Московский физико-технический институт
Физтех-школа прикладной математики и информатики

АЛГЕБРА И ГЕОМЕТРИЯ

II СЕМЕСТР

Лектор: *Богданов Илья Игоревич*

h\nu

Автор: *Даниил Дрябин*
Проект на Github

весна 2020

Содержание

1 Многочлены

1.1 Кольцо многочленов

Замечание. В общем случае, определять многочлены как функции не вполне правильно. Например, полезно различать многочлены $P(x) = x$ и $Q(x) = x^p$ над полем $\mathbb{Z}_p$, однако для любого $x \in \mathbb{Z}_p$ выполнено равенство $P(x) = Q(x)$. Поэтому нам потребуется другое определение.

Определение 1.1. Пусть K — коммутативное кольцо. Последовательность $(a_0, a_1, \dots)$ элементов из K называется *финитной*, если она содержит конечное число ненулевых элементов. Обозначение финитной последовательности — (a_i).

Определение 1.2. Для финитных последовательностей (a_i) и (b_i) можно определить операции сложения и умножения:

▷ $(a_i) + (b_i) := (a_i + b_i)$

▷ $(a_i)(b_i) := (c_k)$, $c_k = \sum_{i+j=k} a_i b_j$

Замечание. Последовательность (c_k) действительно финитна: поскольку (a_i) и (b_i) финитны, то существует число $N \in \mathbb{N}$ такое, что для любого $i \in \mathbb{N}$, $i > N$, выполнены равенства $a_i = 0$ и $b_i = 0$, поэтому для любого $k \in \mathbb{N}$, $k > 2N$, выполнено равенство $c_k = 0$.

Теорема 1.1. *Пусть R — множество всех финитных последовательностей над коммутативным кольцом K. Тогда $(R, +, \cdot)$ также является коммутативным кольцом.*

Доказательство.

1. Покажем сначала, что $(R, +)$ — абелева группа, пользуясь тем, что $(K, +)$ — абелева группа:

   ▷ $\forall (a_i), (b_i) \in R : (a_i) + (b_i) = (a_i + b_i) = (b_i + a_i) = (b_i) + (a_i)$

   ▷ $\forall (a_i), (b_i), (c_i) \in R : ((a_i) + (b_i)) + (c_i) = (a_i + b_i + c_i) = (a_i) + ((b_i) + (c_i))$

   ▷ $\exists 0 := (0, 0, 0, \dots) \in R : \forall (a_i) \in R : (a_i) + 0 = (a_i)$

   ▷ $\forall (a_i) \in R : \exists - (a_i) = (-a_i) \in R : (a_i) + (-(a_i)) = 0$

   Последние два свойства достаточно проверять «с одной стороны» в силу коммутативности сложения в R.

2. Покажем теперь, что $(R, +, \cdot)$ — коммутативное кольцо. Это, в свою очередь, следует из того, что $(K, +, \cdot)$ — коммутативное кольцо:

   ▷ $\forall (a_i), (b_i) \in R : (a_i)(b_i) = \left(\sum_{j+k=i} a_j b_k \right) = (b_i)(a_i)$

   ▷ Заметим, что для любых $(a_i), (b_i), (c_i) \in R$ выполнены следующие равенства:

   $$((a_i)(b_i))(c_i) = \left(\sum_{j+k=i} a_j b_k \right) (c_i) = \left(\sum_{l+m=i} \left(\sum_{j+k=l} a_j b_k \right) c_m \right) = \left(\sum_{j+k+m=i} a_j b_k c_m \right)$$

   Поскольку последовательность $(a_i)((b_i)(c_i))$ можно привести к такому же виду, то $((a_i)(b_i))(c_i) = (a_i)((b_i)(c_i))$.

- $\triangleright$ $(a_i),(b_i),(c_i) \in R : (a_i)((b_i)+(c_i)) = \left(\sum_{j+k=i}(a_j b_k + a_j c_k)\right) = (a_i)(b_i) + (a_i)(c_i)$
- $\triangleright$ $\exists 1 := (1,0,0,\dots) \in R : \forall (a_i) \in R : (a_i)1 = (a_i)$

Последние два свойства также достаточно проверять «с одной стороны» в силу коммутативности умножения в R. $\square$

Замечание. Положим $x := (0,1,0,0,\dots)$, тогда $x^k = (0,\dots,0,\overset{(k)}{1},0,\dots)$ по правилам умножения в R. В таких обозначениях $(a_i) = (a_1,\dots,n,0,\dots)$ можно представить в виде $(a_i) = a_0 + a_1 x + \dots + a_n x^n$, где $a_i \equiv (a_i,0,0,\dots)$ для каждого $i \in \{1,\dots,n\}$. Более того, такое представление единственно: если $(a_i) = b_0 + b_1 x + \dots + b_n x^n$, то $b_i = a_i$ для каждого $i \in \{1,\dots,n\}$.

Определение 1.3. Пусть K — коммутативное кольцо. Кольцо финитных последовательностей элементов из K называется *кольцом многочленов над* K. Обозначение — $K[x]$.

Определение 1.4. Пусть K — коммутативное кольцо. *Степенью многочлена* $P \in K[x]$ называется позиция последнего ненулевого элемента в P. Обозначение — $\deg P$. Считается также, что $\deg 0 = -\infty$.

Замечание. Если не требовать от последовательностей финитности, то построенное аналогичным образом кольцо будет называться *кольцом формальных степенных рядов над* K. Обозначение — $K[[x]]$.

Определение 1.5. Коммутативное кольцо K называется *целостным*, если для любых элементов $a,b \in K\backslash\{0\}$ выполнено $ab \neq 0$.

Пример. Рассмотрим несколько примеров целостных колец:

- $\triangleright$ Поле F является целостным кольцом: если для некоторых $a,b \in F^*$ выполнено равенство $ab = 0$, то, умножая обе его части на a^{-1}, получим, что $b = 0$ — противоречие
- $\triangleright$ Кольцо $\mathbb{Z}$ является целостным

Замечание. В отличие от поля $\mathbb{Z}_p$ при простом p, кольцо $\mathbb{Z}_n$ при составном n не является целостным: если $n = ab$ для некоторых $a,b \in \mathbb{N}$ таких, что $a,b > 1$, то $\overline{a},\overline{b} \neq \overline{0}$, но $\overline{a}\overline{b} = \overline{0}$

Утверждение 1.1. *В целостном кольце K можно «сокращать», то есть для любых $a,b,c \in K$ таких, что $ab = ac$ и $a \neq 0$, выполнено $b = c$.*

Доказательство. Поскольку $a(b-c) = 0$ и кольцо K — целостное, то один из множителей a, $(b-c)$ равен 0. По условию, $a \neq 0$, поэтому $b - c = 0$, откуда $b = c$. $\square$

Утверждение 1.2. *Пусть K — коммутативное кольцо, $P,Q \in K[x]$. Тогда:*

1. $\deg(P+Q) \leqslant \max\{\deg P, \deg Q\}$
2. $\deg PQ \leqslant \deg P + \deg Q$, *причем если K — целостное, то* $\deg PQ = \deg P + \deg Q$

Доказательство.

1. Пусть $n := \max\{\deg P, \deg Q\}$, тогда для любого $i \in \mathbb{N}$, $i > n$, выполнено равенство $p_i + q_i = 0$.

2. Положим $n := \deg P$, $m := \deg Q$ и представим многочлены P и Q в виде $\sum_{i=0}^{n} p_i x^i$ и $\sum_{j=0}^{m} q_j x^j$ соответственно. Тогда выполнено следующее равенство:

$$PQ = \sum_{i=0}^{n} \sum_{j=0}^{m} p_i q_j x^{i+j}$$

Значит, $\deg PQ \leqslant m + n$. Более того, коэффициент при x^{n+m} равен $p_n q_m$, поэтому если K — целостное, то $p_n q_m \neq 0$ и $\deg(PQ) = m + n$. □

Следствие. *Если кольцо K — целостное, то кольцо $K[x]$ — тоже целостное.*

Доказательство. Пусть $P, Q \in K[x] \backslash \{0\}$, тогда $\deg P, \deg Q \geqslant 0$. Но тогда выполнено равенство $\deg PQ = \deg P + \deg Q \geqslant 0$, откуда $PQ \neq 0$. □

Замечание. Если F — поле, то $F[x]$ — алгебра над F.

Определение 1.6. *Гомоморфизмом колец* R и S называется отображение $\varphi : R \to S$ такое, что для любых элементов $a, b \in R$ выполнены равенства $\varphi(a + b) = \varphi(a) + \varphi(b)$и $\varphi(ab) = \varphi(a)\varphi(b)$, а также $\varphi(1) = 1$.

Определение 1.7. *Гомоморфизмом алгебр* R и S над полем F называется отображение $\varphi : R \to S$, являющееся одновременно линейным отображением и гомоморфизмом колец.

Утверждение 1.3. *Пусть A — алгебра над полем F, $a \in A$. Тогда существует единственный гомоморфизм алгебр $\varphi : F[x] \to A$ такой, что $\varphi(x) = a$.*

Доказательство. Покажем, что искомый гомоморфизм φ не более чем единственен. Если он существует, то, в силу свойств гомоморфизма колец, для любого $n \in \mathbb{N} \cup \{0\}$ выполнено равенство $\varphi(x^n) = a^n$, тогда для любого многочлена $P = p_0 + \dots + p_n x^n \in F[x]$ значение $\varphi(P)$ определяется однозначно:

$$\varphi(P) = \sum_{i=0}^{n} p_i a^i$$

Покажем теперь, что определенное таким образом отображение действительно является гомоморфизмом алгебр. Оно, очевидно, является линейным отображением, и, кроме того, $\varphi(1) = 1$. Остается проверить лишь свойство мультипликативности. Действительно, для любых $P = p_0 + \dots + p_n x^n, Q = q_0 + \dots + q_m x^m \in F[x]$ выполнено следующее:

$$\varphi(P)\varphi(Q) = \sum_{i=0}^{n} p_i a^i \sum_{j=0}^{m} q_j a^j = \sum_{k=0}^{n+m} \left(\sum_{i+j=k} p_i q_j \right) a^k = \varphi(PQ)$$

Таким образом, φ — гомоморфизм алгебр. □

Определение 1.8. Пусть A — алгебра над полем F. *Значением многочлена $P \in F[x]$ в точке $a \in A$* называется $P(a) := \varphi(P)$, где φ — *гомоморфизм подстановки* из утверждения выше.

