_3 = w_3$, $f_4 = w_4$.

$$\mathcal{Q} = f_1^2 + f_2^2 - \underline{f_3^2 + 4f_3f_4 - 4f_4^2} = (x_1')^2 + (x_2')^2 - (x_3')^2 + 0 \cdot (x_4')^2,$$

где $x_1' = f_1$, $x_2' = f_2$, $x_3' = f_3 - 2f_4$, $x_4' = f_4$.

Если необходимо найти замену базиса $\mathcal{E} = C^*\mathcal{E}'$, где $\mathcal{E}'$ — базис, в котором форма имеет канонический вид ($x = x_1'e_1' + x_2'e_2' + x_3'e_3' + x_4'e_4'$), то необходимо выразить $\hat{X}' = C^{-1}\hat{X}$ и затем найти матрицу перехода C.

В рассмотренном примере мы выделяли полные квадраты, начиная с первой, а не с последней переменной. Следствие **??** и теорема 11.11 показывают, что для $K = \mathbb{R}$ всегда можно прийти к одному результату.

Замечание 11.13 (метод Якоби). Числа q_{ii}, $1 \leqslant j \leqslant n$, и базис $e_1, \ldots, e_n$ в теореме 11.11 определены неоднозначно. Более того, приведение квадратичной формы к диагональному виду не обязательно осуществлять с помощью метода выделения полных квадратов. В случае, когда все главные миноры матрицы Q квадратичной формы в исходном базисе ненулевые, можно использовать метод Якоби, состоящий в следующем.

Пусть квадратичная форма $\mathcal{Q}(x)$ имеет матрицу Q в базисе $e_1, \ldots, e_n$ пространства V, в котором все главные миноры ненулевые:

$$\Delta_1 = q_{11} > 0, \ \Delta_2 = \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} \neq 0, \ldots, \ \Delta_n = |Q| \neq 0.$$

Тогда существует такой базис $v_1, \ldots, v_n$ пространства V, что для $y = y_1v_1 + \ldots + y_nv_n$ имеем

$$\mathcal{Q}(y) = \frac{1}{\Delta_1}y_1^2 + \frac{\Delta_1}{\Delta_2}y_2^2 + \ldots + \frac{\Delta_{n-1}}{\Delta_n}y_n^2.$$

В частности, это показывает, что если $K = \mathbb{R}$, то при $\Delta_1 > 0, \ldots, \Delta_n > 0$ квадратичная форма $\mathcal{Q}(x)$ является положительно определённой: $\mathcal{Q}(x) > 0$ для всех $0 \neq x \in V$.

Действительно, рассмотрим симметрическую билинейную форму $\mathcal{A}(x, y)$, для которой $\mathcal{A}(x, x) = \mathcal{Q}(x)$. Форма $\mathcal{A}$ имеет матрицу Q в базисе $e_1, \ldots, e_n$.

Будем искать такую матрицу перехода $C \in \mathrm{M}_n(K)$, что $D = (d_{ij}) = C^* \in \mathrm{GL}_n(K)$ — *верхнетреугольная* матрица ($d_{ij} = 0$ при $i > j$),

$$\mathcal{E}' = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = D \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix},$$

при этом коэффициенты d_{ij} будем искать из условий $\mathcal{A}(v_i, v_j) = 0$ при $i \neq j$ (совокупность этих условий эквивалентна тому, что матрица формы $\mathcal{A}$ в базисе $v_1, \dots, v_n$ диагональна). Так как форма $\mathcal{A}$ симметрическая и матрица D верхнетреугольная, то условие $\mathcal{A}(v_i, v_j) = 0$ при $i \neq j$ равносильно тому, что $\mathcal{A}(v_i, e_j) = 0$ для всех $1 \leqslant i, j \leqslant n$, $j < i$. Добавим условие нормировки $\mathcal{A}(v_i, e_i) = 1$, $1 \leqslant i \leqslant n$. Для фиксированного i получаем систему линейных уравнений

$$\begin{pmatrix} \mathcal{A}(e_k, e_l) \\ {\scriptstyle 1 \leqslant k,l \leqslant i} \end{pmatrix} \begin{pmatrix} d_{i1} \\ \vdots \\ d_{ii} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{25}$$

Для каждого i определитель этой системы равен $\Delta_i \neq 0$, следовательно, для каждого i мы однозначно находим коэффициенты d_{ij} ($d_{ij} = 0$ при $j > i$), и матрица $D \in \mathrm{GL}_n(K)$ верхнетреугольная. В частности, применяя правило Крамера к системе (25), получаем, что $d_{ii} = \dfrac{\Delta_{i-1}}{\Delta_i}$, где $\Delta_0 = 1$. Тогда в базисе $v_1, \dots, v_n$ матрица квадратичной формы $\mathcal{Q}$ будет иметь требуемый вид.

Следствие 11.14. *Пусть ${}_{\mathbb{R}}V$ — линейное пространство над $\mathbb{R}$, $\dim_{\mathbb{R}} V = n$, $\mathcal{Q}\colon {}_{\mathbb{R}}V \to \mathbb{R}$ — квадратичная форма. Тогда существует базис $v_1, \dots, v_n$ пространства V, в котором*

$$\mathcal{Q}(x) = \sum_{i=1}^{n} \lambda_i x_i^2, \quad \lambda_i \in \{0, 1, -1\},$$

для $x = \sum\limits_{i=1}^{n} x_i v_i \in V$; число ненулевых коэффициентов λ_i, $1 \leqslant i \leqslant n$, равно $\mathrm{r}(\mathcal{Q})$.

Доказательство. Пусть по теореме 11.11 $e_1, \dots, e_n$ — базис пространства ${}_{\mathbb{R}}V$, в котором $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} q_{ii} x_i^2$ для $x = x_1 e_1 + \ldots + x_n e_n$, $q_{ii} \in \mathbb{R}$ (при этом число ненулевых коэффициентов q_{ii} равно $\mathrm{r}(\mathcal{Q})$).

Для всех i, таких что $q_{ii} \neq 0$, положим $y_i = \sqrt{|q_{ii}|}\, x_i$, а для остальных i положим $y_i = x_i$. Тем самым задана обратимая замена переменных (этой замене соответствует обратимая замена базиса). В новом базисе $v_1, \dots, v_n$ $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$, $\lambda_i \in \{0, 1, -1\}$, и число ненулевых коэффициентов λ_i равно $\mathrm{r}(\mathcal{Q})$. $\square$

Следствие 11.15. *Пусть ${}_{\mathbb{C}}V$ — линейное пространство над $\mathbb{C}$, $\dim_{\mathbb{C}} V = n$, $\mathcal{Q}\colon {}_{\mathbb{C}}V \to \mathbb{C}$ — квадратичная форма. Тогда существует базис пространства ${}_{\mathbb{C}}V$, в котором $\mathcal{Q}(x) = \sum\limits_{i=1}^{r} x_i^2$, где $r = \mathrm{r}(\mathcal{Q})$.*