Замечание. Для любых многочленов $P, Q \in F[x]$ и любого $a \in A$ выполнены следующие равенства:

- ▷ $(PQ)(a) = P(a)Q(a)$

▷ $(P+Q)(a) = P(a) + Q(a)$

Пример. Пусть $A = \mathcal{L}(V)$, где V — некоторое линейное пространство над полем F, $\Theta \in A$ и $P(x) = x^2 + 3x + 2$. Тогда $\varphi(\Theta) = \Theta^2 + 3\Theta + 2 = (\Theta + 1)(\Theta + 2)$, причем под единицей в A понимается тождественное отображение id.

1.2 Делимость многочленов

Определение 1.9. Пусть K — коммутативное кольцо, $a, b \in K$. Говорят, что a *делит* b, или b *делится на* a, если существует элемент $c \in K$ такой, что $ac = b$. Обозначение — $a \mid b$, или $b \,\vdots\, a$.

Определение 1.10. Пусть K — коммутативное кольцо, $a, b \in K$. Элемент $c \in K$ называется *наибольшим общим делителем* элементов a и b, если выполнены следующие условия:

1. $c \mid a, b$
2. Для любого $d \in K$ такого, что $d \mid a, b$, выполнено $d \mid c$

Обозначение — $c = \text{НОД}(a, b)$.

Замечание. Наибольший общий делитель двух элементов произвольного коммутативного кольца не всегда существует и не всегда единственен.

Определение 1.11. Пусть K — коммутативное кольцо. Элементы $a, b \in K$ называются *ассоциированными*, если существует $\alpha \in K^*$ такое, что $a = \alpha b$.

Замечание. Если a и b ассоциированы, то для любого элемента $c \in K$ выполнены равносильности $a \mid c \Leftrightarrow b \mid c$ и $c \mid a \Leftrightarrow c \mid b$.

Пример. Справедливы следующие утверждения о делимости:

▷ Если K — коммутативное кольцо, $a \in K$ и $0 \mid a$, то $a = 0$.

▷ Если K — коммутативное кольцо, то $\forall a \in K : a \mid 0$.

▷ $2 \nmid 3$ в $\mathbb{Z}$, но $2 \mid 3$ в $\mathbb{Q}$.

▷ Если K — коммутативное кольцо и $a, b \in K$, то $\text{НОД}(a, b) = a \Leftrightarrow a \mid b$.

Утверждение 1.4. *Пусть K — целостное кольцо, $a, b \in K$. Тогда любые два наибольших общих элементов делителя a и b ассоциированы.*

Доказательство. Пусть $c = \text{НОД}(a, b)$ и $d = \text{НОД}(a, b)$. Тогда, по определению наибольшего общего делителя, $c \mid d$ и $d \mid c$, то есть $c = \alpha d$ и $d = \beta c$ для некоторых $\alpha, \beta \in K$. Следовательно, $\alpha\beta c = c$. Если $c \neq 0$, то $\alpha\beta = 1$ и $\alpha, \beta \in K^*$. Если же $c = 0$, то $a = b = c = d = 0$. □

Замечание. Пусть F — поле. Поскольку для любых $P, Q \in F[x]$ выполнено равенство $\deg PQ = \deg P + \deg Q$, то $F[x]^* = F^*$, то есть обратимы лишь многочлены, являющиеся ненулевыми скалярами. Значит, ассоциированные многочлены в $F[x]$ отличаются умножением на ненулевой скаляр.

Теорема 1.2. *Пусть F — поле, $A, B \in F[x]$ и $B \neq 0$. Тогда существует единственная пара многочленов $Q, R \in F[x]$ такая, что $A = QB + R$ и $\deg R < \deg B$.*

Доказательство.

1. Пусть $n := \deg A$, $k := \deg B$. Докажем существование индукцией по n. База, $n < k$, тривиальна: $A = 0B + A$. Теперь докажем переход, $n \geqslant k$. Перепишем A в виде $A = ax^n + A'$, B — в виде $B = bx^k + B'$, где $A', B' \in F[x]$ — многочлены такие, что $\deg A' < n$, $\deg B' < k$. Определим многочлен $C \in F[x]$ следующим образом:

$$C := A - ab^{-1}x^{n-k}B = A' - ab^{-1}x^{n-k}B'$$

   Поскольку $\deg C \leqslant \deg A' < n$, то, по предположению, существуют многочлены $Q', R' \in F[x]$ такие, что $C = Q'B + R'$, тогда $A = (Q' + ab^{-1}x^{n-k})B + R'$.

2. Покажем, что набор (Q, R) единственен. Пусть $A = Q_1B + R_1 = Q_2B + R_2$ для некоторых двух наборов многочленов (Q_1, R_1) и (Q_2, R_2), тогда $(Q_1 - Q_2)B = R_1 - R_2$. Поскольку $\deg(R_1 - R_2) < \deg B$, то равенство может выполняться только в том случае, когда $Q_1 - Q_2 = 0$, откуда $Q_1 = Q_2$ и $R_1 = R_2$. □

Определение 1.12. Пусть F — поле, $A, B \in F[x]$, $B \neq 0$, а многочлены $Q, R \in F[x]$ таковы, что $A = QB + R$ и $\deg R < \deg B$. Многочлен Q называется *неполным частным*, а R — *остатком* при делении A на B.

Теорема 1.3 (алгоритм Евклида). *Пусть F — поле, $A, B \in F[x]$. Тогда существует многочлен $C = \text{НОД}(A, B)$, причем для некоторых $\exists P, Q \in F[x]$ выполнено равенство $C = AP + BQ$.*

Доказательство. Проведем индукцию по величине $k := \min\{\deg A, \deg B\}$. База, $k = -\infty$, тривиальна: если без ограничения общности $B = 0$, то $\text{НОД}(A, 0) = A$ и $A = 1A + 0B$. Докажем переход. Пусть без ограничения общности $\deg A \geqslant \deg B = k$. Выберем многочлены $Q, R \in F[x]$ такие, что $A = QB + R$ и $\deg R < k$, и заметим, что выполнена равносильность $D \mid A, B \Leftrightarrow D \mid B, R$. Тогда, по предположению индукции, существует многочлен $C = \text{НОД}(B, R) = \text{НОД}(A, B)$ и многочлены $P', Q' \in F[x]$ такие, что $C = P'B + Q'R$. Тогда $C = P'B + Q'R = Q'A + (P' - QQ')B$. □

Следствие. *Пусть F — поле, $A, B \in F[x]$, и имеют место представления $A = BQ + R$, $B = Q_1R + R_1$, $\ldots$, $R_{k-1} = Q_{k+1}R_k + 0$. Тогда наибольший общий делитель многочленов A и B можно вычислить следующим образом:*

$$\textit{НОД}(A, B) = \textit{НОД}(B, R) = \textit{НОД}(R, R_1) = \cdots = \textit{НОД}(R_k, 0) = R_k$$

Определение 1.13. Пусть F — поле, $P \in F[x]$. Многочлен P называется *неприводимым над F*, если $\deg P > 0$ и P не раскладывается в произведение двух многочленов положительной степени.

Пример. Многочлен $x^2 + 1$ неприводим над $\mathbb{R}$, но приводим над $\mathbb{C}$.

Замечание. Пусть F — поле, $P, Q \in F[x]$. Тогда выполнены следующие свойства:

- ▷ Если P неприводим, то либо $\text{НОД}(P, Q) = 1$, либо $\text{НОД}(P, Q) = P$ с точностью до ассоциированности

- ▷ Если P, Q неприводимы и $P \mid Q$, то P и Q ассоциированы

Утверждение 1.5. *Пусть F — поле, $Q \in F[x]$, $\deg Q > 0$. Тогда Q раскладывается в произведение неприводимых многочленов.*

Доказательство. Докажем утверждение индукцией по $\deg Q$. База тривиальна: если $\deg Q = 1$, то Q уже неприводим. Докажем переход. Если Q неприводим, то получено требуемое, иначе — выполнено равенство $Q = Q_1 Q_2$ для некоторых $Q_1, Q_2 \in F[x]$ таких, что $0 < \deg Q_1, \deg Q_2 < \deg Q$, тогда Q_1 и Q_2 представляются в виде произведения неприводимых по предположению индукции. □

Утверждение 1.6. *Пусть F — поле, $P, Q, R \in F[x]$, многочлен P неприводим и выполнено $P \mid QR$. Тогда $P \mid Q$ или $P \mid R$.*

Доказательство. Предположим, что $P \nmid Q$. Тогда, в силу неприводимости многочлена P, выполнено равенство НОД$(P, Q) = 1$, поэтому существуют многочлены $K, L \in F[x]$ такие, что $KP + LQ = 1$. Умножая обе части равенства на R, получим, что $KPR + LQR = R$, откуда $P \mid KPR + LQR = R$. □

Замечание. Утверждение выше легко обобщить: если $P, Q_1, \ldots, Q_n \in F[x]$, многочлен P неприводим и выполнено $P \mid Q_1 \cdots Q_n$, то существует $i \in \{1, \ldots, n\}$ такое, что $P \mid Q_i$. Для доказательства достаточно провести индукцию по n.

Теорема 1.4 (основная теорема арифметики для многочленов)**.** *Пусть F — поле, и $Q \in F[x] \setminus \{0\}$. Тогда существует такой скаляр $\alpha \in F^*$ и такие неприводимые многочлены $P_1, \ldots, P_k \in F[x]$, что Q можно представить в следующем виде:*

$$Q = \alpha P_1 \ldots P_k$$

Более того, если $Q = \alpha P_1 \ldots P_k = \beta R_1 \ldots R_l$ для некоторого скаляра $\beta \in F^*$ и неприводимых многочленов $R_1, \ldots, R_l \in F[x]$, то $k = l$ и существует перестановка $\sigma \in S_k$ такая, что для каждого $i \in \{1, \ldots, k\}$ многочлены P_i и $R_{\sigma(i)}$ ассоциированы.

Доказательство.

- ▷ (Существование) Случай, когда $\deg Q > 0$, уже был рассмотрен. Если же $\deg Q = 0$, то $Q = \alpha$.
- ▷ (Единственность) Проведем индукцию по k. База, $k = 0$, тривиальна: $\deg Q = 0$, откуда $k = l = 0$ и $Q = \alpha = \beta$. Теперь докажем переход. Пусть $k > 0$ и выполнены равенства $Q = \alpha P_1 \ldots P_k = \beta R_1 \ldots R_l$. Тогда, поскольку $P_k \mid R_1 \ldots R_l$, существует $i \in \{1, \ldots, l\}$ такое, что $P_k \mid R_i$, то есть многочлены P_k и R_i ассоциированы в силу их неприводимости: $R_i = \gamma P_k$, $\gamma \in F^*$. Пусть без ограничения общности $i = l$, тогда $\alpha P_1 \ldots P_{k-1} = (\beta\gamma) Q_1 \ldots Q_{l-1}$, и применимо предположение индукции. □

Следствие. *Пусть $A, B \in F[x]$, $A = \alpha P_1 \ldots P_k$, $B = \beta Q_1 \ldots Q_l$ — разложения многочленов A, B на неприводимые сомножители, и все многочлены $P_1, \ldots, P_k, Q_1, \ldots, Q_l \in F[x]$ попарно неассоциированы. Тогда НОД$(A,B) = 1$.*

Доказательство. Пусть это не так, тогда НОД$(A, B) = C$ для некоторого $C \in F[x]$ такого, что $\deg C > 0$. Но тогда существует неприводимый многочлен $P \in F[x]$ такой, что $P \mid C$, откуда $P \mid A, B$, поэтому существуют индексы $i \in \{1, \ldots, k\}$ и $j \in \{1, \ldots, l\}$ такие, что P_i, Q_j ассоциированы с P, — противоречие. □

Следствие. *Пусть* $A = \alpha P_1 \dots P_k$, $B = \beta Q_1 \dots Q_l$ *— разложения многочленов* $A, B \in F[x]$ *на неприводимые сомножители, и существуют индексы* $i \in \{1, \dots, k\}$ *и* $j \in \{1, \dots, l\}$ *такие, что* P_i *ассоциирован с* Q_j*. Тогда выполнено следующее равенство:*

$$\textit{НОД}(A, B) = \textit{НОД}\left(\frac{A}{P_i}, \frac{B}{Q_j}\right) P_i$$

Определение 1.14. Пусть R — целостное кольцо. Элемент $p \in R \backslash \{0\}$ называется *простым*, если p необратим и не раскладывается в произведение двух необратимых.

Определение 1.15. Целостное кольцо R называется *факториальным*, если любой его необратимый элемент раскладывается в произведение простых единственным образом с точностью до перестановки и ассоциированности.

Определение 1.16. *Нормой* на целостном кольце R называется функция $N : R \to \mathbb{N} \cup \{0\}$ такая, что выполнены следующие условия:

1. $N(a) = 0 \Leftrightarrow a = 0$
2. $N(a + b) \leqslant N(a) + N(b)$
3. $N(ab) = N(a)N(b)$

Определение 1.17. Целостное кольцо R называется *евклидовым* относительно нормы N, если для любых элементов $a \in R$, $b \in K \backslash \{0\}$, существуют элементы $\exists q, r \in R$ такие, что $a = qb + r$ и $N(r) < N(b)$.

Пример. Рассмотрим несколько примеров евклидовых колец:

- ▷ Кольцо $\mathbb{Z}$ является евклидовым относительно нормы N такой, что $N(a) := |a|$ для любого $a \in \mathbb{Z}$
- ▷ Если F — поле, то $F[x]$ является евклидовым относительно нормы N такой, что $N(P) := 2^{\deg P}$ для любого $P \in F[x]$

Замечание. Рассуждения, приведенные в данном разделе, позволяют аналогичным образом доказать, что любое евклидово кольцо является факториальным.

1.3 Корни многочленов

До конца раздела зафиксируем поле F.

Определение 1.18. Пусть $P \in F[x]$. Скаляр $a \in F$ называется *корнем* многочлена P, если выполнено равенство $P(a) = 0$.

Теорема 1.5 (Безу). *Скаляр* $a \in F$ *является корнем многочлена* $P \in F[x] \Leftrightarrow (x - a) \mid P$.

Доказательство. Разделим P с остатком на $(x-a)$, то есть выберем $Q, R \in F[x]$ такие, что $P = Q(x - a) + R$ и $\deg R \leqslant 0$. Заметим, что $P(a) = R$, тогда выполнены равносильности $P(a) = 0 \Leftrightarrow R = 0 \Leftrightarrow (x - a) \mid P$. □

Определение 1.19. Пусть $a \in F$ — корень многочлена $P \in F[x]$. *Кратностью* корня a называется наибольшее $\gamma \in \mathbb{N}$ такое, что $(x - a)^\gamma \mid P$. Если $\gamma > 1$, то корень a называется *кратным*, иначе — *простым*.

Теорема 1.6. *Пусть $P \in F[x]\backslash\{0\}$, и $a_1, \ldots, a_k$ — корни многочлена P, имеющие кратности $\gamma_1, \ldots, \gamma_k \in \mathbb{N}$. Тогда $\gamma_1 + \cdots + \gamma_k \leqslant \deg P$.*

Доказательство. По условию, для любого $i \in \{1, \ldots, k\}$ выполнено $(x - a_i)^{\gamma_i} \mid P$, причем для любых индексов $i, j \in \{1, \ldots, k\}$, $i \neq j$, выполнены следующие равенства:

$$\text{НОД}(x - a_i, x - a_j) = \text{НОД}(x - a_i, a_i - a_j) = 1$$

Значит, многочлены $(x - a_1)^{\gamma_1}, \ldots, (x - a_k)^{\gamma_k}$ попарно неассоциированы, тогда, в силу единственности разложения многочлена P на неприводимые сомножители, выполнено неравенство $\gamma_1 + \cdots + \gamma_k \leqslant \deg P$. □

Замечание. В коммутативном кольце, не являющемся целостным данная теорема неверна, поскольку неверна единственность разложения на неприводимые сомножители. Например, в кольце $\mathbb{Z}_4$ у многочлена $P = x^2 = (x-2)^2$ степени 2 есть корень 0 кратности 2 и корень 2 кратности 2.

Замечание. Над полем $\mathbb{C}$ число корней любого ненулевого многочлена с учетом кратности равно его степени. Это утверждение называется *основной теоремой алгебры*, но в рамках данного курса мы не будем его доказывать.

Определение 1.20. Пусть $P = p_0 + p_1x + \cdots + p_nx^n \in F[x]$. *Формальной производной* многочлена $P(x)$ называется многочлен $P' := p_1 + 2p_2x + \cdots + np_nx^{n-1}$, где целочисленные скаляры понимаются как суммы соответствующего числа единиц.

Утверждение 1.7. *Формальная производная обладает следующими свойствами:*

1. *$\forall \alpha, \beta \in F : \forall P, Q \in F[x] : (\alpha P + \beta Q)' = \alpha P' + \beta Q'$ (линейность)*
2. *$\forall P, Q \in F[x] : (PQ)' = P'Q + PQ'$ (правило Лейбница)*

Доказательство.

1. Пусть $n := \max\{\deg P, \deg Q\}$, тогда многочлены P и Q можно представить в виде $P = \sum_{i=0}^n p_i x^i$ и $Q = \sum_{i=0}^n q_i x^i$, откуда $\alpha P + \beta Q = \sum_{i=0}^n (\alpha p_i + \beta q_i) x^i$. Проверим требуемое равенство непосредственной проверкой:

$$(\alpha P + \beta Q)' = \sum_{i=1}^n i(\alpha p_i + \beta q_i)x^{i-1} = \alpha \sum_{i=1}^n i p_i x^{i-1} + \beta \sum_{i=1}^n i q_i x^{i-1} = \alpha P' + \beta Q'$$

2. Левая и правая части требуемого равенства линейны по P и по Q, поэтому равенство достаточно проверить на некотором базисе пространства многочленов, например, для произвольных многочленов вида $P(x) = x^i$, $Q(x) = x^j$, $i, j \in \mathbb{N} \cup \{0\}$:

$$(PQ)' = (i+j)x^{i+j-1} = ix^{i-1}x^j + jx^ix^{j-1} = P'Q + PQ'$$ □

Замечание. Формальная производная не обладает аналитическими свойствами. В $\mathbb{Z}_p[x]$, например, выполнены равенства $(x^p)' = px^{p-1} = 0$.

Следствие. *Формальная производная обладает следующими свойствами:*

1. *$\forall P_1, \ldots, P_n \in F[x] : (P_1P_2 \ldots P_n)' = P_1'P_2 \ldots P_n + P_1P_2' \ldots P_n + \cdots + P_1P_2 \ldots P_n'$*

2. $\forall P \in F[x] : \forall n \in \mathbb{N} : (P^n)' = nP^{n-1}P'$

3. $\forall P, Q \in F[x] : \big(P(Q)\big)' = P'(Q)Q'$

Доказательство.