Доказательство. Пусть по теореме 11.11 $e_1, \dots, e_n$ — такой базис пространства V, в котором $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} q_{ii} x_i^2$ для $x = x_1 e_1 + \ldots + x_n e_n$, $x_i \in \mathbb{C}$, при этом число ненулевых коэффициентов q_{ii} равно $\mathrm{r}(\mathcal{Q})$. Для всех i, таких что $q_{ii} \neq 0$, положим $y_i = \sqrt{q_{ii}}\, x_i$ (выбрав какое-либо из двух значений комплексного квадратного корня), а для остальных i положим $y_i = x_i$. Тем самым задана обратимая замена переменных (этой замене соответствует обратимая замена базиса). В новом базисе $v_1, \ldots, v_n$ $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} \lambda_i y_i^2$, $\lambda_i \in \{0, 1\}$, и число ненулевых коэффициентов λ_i равно $\mathrm{r}(\mathcal{Q})$. Переупорядочив, если это необходимо, базис $v_1, \ldots, v_n$, получаем базис $v_1', \ldots, v_n'$ пространства V, в котором $\mathcal{Q}(x) = \sum\limits_{i=1}^{r} (x_i')^2$, где $r = \mathrm{r}(\mathcal{Q})$, $x = x_1'v_1' + \ldots + x_n'v_n' \in V$. □

Следствие 11.16 (правило Якоби). *Пусть $K = \mathbb{R}$, $\mathcal{Q}(x)$ — квадратичная форма на пространстве ${}_{\mathbb{R}}V$ над $\mathbb{R}$. Если в исходном базисе $e_1, \ldots, e_n$ все главные миноры $\Delta_1, \ldots, \Delta_n$ матрицы A неотрицательны, то число коэффициентов в канонической форме квадратичной формы, равных -1, равно числу перемен знака в последовательности $1, \Delta_1, \ldots, \Delta_n$, а число коэффициентов, равных 1, равно числу совпадений знаков в этой последовательности. В частности, если все главные миноры матрицы Q положительны, то квадратичная форма $\mathcal{Q}$ положительно определённая: $\mathcal{A}(x) > 0$ для всех $x \in V$, $x \neq 0$.*

Доказательство. Утверждение следует из **??** и замечания **??**. □

Теорема 11.17 (закон инерции квадратичных форм). *Пусть $K = \mathbb{R}$, $\mathcal{Q}(x)$ — квадратичная форма на линейном пространстве ${}_{\mathbb{R}}V$ над полем $\mathbb{R}$, $\dim_{\mathbb{R}} V = n < \infty$. Тогда число коэффициентов 0, 1, -1 в каноническом виде квадратичной формы матрицы Q не зависит от выбора базиса.*

Доказательство. Если $\mathcal{Q}(x) = 0$ для всех $x \in V$, то все коэффициенты λ_i в каноническом виде формы $\mathcal{Q}$ (см. следствие **??**) равны нулю. Пусть теперь $\mathrm{r}(\mathcal{Q}) \geqslant 1$, $e_1, \ldots, e_n$ — базис пространства ${}_{\mathbb{R}}V$, в котором для $x = x_1e_1 + \ldots + x_ne_n$, $x_i \in \mathbb{R}$,

$$\mathcal{Q}(x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2, \tag{26}$$

где $0 \leqslant p, q$, $p + q \leqslant n$, $e_1', \ldots, e_n'$ — другой базис пространства ${}_{\mathbb{R}}V$, в котором для $x = x_1'e_1' + \ldots + x_n'e_n'$, $x_i' \in V$,

$$\mathcal{Q}(x) = (x_1')^2 + \ldots + (x_k')^2 - (x_{k+1}')^2 - \ldots - (x_{k+l}')^2, \tag{27}$$

где $0 \leqslant k, l$, $k + l \leqslant n$. По **??** $p + q = k + l = \mathrm{r}(\mathcal{Q}) \geqslant 1$. Допустим, что $p > k$. Рассмотрим подпространства $U = \langle e_1, \ldots, e_p \rangle$ и $W = \langle e_{k+1}', \ldots, e_n' \rangle$ линейного пространства ${}_{\mathbb{R}}V$, $\dim_{\mathbb{R}} U = p$, $\dim_{\mathbb{R}} W = n - k$. Для любого $x \in U$, $x \neq 0$, из (26) следует, что $\mathcal{Q}(x) > 0$, а для любого $x \in W$, $x \neq 0$, $\mathcal{Q}(x) \leqslant 0$. Следовательно, $U \cap W = \{0\}$, $U + W = U \oplus W$, $\dim(U + W) = \dim U + \dim W = p + (n - k) = n + (p - k) > n$. Но $U + W$ — подпространство линейного пространства V, следовательно, $\dim(U + W) \leqslant \dim V = n$. Полученное противоречие завершает доказательство теоремы. □

Определение 11.18. Пусть ${}_{\mathbb{R}}V$ — линейное пространство над $\mathbb{R}$, $\dim_{\mathbb{R}} V = n < \infty$, $\mathcal{Q}(x)$ — квадратичная форма на V. Форма $\mathcal{Q}(x)$ называется положительно определённой, если $\mathcal{Q}(x) > 0$ для всех $x \in V$, $x \neq 0$, и отрицательно определённой, если $\mathcal{Q}(x) < 0$ для всех $x \in V$, $x \neq 0$.

Предложение 11.19. *Квадратичная форма $\mathcal{Q}(x)$ на линейном пространстве ${}_{\mathbb{R}}V$ над $\mathbb{R}$, $\dim_{\mathbb{R}} V = n$, является положительно определённой в том и только в том случае, когда существует базис $v_1, \ldots, v_n$ пространства ${}_{\mathbb{R}}V$, в котором для любого $x \in {}_{\mathbb{R}}V$, $x = x_1v_1 + \ldots + x_nv_n$, $x_i \in \mathbb{R}$, $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} x_i^2$.*

Доказательство. По теореме **??** существует базис $v_1, \ldots, v_n$ пространства ${}_{\mathbb{R}}V$, в котором для $x = x_1v_1 + \ldots + x_nv_n$, $x_i \in \mathbb{R}$, имеем

$$\mathcal{Q}(x) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_{p+q}^2 + 0 \cdot x_{p+q+1} + \ldots + 0 \cdot x_n^2,$$

где $0 \leqslant p, q,\ p+q \leqslant n$. Если $p < n$, то для $x = v_{p-1} + \ldots + v_n$ имеем $\mathcal{Q}(x) \leqslant 0$, что противоречит положительной определённости формы $\mathcal{Q}(x)$. Следовательно, $\mathcal{Q}(x) = x_1^2 + \ldots + x_n^2$.

Если в некотором базисе $v_1, \ldots, v_n$ линейного пространства ${}_{\mathbb{R}}V$ для $x = x_1 v_1 + \ldots + x_n v_n \in {}_{\mathbb{R}}V$ имеем $\mathcal{Q}(x) = x_1^2 + \ldots + x_n^2$, то очевидно, что форма $\mathcal{Q}$ положительно определённая. □

Теорема 11.20 (критерий Сильвестера). *Пусть ${}_{\mathbb{R}}V$ — линейное пространство над $\mathbb{R}$, $\dim_{\mathbb{R}} V = n < \infty$, $e_1, \ldots, e_n$ — базис линейного пространства ${}_{\mathbb{R}}V$, $\mathcal{Q}(x)$ — квадратичная форма на линейном пространстве ${}_{\mathbb{R}}V$, Q — матрица формы $\mathcal{Q}$ в базисе $e_1, \ldots, e_n$. Тогда форма $\mathcal{Q}(x)$ положительно определённая в том и только в том случае, когда все главные миноры матрицы Q положительны:*

$$\Delta_1 = q_{11} > 0,\ \Delta_2 = \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} > 0, \ldots,\ \Delta_n = |A| > 0.$$

Доказательство. Пусть форма $\mathcal{Q}(x)$ положительно определена на линейном пространстве ${}_{\mathbb{R}}V$, $e_1, \ldots, e_n$ — базис линейного пространства ${}_{\mathbb{R}}V$, Q — матрица формы $\mathcal{Q}$ в базисе $e_1, \ldots, e_n$. Тогда $q_{ii} = \mathcal{Q}(e_i) > 0$, $1 \leqslant i \leqslant n$, в частности, $q_{11} > 0$. По теореме 11.11 существует базис $v_1, \ldots, v_n$ пространства V, в котором матрица Q' формы $\mathcal{Q}$ диагональна:

$$Q' = \begin{pmatrix} q'_{11} & & 0 \\ & \ddots & \\ 0 & & q'_{nn} \end{pmatrix}.$$

Так как форма $\mathcal{Q}$ положительно определённая, то $q'_{ii} > 0$ для всех i, $1 \leqslant i \leqslant n$. Следовательно, $|Q'| = \prod\limits_{i=1}^{n} q'_{ii} > 0$. Пусть C — матрица перехода от базиса $v_1, \ldots, v_n$ к базису $e_1, \ldots, e_n$:

$$\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Тогда $\Delta_n = |Q| = |C^* Q' C|$, $|Q| = |C^*|\,|Q'|\,|C| = |C|^2 |Q'| > 0$.

Пусть $V_i = \langle e_1, \ldots, e_i \rangle$, $1 \leqslant i \leqslant n$. Тогда $\mathcal{Q}|_{V_i}$ — положительно определённая квадратичная форма. Тогда матрица формы $\mathcal{Q}|_{V_k}$ в базисе $e_1, \ldots, e_k$ равна $(q_{ij})_{1 \leqslant i,j \leqslant k} \in \mathrm{M}_k(\mathbb{R})$. Как и выше, $\Delta_k = |(q_{ij})_{1 \leqslant i,j \leqslant k}| > 0$. Итак, $\Delta_i > 0$ для всех i, $1 \leqslant i \leqslant n$.

Пусть теперь $e_1, \ldots, e_n$ — базис линейного пространства ${}_{\mathbb{R}}V$, Q — матрица квадратичной формы $\mathcal{Q}$, все угловые миноры матрицы Q положительные:

$$\Delta_1 = q_{11} > 0, \ldots,\ \Delta_n = |Q| > 0.$$

Докажем индукцией по n, что тогда $\mathcal{Q}$ — положительно определённая квадратичная форма.

Основание индукции: $n = 1$, $\Delta_1 = \Delta_n = |Q| = q_{11} > 0$, для $x = x_1 e_1$, $x_1 \in \mathbb{R}$, имеем $\mathcal{Q}(x) = q_{11} x_1^2$, следовательно, форма $\mathcal{Q}$ положительно определённая.

Пусть наше утверждение доказано для всех $k < n$, и пусть $n > 1$. Рассмотрим линейное пространство ${}_{\mathbb{R}}V_{n-1} = \langle e_1, \ldots, e_{n-1} \rangle$ и квадратичную форму $\mathcal{P}(x) = \mathcal{Q}(x)|_{V_{n-1}}$: $\mathcal{P}(x) = \mathcal{Q}(x)$ для $x \in V_{n-1}$. Тогда матрица квадратичной формы $\mathcal{P}(x)$ в базисе $e_1, \ldots, e_{n-1}$ подпространства V_{n-1} равна $(q_{ij})_{1 \leqslant i,j \leqslant n-1} \in \mathrm{M}_{n-1}(\mathbb{R})$, все угловые миноры $\Delta_1, \ldots, \Delta_{n-1}$ положительны. По предположению индукции $\mathcal{P}(x)$ — положительно определённая квадратичная форма. Следовательно, по **??** существует такой базис $v_1, \ldots, v_{n-1}$ пространства ${}_{\mathbb{R}}V_{n-1}$, что для $x = x_1 v_1 + \ldots + x_{n-1} v_{n-1} \in V_{n-1}$ имеем $\mathcal{P}(x) = x_1^2 + \ldots + x_{n-1}^2$. Но тогда в базисе $v_1, \ldots, v_{n-1}, e_n$ пространства V для $x = x_1 v_1 + \ldots + x_{n-1} v_{n-1} + x_n e_n$

$$\mathcal{Q}(x) = x_1^2 + \ldots + x_{n-1}^2 + 2 \sum_{i=1}^{n-1} p_{in} x_i x_n + q_{nn} x_n^2,$$

где $p_{in}, q_{nn} \in \mathbb{R}$, $1 \leqslant i \leqslant n-1$. Выделяем полные квадраты:

$$\mathcal{Q}(x) = (x_1 + p_{1n}x_n)^2 + \ldots + (x_{n-1} + p_{n-1,n}x_n)^2 + px_n^2,$$

где $p = q_{nn} - p_{1n}^2 - \ldots - p_{n-1,n}^2$. Совершаем обратную замену переменных

$$\hat{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = D\hat{X},$$

где

$$D = \begin{pmatrix} 1 & & & p_{1n} \\ & \ddots & 0 & \vdots \\ & & \ddots & p_{n-1,n} \\ 0 & & & 1 \end{pmatrix} \in \mathrm{GL}_n(\mathbb{R}).$$

Этой замене переменных соответствует замена базиса:

$$\mathcal{E}' = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \\ e_n \end{pmatrix}$$

с матрицей перехода $C = (D^{-1})^* \in \mathrm{GL}_n(\mathbb{R})$. При этом в базисе $w_1, \ldots, w_n$ для любого $x \in V$ имеем: $x = y_1w_1 + \ldots + y_nw_n$, $y_i \in \mathbb{R}$; $\mathcal{Q}(x) = \sum\limits_{i=1}^{n-1} y_i^2 + py_n^2$. Определитель матрицы Q' формы $\mathcal{Q}$ в этом базисе равен p. Если F — матрица перехода от базиса $e_1, \ldots, e_n$ к базису $w_1, \ldots, w_n$, $F \in \mathrm{GL}_n(\mathbb{R})$, то

$$p = |Q'| = |F^*QF| = |F|^2|Q| = |F|^2\Delta_n > 0.$$

Следовательно, $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} y_i^2 + py_n^2$ — положительно определённая квадратичная форма. □

Замечание 11.21 (другое доказательство достаточности в критерии Сильвестера). Пусть $e_1, \ldots, e_n$ — базис линейного пространства ${}_{\mathbb{R}}V$, Q — матрица квадратичной формы $\mathcal{Q}$, все угловые миноры матрицы Q положительны:

$$\Delta_1 = q_{11} > 0, \ldots, \Delta_n = |Q| > 0.$$

Тогда из замечания **??** следует, что существует такой базис $v_1, \ldots, v_n$ пространства ${}_{\mathbb{R}}V$, что для $x \in V$, $x = y_1v_1 + \ldots + y_nv_n$, $y_i \in \mathbb{R}$, имеем

$$\mathcal{Q}(x) = \sum_{i=1}^{n} \lambda_i y_i^2,$$

где

$$\lambda_1 = q_{11} = \Delta_1 > 0,\ \lambda_2 = \frac{\Delta_1}{\Delta_2} > 0, \ldots,\ \lambda_n = \frac{\Delta_{n-1}}{\Delta_n} > 0.$$

Таким образом, $\mathcal{Q}$ — положительно определённая квадратичная форма на линейном пространстве ${}_{\mathbb{R}}V$.