1. Достаточно провести индукцию по n.

2. Достаточно применить первое равенство к многочлену P^n.

3. Считая, что $P(x) = p_0 + p_1x + \dots + p_nx^n$, воспользуемся вторым равенством:

$$\big(P(Q)\big)' = \left(\sum_{i=0}^{m} p_iQ^i\right)' = \sum_{i=0}^{m} ip_iQ^{i-1}Q' = P'(Q)Q' \qquad \square$$

Теорема 1.7. *Пусть $P \in F[x]$, $c \in F$. Тогда следующие условия эквивалентны:*

1. *c — кратный корень P*

2. $P(c) = P'(c) = 0$

3. $(x - c) \mid$ *НОД*(P, P')

Доказательство.

▷ (1 ⇔ 2) Пусть c — корень многочлена P, тогда $P = (x - c)Q$ и $P' = Q + (x - c)Q'$, поэтому c — кратный корень многочлена $P \Leftrightarrow Q(c) = 0 \Leftrightarrow P'(c) = 0$.

▷ (2 ⇔ 3) $P(c) = P'(c) = 0 \Leftrightarrow (x - c) \mid P, P' \Leftrightarrow (x - c) \mid$ НОД(P, P'). $\square$

Теорема 1.8. *Пусть $c \in F$ — корень многочлена $P \in F[x]$ кратности $k \in \mathbb{N}$, $k > 1$. Тогда выполнены следующие свойства:*

1. *c — корень многочлена P' кратности хотя бы $k - 1$*

2. *Если $\operatorname{char} F > k$ или $\operatorname{char} F = 0$, то c — корень многочлена P' кратности ровно $k - 1$*

Доказательство. Многочлен P имеет вид $(x - c)^kQ$ для некоторого $Q \in F[x]$ такого, что $(x - c) \nmid Q$. Тогда:

$$P' = k(x - c)^{k-1}Q + (x - c)^kQ' = (x - c)^{k-1}\big(kQ + (x - c)Q'\big)$$

Из равенства выше уже следует, что c — корень многочлена P' кратности хотя бы $k - 1$. Рассмотрим теперь многочлен $kQ + (x - c)Q'$. Если $\operatorname{char} F > k$ или $\operatorname{char} F = 0$, то $kQ(c) \neq 0$, поэтому кратность корня c у многочлена P' равна $k - 1$. $\square$

Следствие. *Пусть $c \in F$ — корень многочлена $P \in F[x]$ кратности $k \in \mathbb{N}$, $k > 1$. Тогда выполнены равенства $P(c) = P'(c) = \dots = P^{(k-1)}(c) = 0$.*

Доказательство. Заметим, что $(x - c)^k \mid P \Rightarrow (x - c)^{k-1} \mid P' \Rightarrow \dots \Rightarrow (x - c) \mid P^{(k-1)}$. $\square$

Следствие. *Пусть $k \in \mathbb{N}$, $k > 1$, $\operatorname{char} F \geqslant k$ или $\operatorname{char} F = 0$, и пусть для многочлена $P \in F[x]$ выполнены равенства $P(c) = \dots = P^{(k-1)}(c) = 0$. Тогда c — корень многочлена P кратности хотя бы k.*

Доказательство. Предположим, что c — корень кратности $l < k$ многочлена P. Тогда c является простым корнем многочлена $P^{(l-1)}$, откуда $P^{(l)}(c) \neq 0$ — противоречие. □

Следствие (теорема Вильсона)**.** *Пусть p — простое число. Тогда выполнено следующее:*

$$(p-1)! \equiv_p -1$$

Доказательство. Рассмотрим многочлен $P = x^{p-1} - 1 \in \mathbb{Z}_p[x]$. Его производная P' равна $-x^{p-2}$. Заметим, что НОД$(P, P') = 1$, так как все делители многочлена P' имеют вид x^k, а 0 не является корнем P. Значит, все корни многочлена P — простые, причем, по малой теореме Ферма, его корнями являются все элементы $1, \ldots, p-1 \in \mathbb{Z}_p$. Тогда, поскольку степень многочлена P равна $p-1$, выполнено следующее равенство:

$$x^{p-1} - 1 = (x-1)(x-2)\ldots(x-(p-1))$$

Поскольку левая и правая части — это один и тот же многочлен в $\mathbb{Z}_p$, то в $\mathbb{Z}_p$ выполнено равенство $(-1)^{p-1}(p-1)! = -1$, то есть $(-1)^{p-1}(p-1)! \equiv_p -1$. Наконец, $1 \equiv_2 -1$, и все простые числа, отличные от 2, нечетны, поэтому $(p-1)! \equiv_p -1$. □

Следствие. *Пусть p — простое число. Тогда для любого $x \in \mathbb{Z}$ выполнено следующее:*

$$x^p - 1 \equiv_p (x-1)^p$$

Доказательство. Рассмотрим многочлен $Q = x^p - 1 \in \mathbb{Z}_p[x]$. Все его производные тождественно равны нулю, поэтому $\overline{1}$ — корень кратности хотя бы p этого многочлена. Тогда, поскольку степень многочлена Q равна p, для любого $x \in \mathbb{Z}_p$ выполнено равенство $x^p - 1 = (x-1)^p$, то есть для любого $x \in \mathbb{Z}$ выполнено сравнение $x^p - 1 \equiv_p (x-1)^p$. □

2 Линейные операторы

2.1 Инвариантные подпространства

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$.

Определение 2.1. Пусть $\varphi \in \mathcal{L}(V)$. Подпространство $U \leqslant V$ называется *инвариантным* относительно преобразования φ, если $\varphi(U) \leqslant U$.

Пример. Рассмотрим несколько примеров инвариантных подпространств: Инвариантными подпространствами относительно соответствующих преобразований являются:

- ▷ Прямая l в плоскости V_2 и прямая $n \perp l$ инвариантны относительно $\varphi \in \mathcal{L}(V_2)$, где φ — симметрия относительно l
- ▷ Пусть $k \in \mathbb{N}$, тогда подпространство $P_k := \{P \in F[x] : \deg P \leqslant k\} \leqslant F[x]$ инвариантно относительно $\varphi \in \mathcal{F}[x]$, где φ — формальное дифференцирование

Утверждение 2.1. *Пусть $\varphi \in \mathcal{L}(V)$, $e = (\overline{e_1}, \ldots, \overline{e_n})$ — базис в V, $\varphi \leftrightarrow_e A \in M_n(F)$, $k \in \mathbb{N}$, $k \leqslant n$. Тогда $U := \langle \overline{e_1}, \ldots, \overline{e_k} \rangle \leqslant V$ инвариантно относительно $\varphi \Leftrightarrow A$ имеет следующий вид для некоторых $B \in M_k(F)$, $D \in M_{n-k}(F)$:*

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

Доказательство. U инвариантно относительно $\varphi \Leftrightarrow \forall i \in \{1, \ldots, k\} : \varphi(\overline{e_i}) \in U \Leftrightarrow A$ имеет требуемый вид. □

Замечание. В утверждении выше матрица B является матрицей оператора $\varphi|_U \in \mathcal{L}(U)$ в базисе $(\overline{e_1}, \ldots, \overline{e_k})$.

Утверждение 2.2. *Пусть $\varphi \in \mathcal{L}(V)$, $U_1, U_2 \leqslant V$ — инвариантные относительно φ подпространства. Тогда подпространства $U_1 + U_2$ и $U_1 \cap U_2$ тоже инвариантны относительно φ.*

Доказательство.

- ▷ $\varphi(U_1 + U_2) = \varphi(U_1) + \varphi(U_2) \leqslant U_1 + U_2$
- ▷ $\varphi(U_1 \cap U_2) \leqslant \varphi(U_1) \cap \varphi(U_2) \leqslant U_1 \cap U_2$ □

Утверждение 2.3. *Пусть $\varphi \in \mathcal{L}(V)$, подпространства $U, W \leqslant V$ таковы, что $U \leqslant \operatorname{Ker} \varphi$ и $\operatorname{Im} \varphi \leqslant W \leqslant V$. Тогда U и W инвариантны относительно φ.*

Доказательство.

- ▷ $\varphi(U) \leqslant \varphi(\operatorname{Ker} \varphi) = \{\overline{0}\} \leqslant U$
- ▷ $\varphi(W) \leqslant \operatorname{Im} \varphi \leqslant W$ □

Утверждение 2.4. *Пусть $\varphi, \psi \in \mathcal{L}(V)$, причем $\varphi \circ \psi = \psi \circ \varphi$. Тогда $\operatorname{Ker} \psi$ и $\operatorname{Im} \psi$ инвариантны относительно φ.*

Доказательство.

- ▷ Пусть $\overline{u} \in \operatorname{Ker} \psi$, тогда $\psi(\varphi(\overline{u})) = \varphi(\psi(\overline{u})) = \varphi(\overline{0}) = \overline{0}$, откуда $\varphi(\overline{u}) \in \operatorname{Ker} \psi$
- ▷ Пусть $\overline{u} \in \operatorname{Im} \psi$, тогда существует вектор $\overline{v} \in V$ такой, что выполнено равенство $\psi(\overline{v}) = \overline{u}$, откуда $\varphi(\overline{u}) = \varphi(\psi(\overline{v})) = \psi(\varphi(\overline{v})) \in \operatorname{Im} \psi$ □

Замечание. Последнее утверждение полезно в случае, когда $\psi = P(\varphi)$, $P \in F[x]$, частности, когда $\psi = \varphi - \lambda$, где $\lambda \in F$.

Замечание. Пусть $V = U \oplus W$, причем подпространства U, W инвариантны относительно $\varphi \in \mathcal{L}(V)$ и $e' = (\overline{e_1}, \ldots, \overline{e_k})$, $e'' = (\overline{e_{k+1}}, \ldots, \overline{e_n})$ — базисы в U и W, $\varphi|_U \leftrightarrow_{e'} B$, $\varphi|_W \leftrightarrow_{e''} D$. Тогда, по свойству прямой суммы, $e = (\overline{e_1}, \ldots, \overline{e_n})$ — базис в V, причем матрица оператора φ в этом базисе имеет следующий вид:

$$\varphi \leftrightarrow_e \left(\begin{array}{c|c} B & 0 \\ \hline 0 & D \end{array}\right)$$

2.2 Собственные векторы

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$.

Определение 2.2. Пусть $\varphi \in \mathcal{L}(V)$. Вектор $\overline{v} \in V \backslash \{\overline{0}\}$ называется *собственным вектором* оператора φ с *собственным значением* $\lambda \in F$, если $\varphi(\overline{v}) = \lambda \overline{v}$. Скаляр $\mu \in F$ называется *собственным значением* оператора φ, если существует собственный вектор $\overline{u} \in V \backslash \{0\}$ оператора φ с собственным значением μ.

Замечание. Пусть $\lambda \in F$ — собственное значение оператора φ. Тогда $\overline{v} \in V \backslash \{\overline{0}\}$ — собственный вектор оператора φ со значением $\lambda \Leftrightarrow \varphi(\overline{v}) = \lambda\overline{v} \Leftrightarrow \overline{v} \in \operatorname{Ker}(\varphi - \lambda)$.

Определение 2.3. Пусть $\varphi \in \mathcal{L}(V)$, $\lambda \in F$ — собственное значение оператора φ. Подпространство $V_\lambda := \operatorname{Ker}(\varphi - \lambda) \leqslant V$ называется *собственным подпространством* оператора φ, соответствующим собственному значению λ.

Замечание. Вектор $v \in V\backslash\{\overline{0}\}$ является собственным вектором оператора $\varphi \in \mathcal{L}(V)$ подпространство $\Leftrightarrow$ $\langle\overline{v}\rangle$ является инвариантным относительно φ. Значит, любое подпространство в V_λ инвариантно относительно φ.

Теорема 2.1. *Пусть $\varphi \in \mathcal{L}(V)$, $\lambda_1, \ldots, \lambda_k \in F$ — различные собственные значения оператора φ. Тогда сумма $V_{\lambda_1} + \cdots + V_{\lambda_k}$ — прямая.*

Доказательство. Проведем индукцию по k. База, $k = 1$, тривиальна, докажем переход. Пусть для некоторого $k > 1$ утверждение неверно, тогда, по критерию прямой суммы, существует индекс $i \in \{1, \ldots, k\}$ такой, что выполнено следующее:

$$V_{\lambda_i} \cap (V_{\lambda_1} + \ldots + V_{\lambda_{i-1}} + V_{\lambda_{i+1}} + \cdots + V_{\lambda_k}) \neq \{\overline{0}\}$$

Пусть без ограничения общности $i = k$, тогда существуют векторы $\overline{v_1} \in V_{\lambda_1}, \cdots, \overline{v_k} \in V_{\lambda_k}$ такие, что выполнено следующее:

$$\overline{v_1} + \cdots + \overline{v_{k-1}} = \overline{v_k} \neq \overline{0}$$

Применим к равенству выше оператор φ, и вычтем из полученного равенства исходное, умноженное на λ_k, тогда:

$$(\lambda_1 - \lambda_k)\overline{v_1} + \cdots + (\lambda_{k-1} - \lambda_k)\overline{v_{k-1}} = (\lambda_k - \lambda_k)\overline{v_k} = \overline{0}$$

Все коэффициенты в левой части по условию отличны от нуля, а также хотя бы один из векторов в левой части — ненулевой, поскольку сумма этих векторов равна $\overline{v_k} \neq \overline{0}$. Получено нетривиальное разложение нуля, что невозможно по предположению индукции. Значит, сумма $V_{\lambda_1} + \cdots + V_{\lambda_k}$ — прямая. $\square$

Следствие. *Количество собственных значений оператора $\varphi \in \mathcal{L}(V)$ не превосходит величины* $\dim V$.

Доказательство. Если собственных значений у φ больше, чем $\dim V$, то соответствующие им собственные подпространства образуют прямую сумму размерности большей, чем $\dim V$, что невозможно. $\square$

Замечание. Пусть $V_{\lambda_1}, \ldots, V_{\lambda_k} \leqslant V$ — собственные подпространства оператора φ. Поскольку $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$ — прямая сумма, то объединение базисов в этих подпространствах можно дополнить до базиса e в V. В полученном базисе матрица преобразования φ принимает следующий вид:

$$\varphi \leftrightarrow_e \left(\begin{array}{ccc|ccc}
\lambda_1 & \ldots & 0 & * & \ldots & * \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \ldots & \lambda_k & \vdots & \ddots & \vdots \\
\hline
0 & \ldots & 0 & \vdots & \ddots & \vdots \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & * & \ldots & *
\end{array}\right)$$

Для каждого индекса $i \in \{1, \dots, k\}$ значение λ_i встречается в диагональном блоке матрицы выше ровно $\dim V_{\lambda_i}$ раз.

Определение 2.4. Пусть $A \in M_n(F)$. *Характеристическим многочленом* матрицы A называется многочлен $\chi_A(\lambda) := |A - \lambda E|$.

Замечание. Степень характеристического многочлена χ_A равна n, поскольку единственное слагаемое с λ^n в формуле определителя получается при $\sigma = \mathrm{id}$, когда значение $(-\lambda)$ перемножается n раз. В частности, коэффициент при λ^n равен $(-1)^n$.

Утверждение 2.5. *Пусть $\varphi \in \mathcal{L}(V)$, $\varphi \leftrightarrow_e A \in M_n(F)$. Тогда скаляр $\lambda_0 \in F$ является собственным значением оператора $\varphi \Leftrightarrow \chi_A(\lambda_0) = 0 \Leftrightarrow (\lambda - \lambda_0) \mid \chi_A(\lambda)$.*

Доказательство. Скаляр λ_0 является собственным значением тогда и только тогда, когда $\mathrm{Ker}\,(\varphi - \lambda_0) \neq \{\overline{0}\}$. Выполнены следующие равносильности:

$$\mathrm{Ker}\,(\varphi - \lambda_0) \neq \{\overline{0}\} \Leftrightarrow \mathrm{rk}\,(A - \lambda_0 E) < n \Leftrightarrow |A - \lambda_0 E| = 0 \Leftrightarrow \chi_A(\lambda_0) = 0$$ □

Определение 2.5. Матрицы $A, B \in M_n(F)$ называются *подобными*, если существует матрица $S \in \mathrm{GL}_n(F)$ такая, что $B = S^{-1}AS$.

Замечание. Подобные матрицы — это матрицы одного и того же оператора в разных базисах.

Утверждение 2.6. *Пусть $A, B \in M_n(F)$ — подобные матрицы. Тогда $\chi_A(\lambda) = \chi_B(\lambda)$.*

Доказательство. Зафиксируем значение $\lambda \in F$, тогда выполнены следующие равенства:

$$\chi_A(\lambda) = |A - \lambda E| = \left|S^{-1}(B - \lambda E)S\right| = \left|S^{-1}\right| |B - \lambda E||S| = |B - \lambda E| = \chi_B(\lambda)$$

Получено требуемое. □

Определение 2.6. Пусть $\varphi \in \mathcal{L}(V)$. *Характеристическим многочленом* оператора φ называется характеристический многочлен его матрицы в произвольном базисе. Обозначение — $\chi_\varphi(\lambda)$.

Определение 2.7. Пусть $A = (a_{ij}) \in M_n(F)$. *Следом* матрицы A называется величина $\mathrm{tr} := \sum_{i=1}^n a_{ii}$.

Утверждение 2.7. *Пусть $A \in M_n(F)$. Тогда в характеристическом многочлене $\chi_A(\lambda)$ коэффициент при λ^{n-1} равен $(-1)^{n-1}\,\mathrm{tr}\,A$, а свободный член равен $\det A$.*

Доказательство. Во всех нетождественных перестановках степень получаемых в $\chi_A(\lambda)$ мономов не превосходит $n-2$, поэтому слагаемое с λ^{n-1} может возникнуть только при $\sigma = \mathrm{id}$, когда число $(-\lambda)$ перемножается $n-1$ раз и умножается на один из диагональных элементов, поэтому коэффициент при λ^{n-1} равен $(-1)^{n-1}\,\mathrm{tr}\,A$. Свободный член в $\chi_A(\lambda)$ равен $\chi_A(0) = |A|$. □

Следствие. *Если матрицы $A, B \in M_n(F)$ подобны, то $\mathrm{tr}\,A = \mathrm{tr}\,B$ и $\det A = \det B$.*

Определение 2.8. Пусть $\varphi \in \mathcal{L}(V)$. *Следом оператора* φ называется след матрицы φ в произвольном базисе, *определителем оператора* — определитель матрицы φ в произвольном базисе. Обозначения — $\mathrm{tr}\,\varphi$ и $\det\varphi$ соответственно.