Теорема 11.22 (критерий Сильвестера отрицательной определённости квадратичной формы). *Пусть ${}_{\mathbb{R}}V$ — линейное пространство над $\mathbb{R}$, $\dim_{\mathbb{R}} V = n < \infty$, $e_1, \ldots, e_n$ — базис линейного пространства ${}_{\mathbb{R}}V$, $\mathcal{Q}(x)$ — квадратичная форма на линейном пространстве V, Q — матрица формы $\mathcal{Q}$ в базисе $e_1, \ldots, e_n$. Тогда форма $\mathcal{Q}(x)$ отрицательно определённая в том и только в том случае, когда все главные миноры Q ненулевые,*

$$\Delta_1 = q_{11} < 0,\ \Delta_2 > 0,\ \Delta_3 < 0, \ldots,\ \Delta_i \cdot (-1)^i > 0, \ldots,\ \Delta_n \cdot (-1)^n > 0$$

(знаки чисел в последовательности $\Delta_1, \Delta_2, \ldots, \Delta_n$ чередуются).

Доказательство. Форма $\mathcal{Q}(x)$ на пространстве ${}_{\mathbb{R}}V$ отрицательно определена в том и только в том случае, когда квадратичная форма $\mathcal{Q}'(x) = -\mathcal{Q}(x)$ положительно определена на пространстве ${}_{\mathbb{R}}V$. Матрица квадратичной формы Q' в базисе $e_1, \ldots, e_n$ равна $-Q$. Теперь утверждение теоремы следует из теоремы 11.20 и свойств определителя матрицы ??. □

Из определения ?? скалярного произведения следует, что если $\mathcal{A}(x, y) = (x, y)$ для всех элементов x, y евклидова пространства ${}_{\mathbb{R}}V$, то $\mathcal{Q}(x) = \mathcal{A}(x, x)$ — положительно определённая квадратичная форма на пространстве V. Если $e_1, \ldots, e_n$ — ортонормированный базис евклидова пространства V, то матрица Q квадратичной формы $\mathcal{Q}(x)$ в этом базисе единичная.

Следующее предложение показывает, что конечномерное линейное пространство ${}_{\mathbb{R}}V$ над $\mathbb{R}$ с помощью положительно определённой квадратичной формы может быть превращено в евклидово пространство.

Предложение 11.23. *Пусть $\mathcal{Q}(x)$ — положительно определённая квадратичная форма на конечномерном пространстве ${}_{\mathbb{R}}V$ над $\mathbb{R}$, $\mathcal{A}(x, y)$ — соответствующая симметрическая билинейная форма,*

$$\mathcal{A}(x, y) = \frac{1}{2}\big(\mathcal{Q}(x + y) - \mathcal{Q}(x) - \mathcal{Q}(y)\big), \quad \mathcal{Q}(x) = \mathcal{A}(x, x).$$

Тогда билинейная форма $\mathcal{A}(x, y)$ задаёт скалярное произведение на пространстве ${}_{\mathbb{R}}V$, пространство ${}_{\mathbb{R}}V$ со скалярным произведением $(x, y) = \mathcal{A}(x, y)$ является евклидовым пространством. В любом ортонормированном базисе относительно этого скалярного произведения матрица квадратичной формы $\mathcal{Q}$ единичная.

Доказательство. Так как форма $\mathcal{A}(x, y)$ билинейная, то из свойств скалярного произведения остаётся проверить, что $\mathcal{A}(x, x) \geqslant 0$ для всех $x \in V$ и что $\mathcal{A}(x, x) = 0$, только если $x = 0$. Но это следует из положительной определённости квадратичной формы $\mathcal{Q}(x)$ на пространстве ${}_{\mathbb{R}}V$.

Ясно, что в ортонормированном базисе $v_1, \ldots, v_n$ евклидова пространства V относительно введённого скалярного произведения для $x = x_1 v_1 + \ldots + x_n v_n \in V$, $x_i \in \mathbb{R}$, имеем $\mathcal{Q}(x) = \mathcal{A}(x, x) = \sum\limits_{i=1}^{n} x_i^2$, то есть матрица квадратичной формы $\mathcal{Q}(x)$ в базисе $v_1, \ldots, v_n$ единичная. □

Замечание 11.24. Предложение 11.23 показывает, что базис, в котором положительно определённая квадратичная форма принимает канонический вид (её матрица в этом базисе единичная) можно находить с помощью процесса ортогонализации: в ортонормированном базисе матрица скалярного произведения единичная. Метод Якоби (см. замечание 11.13) в этом случае также сводится к процессу ортогонализации.

Рассмотрим теперь квадратичные формы на евклидовых пространствах. Будем совершать переход к новому базису только ортогональными преобразованиями (меняя один ортонормированный базис на другой). Следующая теорема определяет канонический вид квадратичной формы на евклидовом пространстве.

Теорема 11.25 (приведение квадратичной формы к главным осям). *Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $\dim_{\mathbb{R}} V = n < \infty$, $\mathcal{Q}(x)$ — квадратичная форма на пространстве ${}_{\mathbb{R}}V$, $e_1, \ldots, e_n$ — ортонормированный базис пространства ${}_{\mathbb{R}}V$, Q — матрица формы $\mathcal{Q}$ в базисе $e_1, \ldots, e_n$. Тогда существует ортонормированный базис $v_1, \ldots, v_n$ пространства ${}_{\mathbb{R}}V$, в котором для $x = x_1 v_1 + \ldots + x_n v_n$, $x_i \in \mathbb{R}$, $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$, где $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, — все (считая кратность) собственные числа матрицы Q.*

Доказательство. Пусть $\mathcal{Q}\colon {}_{\mathbb{R}}V \to {}_{\mathbb{R}}V$ — линейный оператор, имеющий матрицу Q в базисе $e_1, \ldots, e_n$. Так как $Q^* = Q$ (Q — матрица квадратичной формы) и $e_1, \ldots, e_n$ — ортонормированный базис, то Q — самосопряжённый оператор. По теореме ?? существует ортонормированный базис $v_1, \ldots, v_n$ пространства V, в котором матрица Q' линейного оператора $\mathcal{Q}$ диагональна:

$$Q' = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

где $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, — все корни (считая кратные) характеристического многочлена оператора $\mathcal{Q}$ (матрицы Q). Пусть $C \in \mathrm{GL}_n(\mathbb{R})$ — матрица перехода от базиса $e_1, \ldots, e_n$ к базису $v_1, \ldots, v_n$. Тогда $Q' = C^{-1}QC$. Матрица квадратичной формы $\mathcal{Q}$ в базисе $v_1, \ldots, v_n$ равна C^*QC. Но так как C — матрица перехода от одного ортонормированного базиса к другому, то C — ортогональная матрица, $C^* = C^{-1}$, следовательно, матрица квадратичной формы Q в базисе $v_1, \ldots, v_n$ совпадает с Q' для $x = x_1 v_1 + \ldots + x_n v_n \in V$, $x_i \in \mathbb{R}$, $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$, где $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, — все собственные числа матрицы Q. □

Замечание 11.26. Канонический вид квадратичной формы, полученный в теореме о приведении квадратичной формы к главным осям, единствен с точностью до перестановки чисел λ_i на главной диагонали. Сам набор чисел $\lambda_1, \ldots, \lambda_n$ определён однозначно как набор всех (считая кратность) корней характеристического многочлена матрицы Q.