Теорема 2.2. *Пусть $\varphi \in \mathcal{L}(V)$, и $\chi_\varphi(\lambda)$ имеет n различных корней $\lambda_1, \ldots, \lambda_n$. Тогда в V существует базис e, в котором матрица оператора φ имеет следующий вид:*

$$\varphi \leftrightarrow_e \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

Доказательство. Поскольку корни многочлена $\chi_\varphi(\lambda)$ — это собственные значения V, то $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n} = V$, и объединение базисов в $V_{\lambda_1}, \ldots, V_{\lambda_n}$ образует искомый базис e в V. □

Следствие. *Пусть $\varphi \in \mathcal{L}(V)$, и $\chi_\varphi(\lambda)$ имеет n различных корней $\lambda_1, \ldots, \lambda_n$. Тогда выполнены равенства $\operatorname{tr} \varphi = \sum_{i=1}^n \lambda_i$ и $\det \varphi = \prod_{i=1}^n \lambda_i$.*

Определение 2.9. Оператор $\varphi \in \mathcal{L}(V)$ называется *диагонализуемым*, если существует базис в V, в котором матрица φ имеет диагональный вид. Матрица $A \in M_n(F)$ называется *диагонализуемой*, если она подобна некоторой диагональной.

Определение 2.10. Пусть $\varphi \in \mathcal{L}(V)$, $\lambda_0 \in F$ — собственное значение оператора φ. *Алгебраической кратностью* собственного значения λ_0 называется кратность корня λ_0 в $\chi_\varphi(\lambda)$, *геометрической кратностью* — величина $\dim V_{\lambda_0} = \dim \operatorname{Ker}(\varphi - \lambda_0)$.

Теорема 2.3. *Пусть $\varphi \in \mathcal{L}(V)$, $\lambda_0 \in F$ — собственное значение оператора φ. Тогда алгебраическая кратность значения λ_0 не меньше его геометрической кратности.*

Доказательство. Пусть геометрическая кратность значения λ_0 равна $k \in \mathbb{N}$. Выберем базис $(\overline{e_1}, \ldots, \overline{e_k})$ в V_{λ_0} и дополним этот базис до базиса $e = (\overline{e_1}, \ldots, \overline{e_n})$ в V. Тогда матрица оператора φ в этом базисе имеет следующий вид для некоторой матрицы $D \in M_{n-k}(F)$:

$$\varphi \leftrightarrow_e A := \left(\begin{array}{c|c} \lambda_0 E_k & * \\ \hline 0 & D \end{array}\right)$$

По теореме об определителе с углом нулей, выполнены следующие равенства:

$$\chi_\varphi(\lambda) = |A - \lambda E_k| = |(\lambda_0 - \lambda)E_k||D - \lambda E_{n-k}| = (\lambda_0 - \lambda)^k |D - \lambda E_{n-k}|$$

Значит, λ_0 — корень кратности не меньше k в $\chi_\varphi(\lambda)$. □

Замечание. Неравенство в теореме выше может быть строгим. Рассмотрим, например, следующую матрицу:

$$A := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

Тогда $\chi_\varphi(\lambda) = \lambda^2$, поэтому 0 является корнем кратности 2 в $\chi_\varphi(\lambda)$, при этом выполнены равенства $\dim V_0 = \dim \operatorname{Ker} \varphi = 2 - \operatorname{rk} A = 1$.

Замечание. Пусть $\varphi \in \mathcal{L}(V)$, $\alpha \in F$. Тогда для оператора $\varphi - \alpha$ выполнено следующее:

$$\chi_{\varphi-\alpha}(\lambda) = |A - (\lambda + \alpha)E| = \chi_\varphi(\lambda + \alpha)$$

Значит, $\lambda_0 \in F$ — собственное значение оператора $\varphi - \alpha \Leftrightarrow \lambda_0 + \alpha$ — собственное значение оператора φ. Кроме того, собственные векторы операторов $\varphi - \alpha$ и φ совпадают.

Теорема 2.4. *Пусть* $\varphi \in \mathcal{L}(V)$, $U \leqslant V$ — *инвариантное относительно* φ *подпространство. Тогда для оператора* $\psi := \varphi|_U \in \mathcal{L}(U)$ *выполнено* $\chi_\psi \mid \chi_\varphi$.

Доказательство. Дополним базис $e' = (\overline{e_1}, \ldots, \overline{e_k})$ в U до базиса $e = (\overline{e_1}, \ldots, \overline{e_n})$ в V. Тогда в базисе e матрица оператора φ имеет следующий вид для некоторых $B \in M_k(F)$, $C \in M_{k\times(n-k)}(F)$, $D \in M_{n-k}(F)$:

$$\varphi \leftrightarrow_e A := \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

По теореме об определителе с углом нулей, $\chi_A(\lambda) = |B - \lambda E_k||D - \lambda E_{n-k}|$, тогда, поскольку $\psi = \varphi|_U \leftrightarrow_{e'} B$, выполнено соотношение $\chi_\psi \mid \chi_\varphi$. $\square$

Замечание. Предыдущую теорему можно вывести из только что доказанной. Действительно, если $V_{\lambda_0} \leqslant V$ — собственное подпространство значения $\lambda_0 \in F$, то геометрическая кратность значения λ_0 равна $\dim V_{\lambda_0} = k$, причем $\chi_{\varphi|_U}(\lambda) = (\lambda_0 - \lambda)^k \mid \chi_\varphi(\lambda)$.

Теорема 2.5. *Пусть* $\varphi \in \mathcal{L}(V)$. *Тогда равносильны следующие условия:*

1. *Оператор* φ *диагонализуем*

2. *Алгебраическая кратность каждого собственного значения оператора* φ *равна геометрической, и* χ_φ *раскладывается на линейные сомножители, то есть имеет следующий вид при некоторых* $\lambda_1, \ldots, \lambda_k \in F$ *и* $\alpha_1, \ldots, \alpha_k \in \mathbb{N}$ *таких, что* $\sum_{i=1}^k \alpha_i = n$:

$$\chi_\varphi(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\alpha_i}$$

3. $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$, *где* $V_{\lambda_1}, \ldots, V_{\lambda_k}$ — *собственные подпространства оператора* φ

4. *В* V *есть базис, состоящий из собственных векторов оператора* φ

Доказательство.

- ▷ $(1 \Rightarrow 2)$ Пусть в некотором базисе e в V матрица оператора φ имеет диагональный вид, $\lambda_1, \ldots, \lambda_k \in F$ — различные элементы на диагонали, $\alpha_1, \ldots, \alpha_k \in \mathbb{N}$ — количества их вхождений в матрицу, тогда $\chi_\varphi(\lambda) = \prod_{i=1}^k (\lambda_i - \lambda)^{\alpha_i}$. Для любого $i \in \{1, \ldots, k\}$ алгебраическая кратность значения λ_i равна α_i, при этом α_i базисных векторов из e являются собственными векторами со значением λ_i, откуда $\dim V_{\lambda_i} \geqslant \alpha_i$, и обратное неравенство тоже верно.

- ▷ $(2 \Rightarrow 3)$ Пусть $V_{\lambda_1}, \ldots, V_{\lambda_k} \leqslant V$ — собственные подпространства оператора φ. Их сумма — прямая, и по условию $\sum_{i=1}^k \dim V_{\lambda_i} = \sum_{i=1}^k \alpha_i = n$, поэтому $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k} = V$.

- ▷ $(3 \Rightarrow 4)$ Выберем базисы $e_1, \ldots, e_k$ в пространствах $V_{\lambda_1}, \ldots, V_{\lambda_k}$. Тогда, так как сумма $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$ — прямая, то объединение этих базисов дает базис в V, который и является искомым.

- ▷ $(4 \Rightarrow 1)$ Если e — базис из собственных векторов, то именно в этом базисе матрица оператора φ имеет требуемый диагональный вид. $\square$

Замечание. Рассмотрим пространство V_2 над $\mathbb{R}$ и $\varphi \in \mathcal{L}(V)$ — поворот на угол $\alpha \in (0, \pi)$. Тогда ни один ненулевой вектор из V_2 не переходит в коллинеарный себе под действием φ, поэтому φ нет собственных значений.

2.3 Теорема Гамильтона-Кэли

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$. Отметим также, что в следующих разделах на рассматриваемый оператор $\varphi \in \mathcal{L}(V)$ часто будет налагаться требование, что χ_φ раскладывается в произведение линейных сомножителей, то есть имеет следующий вид при некоторых $\lambda_1, \ldots, \lambda_k \in F$ и $\alpha_1, \ldots, \alpha_k \in \mathbb{N}$ таких, что $\sum_{i=1}^{k} \alpha_i = n$:

$$\chi_\varphi(\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{\alpha_i} \qquad (\star)$$

Определение 2.11. Введем следующие обозначения:

- ▷ Для произвольных $\varphi \in \mathcal{L}(V)$ и $\mu \in F$ положим $\varphi_\mu := \varphi - \mu$
- ▷ Для произвольных $A \in M_n(F)$ и $\mu \in F$ положим $A_\mu := A - \mu E$

Утверждение 2.8. *Пусть $\varphi \in \mathcal{L}(V)$. Тогда $U \leqslant V$ является инвариантным относительно $\varphi \Leftrightarrow U$ является инвариантным относительно φ_μ.*

Доказательство.