Пример 11.27 (приведение квадратичной формы к главным осям). Пусть ${}_{\mathbb{R}}V = \mathbb{R}^3$ со стандартным скалярным произведением, $\mathcal{Q}(x)$ — квадратичная форма на пространстве ${}_{\mathbb{R}}V$, для $x = (x_1, x_2, x_3) \in {}_{\mathbb{R}}V$,

$$\mathcal{Q}(x) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3.$$

В стандартном ортонормированном базисе $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ матрица Q формы $\mathcal{Q}$ равна

$$Q = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}.$$

Рассмотрим линейный оператор $\mathcal{Q}$, имеющий матрицу Q в этом базисе. В примере ?? мы построили базис v_1, v_2, v_3, состоящий из собственных векторов оператора $\mathcal{Q}$, матрица оператора $\mathcal{Q}$ в ортонормированном базисе v_1, v_2, v_3 равна

$$\begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Следовательно, для любого элемента $x \in V$, $x = y_1 v_1 + y_2 v_2 + y_3 v_3$, $y_i \in \mathbb{R}$,

$$\mathcal{Q}(x) = 10y_1^2 + y_2^2 + y_3^2.$$

Теорема 11.25 (приведение квадратичной формы к главным осям). *Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $\dim_{\mathbb{R}} V = n < \infty$, $\mathcal{Q}(x)$ — квадратичная форма на пространстве ${}_{\mathbb{R}}V$, $e_1, \ldots, e_n$ — ортонормированный базис пространства ${}_{\mathbb{R}}V$, Q — матрица формы $\mathcal{Q}$ в базисе $e_1, \ldots, e_n$. Тогда существует ортонормированный базис $v_1, \ldots, v_n$ пространства ${}_{\mathbb{R}}V$, в котором для $x = x_1v_1 + \ldots + x_nv_n$, $x_i \in \mathbb{R}$, $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$, где $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, — все (считая кратность) собственные числа матрицы Q.*

Доказательство. Пусть $\mathcal{Q}\colon {}_{\mathbb{R}}V \to {}_{\mathbb{R}}V$ — линейный оператор, имеющий матрицу Q в базисе $e_1, \ldots, e_n$. Так как $Q^* = Q$ (Q — матрица квадратичной формы) и $e_1, \ldots, e_n$ — ортонормированный базис, то Q — самосопряжённый оператор. По теореме ?? существует ортонормированный базис $v_1, \ldots, v_n$ пространства V, в котором матрица Q' линейного оператора $\mathcal{Q}$ диагональна:

$$Q' = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

где $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, — все корни (считая кратные) характеристического многочлена оператора $\mathcal{Q}$ (матрицы Q). Пусть $C \in \mathrm{GL}_n(\mathbb{R})$ — матрица перехода от базиса $e_1, \ldots, e_n$ к базису $v_1, \ldots, v_n$. Тогда $Q' = C^{-1}QC$. Матрица квадратичной формы $\mathcal{Q}$ в базисе $v_1, \ldots, v_n$ равна C^*QC. Но так как C — матрица перехода от одного ортонормированного базиса к другому, то C — ортогональная матрица, $C^* = C^{-1}$, следовательно, матрица квадратичной формы Q в базисе $v_1, \ldots, v_n$ совпадает с Q' для $x = x_1v_1 + \ldots + x_nv_n \in V$, $x_i \in \mathbb{R}$, $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$, где $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, — все собственные числа матрицы Q. □

Замечание 11.26. Канонический вид квадратичной формы, полученный в теореме о приведении квадратичной формы к главным осям, единствен с точностью до перестановки чисел λ_i на главной диагонали. Сам набор чисел $\lambda_1, \ldots, \lambda_n$ определён однозначно как набор всех (считая кратность) корней характеристического многочлена матрицы Q.

Пример 11.27 (приведение квадратичной формы к главным осям). Пусть ${}_{\mathbb{R}}V = \mathbb{R}^3$ со стандартным скалярным произведением, $\mathcal{Q}(x)$ — квадратичная форма на пространстве ${}_{\mathbb{R}}V$, для $x = (x_1, x_2, x_3) \in {}_{\mathbb{R}}V$,

$$\mathcal{Q}(x) = 2x_1^2 + 5x_2^2 + 5x_3^2 + 4x_1x_2 - 4x_1x_3 - 8x_2x_3.$$

В стандартном ортонормированном базисе $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ матрица Q формы $\mathcal{Q}$ равна

$$Q = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}.$$

Рассмотрим линейный оператор $\mathcal{Q}$, имеющий матрицу Q в этом базисе. В примере ?? мы построили базис v_1, v_2, v_3, состоящий из собственных векторов оператора $\mathcal{Q}$, матрица оператора $\mathcal{Q}$ в ортонормированном базисе v_1, v_2, v_3 равна

$$\begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Следовательно, для любого элемента $x \in V$, $x = y_1v_1 + y_2v_2 + y_3v_3$, $y_i \in \mathbb{R}$,

$$\mathcal{Q}(x) = 10y_1^2 + y_2^2 + y_3^2.$$

Предложение 11.28 (одновременное приведение пары квадратичных форм к диагональному виду). *Пусть $\mathcal{Q}(x)$ и $\mathcal{P}(x)$ — квадратичные формы на линейном пространстве ${}_{\mathbb{R}}V$, $\dim_{\mathbb{R}} V = n < \infty$, форма $\mathcal{Q}(x)$ положительно (отрицательно) определённая. Тогда существует базис $v_1, \ldots, v_n$, в котором для $x = x_1 v_1 + \ldots + x_n v_n \in V$, $x_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$: $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} x_i^2$, если форма $\mathcal{Q}(x)$ положительно определена на пространстве ${}_{\mathbb{R}}V$; $\mathcal{Q}(x) = \sum\limits_{i=1}^{n} (-x_i^2)$, если форма $\mathcal{Q}(x)$ отрицательно определена на пространстве ${}_{\mathbb{R}}V$; $\mathcal{P}(x) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$, $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$.*

Доказательство. Пусть $\mathcal{Q}(x)$ — положительно определённая квадратичная форма. Пусть в соответствии с предложением **??** $\mathcal{A}(x, y)$ — такая симметрическая билинейная форма на пространстве V, $\mathcal{A}(x,y) = \dfrac{1}{2}\big(\mathcal{Q}(x+y) - \mathcal{Q}(x) - \mathcal{Q}(y)\big)$; $(x, y) = \mathcal{A}(x, y)$ — скалярное произведение на пространстве ${}_{\mathbb{R}}V$, $w_1, \ldots, w_n$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$ относительно введённого скалярного произведения. Матрица Q в базисе $w_1, \ldots, w_n$ единичная. По теореме 11.25 приведём квадратичную форму $\mathcal{P}(x)$ к главным осям: существует ортонормированный базис $v_1, \ldots, v_n$ евклидова пространства ${}_{\mathbb{R}}V$, в котором $\mathcal{P}(x) = \sum\limits_{i=1}^{n} \lambda_i x_i^2$, $\lambda_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, для $x = x_1 v_1 + \ldots + x_n v_n$, $x_i \in \mathbb{R}$.