- $\Rightarrow$ Если U инвариантно относительно φ, то $\varphi_\mu(U) \leqslant \varphi(U) + \mu U \leqslant U$
- $\Leftarrow$ Если U инвариантно относительно φ_μ, то $\varphi(U) = (\varphi_\mu + \mu)(U) \leqslant \varphi_\mu(U) + \mu U \leqslant U$ □

Утверждение 2.9. *Пусть оператор $\varphi \in \mathcal{L}(V)$ имеет собственное значение. Тогда существует инвариантное относительно φ подпространство $U \leqslant V$ размерности $n - 1$.*

Доказательство. Пусть $\mu \in F$ — собственное значение оператора φ. Тогда φ_μ — вырожденный оператор, то есть $\dim \operatorname{Im} \varphi_\mu \leqslant n-1$. Дополним базис в $\operatorname{Im} \varphi_\mu$ до базиса в некотором подпространстве $U \leqslant V$ размерности $n - 1$, тогда U инвариантно относительно φ_μ и, следовательно, относительно φ. □

Теорема 2.6. *Пусть $\varphi \in \mathcal{L}(V)$, и χ_φ имеет вид $(\star)$. Тогда в V существует такой базис $e = (\overline{e_1}, \ldots, \overline{e_n})$, что для любого $i \in \{1, \ldots, n\}$ подпространство $\langle \overline{e_1}, \ldots, \overline{e_i} \rangle \leqslant V$ инвариантно относительно φ.*

Доказательство. Проведем индукцию по n. База, $n = 1$, тривиальна, докажем переход, $n > 1$. Многочлен χ_φ имеет корни, поэтому существует инвариантное относительно φ подпространство $U \leqslant V$ размерности $n - 1$. Положим $\psi := \varphi|_U \in \mathcal{L}(U)$, тогда $\chi_\psi \mid \chi_\varphi$, поэтому χ_ψ также имеет вид $(\star)$, и к нему применимо предположение индукции. Выберем подходящий базис e' в U и дополним его до базиса в V, получим требуемое. □

Следствие. *Пусть $\varphi \in \mathcal{L}(V)$, и χ_φ имеет вид $(\star)$. Тогда в V существует базис e, в котором матрица преобразования φ имеет верхнетреугольный вид.*

Доказательство. В базисе из теоремы выше матрица оператора φ имеет верхнетреугольный вид, что и требовалось. □

Замечание. В базисе e из утверждения выше матрица оператора φ не только имеет верхнетреугольный вид, но и значения на ее диагонали в точности совпадают с набором корней многочлена χ_φ с учетом кратности.

Теорема 2.7 (Гамильтона-Кэли)**.** *Для любого оператора $\varphi \in \mathcal{L}(V)$ выполнено следующее равенство:*

$$\chi_\varphi(\varphi) = 0$$

Доказательство для случая, когда χ_φ имеет вид $(\star)$. Выберем базис e в V, в котором для любого $i \in \{1, \ldots, n\}$ подпространство $V_i := \langle \overline{e_1}, \ldots, \overline{e_i} \rangle \leqslant V$ инвариантно относительно φ, тогда матрица оператора φ в этом базисе имеет верхнетреугольный вид. Пусть $\varphi \leftrightarrow_e \operatorname{diag}(\lambda_1, \ldots, \lambda_n) \in M_n(F)$. Покажем, что тогда $\varphi_{\lambda_i}(V_i) \leqslant V_{i-1}$ для любого $i \in \{1, \ldots, n\}$. Действительно, так как V_i инвариантно относительно φ, то оно также инвариантно относительно φ_{λ_i}, и матрица сужения $\psi := \varphi_{\lambda_i}|_{V_i} \in \mathcal{L}(V_i)$ имеет следующий вид:

$$\psi \xleftrightarrow[(\overline{e_1}, \ldots, \overline{e_i})]{} \begin{pmatrix} \lambda_1 - \lambda_i & * & \ldots & * \\ 0 & \lambda_2 - \lambda_i & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

Последняя координата у образов всех базисных векторов нулевая, поэтому $\psi(V_i) \leqslant V_{i-1}$. Следовательно, выполнены следующие включения:

$$\chi_\varphi(\varphi)(V) = (\varphi_{\lambda_1} \ldots \varphi_{\lambda_n})(V) \leqslant (\varphi_{\lambda_1} \ldots \varphi_{\lambda_{n-1}})(V_{n-1}) \leqslant \varphi_{\lambda_1}(V_1) \leqslant V_0 = \{\overline{0}\}$$

Таким образом, $\chi_\varphi(\varphi) = 0$, что и требовалось. □

Замечание. Теорема Гамильтона-Кэли справедлива и в общем случае. Это можно доказать, воспользовавшись фактом, который будет доказан позднее: если F — поле и $P \in F[x]$, то существует надполе $K \supset F$ такое, что P раскладывается на линейные сомножители над K. Тогда $A \in M_n(K)$, и в K теорема верна, но поскольку все действия в вычислении $\chi_A(A)$ происходят в F, то и в F теорема верна.

2.4 Аннулирующие многочлены

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$.

Определение 2.12. Пусть $\varphi \in \mathcal{L}(V)$. Многочлен $P \in F[x] \backslash \{0\}$ называется *аннулирующим многочленом* оператора φ, если $P(\varphi) = 0$. *Минимальным многочленом* оператора φ называется аннулирующий многочлен μ_φ наименьшей степени.

Замечание. Из теоремы Гамильтона-Кэли следует, что многочлен χ_φ — аннулирующий для φ. Но и без этой теоремы можно установить существование аннулирующего многочлена у произвольного оператора φ: система $(1, \varphi, \ldots, \varphi^{n^2})$ линейно зависима в $\mathcal{L}(V)$, поскольку $\dim \mathcal{L}(V) = n^2$, значит, у нее есть нетривиальная линейная комбинация, равная нулю, которая и является искомым многочленом.

Утверждение 2.10. *Пусть $\varphi \in \mathcal{L}(V)$, μ_φ — минимальный многочлен для φ. Тогда многочлен $P \in F[x]$ — аннулирующий для $\varphi \Leftrightarrow \mu_\varphi \mid P$.*

Доказательство. Разделим P на μ_φ с остатком, то есть выберем $Q, R \in F[x]$ такие, что $P = Q\mu_\varphi + R$ и $\deg R < \deg \mu_\varphi$, тогда $P(\varphi) = R(\varphi)$. Тогда, поскольку μ_φ — минимальный, выполнены следующие равносильности:

$$P(\varphi) = 0 \Leftrightarrow R(\varphi) = 0 \Leftrightarrow R = 0 \Leftrightarrow \mu_\varphi \mid P$$ □

Следствие. *Минимальный многочлен оператора $\varphi \in \mathcal{L}(V)$ единственен с точностью до ассоциированности.*

Доказательство. Пусть $\mu_1, \mu_2 \in F[x]$ — различные минимальные многочлены для φ. Тогда, поскольку оба они ненулевые, выполнено следующее:

$$\begin{cases} \mu_1 \mid \mu_2 \\ \mu_2 \mid \mu_1 \end{cases} \Rightarrow \begin{cases} \deg \mu_1 \leqslant \deg \mu_2 \\ \deg \mu_2 \leqslant \deg \mu_1 \end{cases} \Rightarrow \deg \mu_1 = \deg \mu_2$$

Таким образом, $\mu_1 \mid \mu_2$ и $\deg \mu_1 = \deg \mu_2$, откуда $\mu_2 = \alpha \mu_1$ для некоторого $\alpha \in F^*$. □

Замечание. Из теоремы Гамильтона-Кэли и утверждения выше следует, что $\mu_\varphi \mid \chi_\varphi$. В частности, если многочлен χ_φ имеет вид ($\star$), то многочлен μ_φ имеет следующий вид при некоторых $\beta_1, \ldots, \beta_k \in \mathbb{N} \cup \{0\}$ таких, что $\beta_i \leqslant \alpha_i$ для каждого $i \in \{1, \ldots, n\}$:

$$\mu_\varphi(\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{\beta_i}$$

Утверждение 2.11. *Пусть $\varphi \in \mathcal{L}(V)$, $\lambda_0 \in F$ — собственное значение оператора φ. Тогда $(\lambda - \lambda_0) \mid \mu_\varphi$.*

Доказательство. Достаточно показать, что λ_0 — корень многочлена μ_φ. Рассмотрим собственный вектор $\overline{v} \in V \backslash \{\overline{0}\}$ оператора φ со значением λ_0, тогда для любого $k \in \mathbb{N}$ выполнено равенство $\varphi^k(\overline{v}) = \lambda_0^k \overline{v}$. В частности, для многочлена μ_φ выполнены следующие равенства:

$$\mu_\varphi(\varphi)(\overline{v}) = \mu_\varphi(\lambda_0)\overline{v} = \overline{0}$$

Но вектор $\overline{v}$ — ненулевой, поэтому $\mu_\varphi(\lambda_0) = 0$. □

Замечание. Можно также показать, что любой неприводимый делитель многочлена χ_φ делит многочлен μ_φ.

Теорема 2.8. *Пусть $\varphi \in \mathcal{L}(V)$, $P \in F[x]$ — аннулирующий многочлен оператора φ, и $P = P_1 P_2$ для многочленов $P_1, P_2 \in F[x]$ таких, что НОД$(P_1, P_2) = 1$. Тогда $V = V_1 \oplus V_2$, где $V_1 := \operatorname{Ker} P_1(\varphi), V_2 := \operatorname{Ker} P_2(\varphi)$ — инвариантные относительно φ подпространства.*

Доказательство. Подпространства из условия инвариантны относительно φ, поскольку операторы $P_1(\varphi), P_2(\varphi)$ коммутируют с φ. Покажем, что $\operatorname{Im} P_1(\varphi) \leqslant V_2$. Действительно, $P_2(\varphi)(P_1(\varphi)(V)) = P(\varphi)(V) = \{\overline{0}\}$, то есть $\operatorname{Im} P_1(\varphi) \leqslant \operatorname{Ker} P_2(\varphi) = V_2$. Аналогично, выполнено включение $\operatorname{Im} P_2(\varphi) \leqslant V_1$. Поскольку НОД$(P_1, P_2) = 1$, то существуют многочлены $Q_1, Q_2 \in F[x]$ такие, что выполнено равенство $P_1 Q_1 + P_2 Q_2 = 1$. Подставим φ в это равенство и получим следующее:

$$P_1(\varphi)Q_1(\varphi) + P_2(\varphi)Q_2(\varphi) = \mathrm{id}$$

Значит, для произвольного $\overline{v} \in V$ выполнены следующие равенства:

$$\overline{v} = \mathrm{id}(\overline{v}) = Q_1(\varphi)(P_1(\varphi)(\overline{v})) + Q_2(\varphi)(P_2(\varphi)(\overline{v}))$$