Если $\mathcal{Q}(x)$ — отрицательно определённая форма на линейном пространстве ${}_{\mathbb{R}}V$, то $-\mathcal{Q}(x)$ — положительно определённая квадратичная форма, и достаточно применить доказанное утверждение теоремы для формы $-\mathcal{Q}(x)$. □

Замечание 11.29. Для любой квадратичной формы $\mathcal{Q}(x)$ в евклидовом пространстве ${}_{\mathbb{R}}V$ существует единственный самосопряжённый оператор $\mathcal{A}$ на пространстве ${}_{\mathbb{R}}V$, такой что $\mathcal{Q}(x) = (\mathcal{A}(x), x)$ для всех $x \in {}_{\mathbb{R}}V$.

Напомним, что единичная сфера в евклидовом пространстве ${}_{\mathbb{R}}V$ — это множество таких элементов $v \in {}_{\mathbb{R}}V$, что $(v, v) = 1$ (множество всех элементов единичной длины).

Теорема 11.30. *Пусть $\mathcal{A}$ — самосопряжённый линейный оператор на евклидовом пространстве ${}_{\mathbb{R}}V$. Тогда квадратичная форма $\mathcal{Q}(x) = \big(x, \mathcal{A}(x)\big)$ достигает на единичной сфере минимума $C \in \mathbb{R}$ в некоторой точке $c \in V$, $(c, c) = 1$, и максимума $B \in \mathbb{R}$ в некоторой точке $b \in V$, $(b, b) = 1$, при этом C — минимальное, B — максимальное собственные значения линейного оператора $\mathcal{A}$, $\mathcal{A}(c) = C \cdot c$, $\mathcal{A}(b) = B \cdot b$ (b и c — собственные векторы единичной длины оператора $\mathcal{A}$ относительно собственных значений C и B соответственно).*

Доказательство. Функция $\mathcal{Q}(x)$ непрерывна на единичной сфере — замкнутом ограниченном подмножестве n-мерного точечного пространства $V \cong \mathbb{R}^n$, которое является компактом. Следовательно, функция $\mathcal{Q}(x) = \big(x, \mathcal{A}(x)\big)$ достигает минимума C в некоторой точке c, $(c, c) = 1$, и максимума B в некоторой точке b, $(b, b) = 1$. Поэтому $B \geqslant \mathcal{Q}(x) \geqslant C$ для всех $x \in {}_{\mathbb{R}}V$, $(x, x) = 1$. Если $v \in V$, то $v = |v| \cdot x$, где $x \in V$, $(x, x) = 1$, $\mathcal{Q}(x) = |v|^2 \mathcal{Q}(x) = (v, v)\mathcal{Q}(x)$. Из $B \geqslant \mathcal{Q}(x) \geqslant C$ следует, что $B \cdot (v, v) \geqslant \mathcal{Q}(x) \cdot (v, v) \geqslant C \cdot (v, v)$. Следовательно, $B \cdot (v, v) \geqslant \mathcal{Q}(v) \geqslant C \cdot (v, v)$ для всех $v \in V$. Тогда $(v, \mathcal{A}(v) - Cv) \geqslant 0$, $(v, \mathcal{A}(v) - Bv) \leqslant 0$ для всех $v \in V$, $(c, \mathcal{A}(c) - Cc) = 0$, $(b, \mathcal{A}(b) - Bb) = 0$.

Обозначим $\mathcal{F} = \mathcal{A} - C\mathcal{I}$, $\mathcal{H} = \mathcal{A} - B\mathcal{I}$, где $\mathcal{I}$ — тождественный линейный оператор. Тогда $\mathcal{F}$ и $\mathcal{H}$ — самосопряжённые линейные операторы. Полагая $v = c + tz$, $w = b + tz$, $t \in \mathbb{R}$, $z \in {}_{\mathbb{R}}V$, имеем:

$$\big(v, \mathcal{F}(v)\big) = \big(c + tz, \mathcal{F}(c) + t\mathcal{F}(z)\big) = 2t\big(z, \mathcal{F}(c)\big) + t^2\big(z, \mathcal{F}(z)\big) \geqslant 0$$

для всех $t \in \mathbb{R}$, $z \in V$. Следовательно, $\big(z, \mathcal{F}(c)\big) = 0$ для всех $z \in V$, и по **??** $\mathcal{F}(c) = \mathcal{A}(c) - Cc = 0 \in V$, $\mathcal{A}(c) = Cc$. Аналогично,

$$\big(w, \mathcal{H}(w)\big) = \big(b + tz, \mathcal{H}(b) + t\mathcal{H}(z)\big) = 2t\big(z, \mathcal{H}(b)\big) + t^2\big(z, \mathcal{H}(z)\big) \leqslant 0$$

для всех $t \in \mathbb{R}$, $z \in V$. Следовательно, $\big(z, \mathcal{H}(b)\big) = 0$ для всех $z \in V$, и по ?? $\mathcal{H}(b) = \mathcal{A}(b) - Bb = 0 \in V$, поэтому $\mathcal{A}(b) = Bb$. Итак, b и c — собственные векторы линейного оператора $\mathcal{A}$.

Так как $\mathcal{A}$ — самосопряжённый оператор, то по ?? все его собственные значения действительны. Пусть $\dim V = n$, $\lambda_1 \leqslant \ldots \leqslant \lambda_n$ — все собственные значения линейного оператора $\mathcal{A}$, $v_1, \ldots, v_n$ — соответствующие собственные векторы единичной длины. Тогда $\mathcal{Q}(v_i) = \big(v_i, \mathcal{A}(v_i)\big) = \lambda_i$. Следовательно, так как B и C соответственно максимум и минимум функции $\mathcal{Q}(x)$ на единичной сфере, то $C = \lambda_1$ и $B = \lambda_n$. □

Замечание 11.31. Если для квадратичной формы $\mathcal{Q}(x)$ найден собственный вектор c относительно минимального собственного значения $C = \lambda_1$, то для нахождения следующего собственного значения $\lambda_2 \geqslant \lambda_1$ можно рассмотреть разложение $V = \langle c\rangle \oplus \langle c\rangle^\perp$, $\dim(\langle c\rangle^\perp) = n-1$, и минимум квадратичной формы $\mathcal{Q}(x)$, $x \in \langle c\rangle^\perp$, этот минимум будет равен λ_2, и так далее.

Теорема ?? и замечание ?? дают доказательство, отличное от доказательства теоремы ??, того факта, что все собственные числа самосопряжённого оператора вещественны.

Задача 11.32. Пусть ${}_{\mathbb{R}}V$ — евклидово пространство, $\dim V = n$, $\mathcal{A}$ — самосопряжённый оператор на пространстве V, $\lambda_1 \leqslant \ldots \leqslant \lambda_n \in \mathbb{R}$ — собственные значения линейного оператора $\mathcal{A}$. Докажите, что

$$\max_{\substack{U \text{ — подпространство в } V\\ \dim U = n-k+1}} \left\{ \min_{\substack{x \in U\\ (x,x)=1}} \big\{\big(x, \mathcal{A}(x)\big)\big\} \right\} = \lambda_k.$$

Теорема 11.33 (разложение Холеского). *Пусть $A \in \mathrm{M}_n(\mathbb{R})$, $A = A^*$ и A — положительно определённая матрица ($XA\hat{X} > 0$ для любой строки $X \in \mathbb{R}^n$, $X \neq (0, \ldots, 0)$). Тогда существует единственное представление матрицы A в виде $A = C^*C$, где C — верхнетреугольная матрица, в которой все диагональные элементы положительны.*

Доказательство.

1) Рассмотрим матрицу A как матрицу самосопряжённого оператора в стандартном ортонормированном базисе евклидова пространства $\mathbb{R}^n$. По теореме ?? существует такая ортогональная матрица D, что

$$D^{-1}AD = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \lambda_i > 0, \quad 1 \leqslant i \leqslant n.$$

Пусть

$$B = D^{-1} \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} D.$$

Тогда $B^2 = A$,

$$B^* = D^* \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} (D^{-1})^* = B,$$

$A = B^*B$.

Пусть $B = QR$ — QR-разложение матрицы B (см. ??, Q — ортогональная матрица, R — верхнетреугольная матрица с положительными диагональными элементами). Тогда $A = B^*B = R^*Q^*QR = R^*R$. Положив $C = R$, получаем разложение Холеского матрицы A.

2) Докажем теперь индукцией по n единственность разложение Холеского. Основание индукции: $n = 1$, $A = (a_{11})$, $a_{11} > 0$, $a_{11} = (c_{11})^2$, $c_{11} > 0$.

Пусть $n > 1$ и наше утверждение доказано для всех $n' < n$, $A = C^*C$ — разложение Холеского:

$$\left(\begin{array}{ccc|c} a_{11} & \ldots & \ldots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \hline a_{n1} & \ldots & \ldots & a_{nn} \end{array}\right) = \left(\begin{array}{cccc|c} c_{11} & 0 & \ldots & \ldots & 0 \\ c_{12} & c_{22} & \ddots & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ \hline c_{n1} & c_{2n} & \ldots & \ldots & c_{nn} \end{array}\right) \left(\begin{array}{cccc|c} c_{11} & c_{12} & \ldots & \ldots & c_{1n} \\ 0 & c_{22} & \ldots & \ldots & c_{2n} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ \hline 0 & 0 & \ldots & 0 & c_{nn} \end{array}\right),$$

$c_{11} > 0, \ldots, c_{nn} > 0$, $0 < |A| = \prod\limits_{i=1}^{n} c_{ii}^2$. Тогда $c_{nn}^2 = \dfrac{|A|}{c_{11}^2 \cdots c_{n-1,n-1}^2}$, поэтому элемент c_{nn} определён однозначно. Более того, $\bar{A} = \bar{C}^*\bar{C}$, где для матрицы $M \in \mathrm{M}_n(\mathbb{R})$ через $\bar{M}$ мы обозначаем $(n-1)\times(n-1)$-матрицу, стоящую в левом верхнем углу матрицы M. По предположению индукции все элементы матрицы $\bar{C}$ определены однозначно, $|\bar{C}| > 0$. Так как $A = C^*C$, то

$$\bar{C}\begin{pmatrix} c_{1n} \\ \vdots \\ c_{n-1,n} \end{pmatrix} = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{n-1,n} \end{pmatrix},$$

и поскольку $|\bar{C}| \neq 0$, то элементы $c_{1n}, \ldots, c_{n-1,n}$ определены однозначно. Следовательно, все элементы матрицы C определены однозначно. □

Упражнение 11.34. Найдите с помощью указанного алгоритма разложение Холеского следующей положительно определённой матрицы:

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

Замечание 11.35. Разложение Холеского полезно при решении систем линейных уравнений с положительно определённой матрицей: если решается система $A\hat{X} = \hat{B} \in \hat{\mathbb{R}}^n$, $A \in \mathrm{GL}_n(\mathbb{R})$, A — положительно определённая матрица, $A = R^*R$ — разложение Холеского матрицы A, то эта система эквивалентна системе

$$\begin{cases} R^*\hat{Y} = \hat{B}, \\ R\hat{X} = \hat{Y}, \end{cases}$$

где обе системы треугольные с обратимыми матрицами R^* и R, и решение $\hat{X}$ находится эффективно.

В вычислительной линейной алгебре имеется ряд эффективных алгоритмов построения разложения Холеского.

§ 12. Аффинные (или точечно-векторные) пространства

Пусть ${}_KV$ — линейное пространство над полем K, Ω — такое множество (элементы множества Ω будем называть точками), что существует отображение

$$\Omega \times \Omega \to {}_KV, \quad (M, N) \to \overrightarrow{MN} \in V \quad \text{для} \quad M, N \in \Omega,$$

для которого

1) $\forall M \in \Omega\ \forall x \in V\ \exists!\, N,\ \overrightarrow{MN} = x$;

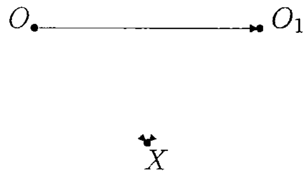
2) $\forall M, N, P \in \Omega \colon \overrightarrow{MN} + \overrightarrow{NP} = \overrightarrow{MP}$.

Пара $(\Omega, {}_K V)$ называется аффинным (точечно-векторным) пространством над полем K. Размерностью этого аффинного пространства называется размерность линейного пространства ${}_K V$.

Пример 12.1. $(\mathbb{R}^n, \mathbb{R}^n)$.

Выберем точку $O \in \Omega$ и базис $v_1, \ldots, v_n$ линейного пространства ${}_K V$. Система координат аффинного пространства $(\Omega, {}_K V)$ — это совокупность $\{O; v_1, \ldots, v_n\}$. Для любой точки $X \in \Omega$ координаты точки X в фиксированной системе координат $\{O; v_1, \ldots, v_n\}$ — это координаты элемента $\overrightarrow{OX}$ в базисе $v_1, \ldots, v_n$ линейного пространства ${}_K V$. Отождествляя точку $X \in \Omega$ с её координатами, $X \in K^n$, имеем $\overrightarrow{XY} = \overrightarrow{OY} - \overrightarrow{OX} = Y - X \in K^n$.

Пусть $\{O_1; v_1', \ldots, v_n'\}$ — другая система координат аффинного пространства (Ω, V). Тогда для пересчёта координат точек в разных системах координат имеем



$$\overrightarrow{OX} - \overrightarrow{O_1X} = \overrightarrow{OO_1},$$

$$\mathcal{E} = \begin{pmatrix} v_1' \\ \vdots \\ v_n' \end{pmatrix} = C^* \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = C^* \mathcal{E},$$

где $C \in \mathrm{GL}_n(K)$ — матрица перехода от базиса $v_1, \ldots, v_n$ линейного пространства ${}_K V$ к базису $v_1', \ldots, v_n'$. Следовательно, $\hat{O}_1 = \hat{X} - C \cdot \hat{X}'$, или $\hat{X} = \hat{O}_1 + C \cdot \hat{X}'$, где $\hat{O}_1$, $\hat{X}$ — столбцы координат точек O_1 и X соответственно в «старой» системе координат $\{O; v_1, \ldots, v_n\}$, $\hat{X}'$ — координаты элемента $\overrightarrow{O_1X}$ в базисе $v_1', \ldots, v_n'$ линейного пространства ${}_K V$.

Пусть зафиксирована система координат $\{O; v_1, \ldots, v_n\}$ точечно-векторного аффинного пространства (Ω, V), $n = \dim V$, $A \in \mathrm{M}_{m,n}(K)$, $\mathrm{r}(A) = r < n$, $k = n - r$. Множество всех точек X пространства (Ω, V), чьи координаты в этой системе координат аффинного пространства (Ω, V) удовлетворяют совместной системе уравнений

$$A \cdot \hat{X} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, \quad b_i \in K, \quad 1 \leqslant i \leqslant m,$$

называется *k-мерной плоскостью* (если $k = n - 1$, то *гиперплоскостью*). Любая гиперплоскость может быть задана линейным уравнением

$$\sum_{i=1}^{n} a_i x_i = b,$$

где $b \in K$, $a_i \in K$, $1 \leqslant i \leqslant n$.

Пусть теперь $K = \mathbb{R}$, ${}_{\mathbb{R}} V$ — евклидово пространство. Тогда аффинное точечно-векторное пространство (Ω, V) называется евклидовым аффинным (точечно-векторным) пространством. *Расстоянием* между точками $M, N \in \Omega$ называется число $|\overrightarrow{MN}| \in \mathbb{R}$.

Пусть k-мерная плоскость W в евклидовом аффинном пространстве задана системой

$$A \cdot \hat{X} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n,$$

где $A \in \mathrm{M}_{m,n}(\mathbb{R})$, $\mathrm{r}(A) = r$, $k = n - r$. Так как $O \in W$, то W — линейное подпространство в ${}_{\mathbb{R}}V$, ${}_{\mathbb{R}}V = W \oplus W^{\perp}$, и для любого $X \in \Omega$ элемент $\overrightarrow{OX}$ однозначно представляется в виде $\overrightarrow{OX} = y + z$, где $y \in W$, $z \in W^{\perp}$. Число $|z|$ называется *расстоянием от точки X до плоскости* W. Ясно, что (см. ??) $|z|$ — минимальное возможное число среди всех чисел $|\overrightarrow{OX} - y|$, где $y \in W$.

Если теперь k-мерная плоскость W задана совместной системой линейных уравнений

$$A \cdot \hat{X} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}, \tag{28}$$

где $A \in \mathrm{M}_{m,n}(\mathbb{R})$, $\mathrm{r}(A) = r$, $k = n - r$, $b_i \in \mathbb{R}$, $1 \leqslant i \leqslant m$, то пусть $\hat{X}_0 \in W$ — частное решение системы (28). Тогда все решения системы (28) имеют вид $\hat{X}_0 + \hat{Y}$, где $\hat{Y}$ — решения однородной системы

$$A \cdot \hat{Y} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^m;$$

все решения $\hat{Y}$ образуют k-мерную гиперплоскость W', проходящую через начало координат. Если $V = W' \oplus (W')^{\perp}$, то для любой точки $X \in \Omega$ элемент $\overrightarrow{X_0X}$ имеет однозначное представление в виде $\overrightarrow{X_0X} = y + z$, где $y \in W'$, $z \in (W')^{\perp}$. Число $|z|$ называется *расстоянием* от точки X до плоскости W.

Пусть $K = \mathbb{R}$, (Ω, V) — n-мерное аффинное пространство, $\mathcal{Q}(x)$ — ненулевая квадратичная форма на линейном пространстве ${}_{\mathbb{R}}V$, $b(x)$ — линейная форма на линейном пространстве ${}_{\mathbb{R}}V$. Пусть $\{O; v_1, \ldots, v_n\}$ — система координат в (Ω, V). *Поверхностью второго порядка* (квадрикой) называется множество таких точек $X \in \Omega$, что

$$\mathcal{Q}(\overrightarrow{OX}) + 2b(\overrightarrow{OX}) + c = 0, \quad \text{где} \quad c \in \mathbb{R}. \tag{29}$$

Отметим, что если $\overrightarrow{OX} = x_1v_1 + \ldots + x_nv_n$, $x_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$, то $b(\overrightarrow{OX}) = b_1x_1 + \ldots + b_nx_n$, $b_i \in \mathbb{R}$, $1 \leqslant i \leqslant n$. С помощью замены базиса линейного пространства ${}_{\mathbb{R}}V$ квадратичная форма $\mathcal{Q}(\overrightarrow{OX})$ может быть приведена в каноническому виду (см. ??), после чего, с помощью выделения полных квадратов (и переноса начала координат), в новой системе координат уравнение (29) преобразуется к одному из следующих типов:

$$\begin{aligned}
&\pm y_1^2 \pm y_2^2 \ldots \pm y_r^2 = 1 && (1 \leqslant r \leqslant n); \\
&\pm y_1^2 \pm y_2^2 \ldots \pm y_r^2 = 0 && (1 \leqslant r \leqslant n); \\
&\pm y_1^2 \pm y_2^2 \ldots \pm y_r^2 - 2y_n = 0 && (1 \leqslant r \leqslant n-1).
\end{aligned}$$

Если (Ω, V) — евклидово n-мерное аффинное пространство, $\{v_1, \ldots, v_n\}$ — ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$, то, изменяя ортонормированный базис евклидова пространства ${}_{\mathbb{R}}V$ на ортонормированный базис $v_1', \ldots, v_n'$ евклидова пространства ${}_{\mathbb{R}}V$ (в процессе приведения квадратичной формы $\mathcal{Q}(\overrightarrow{OX})$ к главным осям и возможного переноса начала координат O в O_1), в новой системе координат $\{O_1; v_1', \ldots, v_n'\}$ уравнение (29) преобразуется к одному из следующих типов:

$$\begin{aligned}
&\lambda_1y_1^2 + \ldots + \lambda_ry_r^2 = H && (r \leqslant n,\ H \in \mathbb{R},\ \lambda_i \in \mathbb{R},\ 1 \leqslant i \leqslant r); \\
&\lambda_1y_1^2 + \ldots + \lambda_ry_r^2 - 2\mu y_n = 0 && (r \leqslant n-1,\ \mu, \lambda_1, \ldots, \lambda_r \in \mathbb{R}).
\end{aligned}$$

При этом $\lambda_1, \ldots, \lambda_r$ — ненулевые коэффициенты канонического вида формы $\mathcal{Q}$ при приведении её к главным осям (см. теорему 11.25). Случаи $n = 2, 3$ подробно рассмотрены в ??.