Заметим теперь, что $P_1(\varphi)(\overline{v}) \in \operatorname{Im} P_1(\varphi) \leqslant V_2$ и $P_2(\varphi)(\overline{v}) \in \operatorname{Im} P_2(\varphi) \leqslant V_1$, откуда $Q_1(\varphi)(P_1(\varphi)(\overline{v})) \in V_2$ и $Q_2(\varphi)(P_2(\varphi)(\overline{v})) \in V_1$ в силу инвариантности подпространств V_1, V_2

относительно φ. Значит, $V_1 + V_2 = V$, причем эта сумма — прямая, поскольку для любого вектора $\overline{w} \in V_1 \cap V_2$ выполнены следующие равенства:

$$\overline{w} = \mathrm{id}(\overline{w}) = Q_1(\varphi)(P_1(\varphi)(\overline{w})) + Q_2(\varphi)(P_2(\varphi)(\overline{w})) = Q_1(\varphi)(\overline{0}) + Q_2(\varphi)(\overline{0}) = \overline{0}$$

Таким образом, $V = V_1 \oplus V_2$. □

Замечание. На самом деле, в теореме выше выполнены равенства $\operatorname{Im} P_1(\varphi) = V_2$ и $\operatorname{Im} P_2(\varphi) = V_1$. Согласно теореме, $V_1 \oplus V_2 = V$, поэтому выполнено следующее:

$$\dim V_1 + \dim V_2 = \dim V = \dim \operatorname{Im} P_1(\varphi) + \dim \operatorname{Ker} P_1(\varphi)$$

Следовательно, $\dim V_2 = \dim \operatorname{Im} P_1(\varphi)$, и, в силу включения $\operatorname{Im} P_1(\varphi) \leqslant V_2$, имеем $\operatorname{Im} P_1(\varphi) = V_2$. Аналогичное рассуждение показывает, что $\operatorname{Im} P_2(\varphi) = V_1$.

Следствие. *Пусть $\varphi \in \mathcal{L}(V)$, $P \in F[x]$ — аннулирующий многочлен оператора φ, и $P = P_1 \cdots P_n$ для попарно взаимно простых многочленов $P_1, \ldots, P_n \in F[x]$. Тогда выполнено равенство $V = V_1 \oplus \cdots \oplus V_n$, где $V_i := \operatorname{Ker} P_i(\varphi)$ — инвариантное относительно φ подпространство для каждого $i \in \{1, \ldots, n\}$.*

Доказательство. Проведем индукцию по n. База, $n = 2$, уже доказан, докажем переход. Пусть $n > 2$, тогда $P = P_1 \ldots P_n = (P_1 \ldots P_{n-1})P_n$, причем многочлены $(P_1 \ldots P_{n-1})$ и P_n взаимно просты, тогда выполнено следующее:

$$V = \operatorname{Ker}(P_1 \ldots P_{n-1})(\varphi) \oplus \operatorname{Ker} P_n(\varphi)$$

Положим $\widetilde{V} := \operatorname{Ker}(P_1 \ldots P_{n-1})(\varphi)$, $V_n := \operatorname{Ker} P_n(\varphi)$, и применим предположение индукции к оператору $\psi := \varphi|_{\widetilde{V}} \in \mathcal{L}(\widetilde{V})$ и $P_1 \ldots P_{n-1}$, тогда:

$$\widetilde{V} = \operatorname{Ker} P_1(\psi) \oplus \cdots \oplus \operatorname{Ker} P_{n-1}(\psi)$$

Остается заметить, что для каждого индекса $i \in \{1, \ldots, n-1\}$ выполнено включение $V_i \leqslant \operatorname{Ker}(P_1 \ldots P_{n-1})(\varphi) = \widetilde{V}$, откуда $\operatorname{Ker} P_i(\psi) = V_i \cap \widetilde{V} = V_i$, поэтому $V_i' = V_i$, Таким образом, $V = V_1 \oplus \cdots \oplus V_n$. □

3 Жорданова нормальная форма и ее приложения

3.1 Жорданова нормальная форма

Определение 3.1. Пусть F — поле, $\lambda_0 \in F$. *Жордановой клеткой* размера $k \in \mathbb{N}$ с собственным значением λ_0 называется матрица $J_k(\lambda_0) \in M_k(F)$, имеющая следующий вид:

$$J_k(\lambda_0) := \begin{pmatrix} \lambda_0 & 1 & \ldots & 0 & 0 \\ 0 & \lambda_0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \lambda_0 & 1 \\ 0 & 0 & \ldots & 0 & \lambda_0 \end{pmatrix}$$

Считается также, что $J_k := J_k(0)$.

Определение 3.2. Пусть F — поле, $A \in M_n(F)$. Матрица A имеет *жорданов вид*, если она имеет блочно-диагональный вид, в котором каждый блок является жордановой клеткой, то есть имеет следующий вид для некоторых $k_1, \ldots, k_m \in \mathbb{N}$ и $\lambda_1, \ldots, \lambda_m \in F$:

$$A = \begin{pmatrix} \boxed{J_{k_1}(\lambda_1)} & 0 & \ldots & 0 \\ 0 & \boxed{J_{k_2}(\lambda_2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \boxed{J_{k_m}(\lambda_m)} \end{pmatrix}$$

Определение 3.3. Пусть V — линейное пространство, $\varphi \in \mathcal{L}(V)$. Если в некотором базисе матрица оператора φ имеет жорданов вид, то она называется *жордановой нормальной формой* оператора φ, а соответствующий базис — его *жордановым базисом.*

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$, а также зафиксируем оператор $\varphi \in \mathcal{L}(V)$ такой, что $\chi_\varphi(\lambda)$ имеет вид $(\star)$.

Замечание. В силу предположения выше, многочлен $\chi_\varphi(\lambda)$ можно представить в виде произведения многочленов $P_1 := (\lambda_1 - \lambda)^{\alpha_1}, \ldots, P_k := (\lambda_k - \lambda)^{\alpha_k}$. Эти многочлены попарно взаимно просты.

Определение 3.4. Пусть λ_i — собственное значение оператора φ. *Корневым подпространством*, соответствующим λ_i, называется $V^{\lambda_i} := \operatorname{Ker} P_i(\varphi) = \operatorname{Ker}(\varphi_{\lambda_i})^{\alpha_i}$.

Замечание. Поскольку $V_{\lambda_i} = \operatorname{Ker}\varphi_{\lambda_i}$, то $V_{\lambda_i} \leqslant V^{\lambda_i}$, но обратное включение верно не всегда. Кроме того, $V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_k}$ по уже доказанному утверждению.

Утверждение 3.1. *Пусть λ_i, λ_j — различные собственные значения оператора φ. Тогда оператор $\varphi_{\lambda_i}|_{V^{\lambda_j}} \in \mathcal{L}(V^{\lambda_j})$ — невырожденный.*

Доказательство. Отметим сначала, что сужение $\varphi_{\lambda_i}|_{V^{\lambda_j}}$ корректно, поскольку V^{λ_j} инвариантно относительно φ. Тогда, поскольку выполнены равенства $V^{\lambda_i} \cap V^{\lambda_j} = \{\overline{0}\}$ и $V_{\lambda_i} = \operatorname{Ker}\varphi_{\lambda_i} \leqslant V^{\lambda_i}$, имеем $\operatorname{Ker}\varphi_{\lambda_i}|_{V^{\lambda_j}} = V^{\lambda_j} \cap \operatorname{Ker}\varphi_{\lambda_i} = \{\overline{0}\}$, что и означает невырожденность оператора. □

Утверждение 3.2. *Пусть λ_i — собственное значение оператора φ. Тогда выполнено следующее равенство:*

$$V^{\lambda_i} = \{\overline{v} \in V : \exists m \in \mathbb{N} : (\varphi_{\lambda_i})^m(\overline{v}) = \overline{0}\}$$

Доказательство. Нетривиально только включение $(\geqslant)$. Пусть $\overline{v} \in V$ — вектор такой, что $(\varphi_{\lambda_i})^m(\overline{v}) = \overline{0}$ для некоторого $m \in \mathbb{N}$. Представим его в виде $\overline{v} = \overline{v_1} + \cdots + \overline{v_k}$, где $\overline{v_j} \in V^{\lambda_j}$ для каждого $j \in \{1, \ldots, k\}$, тогда выполнено следующее равенство:

$$(\varphi_{\lambda_i})^m(\overline{v_1}) + \cdots + (\varphi_{\lambda_i})^m(\overline{v_k}) = \overline{0}$$

В силу инвариантности, для каждого $j \in \{1, \ldots, k\}$ выполнено $(\varphi_{\lambda_i})^m(\overline{v_j}) \in V^{\lambda_j}$, поэтому каждый такой вектор равен $\overline{0}$ по свойству прямой суммы. Но для каждого индекса $j \in \{1, \ldots n\} \setminus \{i\}$ оператор $\varphi_{\lambda_i}|_{V^{\lambda_j}}$ — невырожденный, откуда $\overline{v_j} = \overline{0}$. Значит, $\overline{v} = \overline{v_i} \in V^{\lambda_i}$, и получено требуемое. □

Замечание. Утверждение выше означает, что если вектор $\overline{v} \in V$ обнуляется под действием какой-либо степени оператора φ_{λ_i}, то он также обнуляется под действием этого оператора в некоторой степени, не превосходящей α_i.

Определение 3.5. Пусть $\psi \in \mathcal{L}(V)$. Оператор ψ называется *нильпотентным*, если существует $m \in \mathbb{N}$ такое, что $\psi^m = 0$.

Замечание. Все собственные значения оператора являются корнями любого аннулирующего многочлена этого оператора, а для нильпотентного оператора ψ многочлен x^m является аннулирующим, поэтому его единственное собственное значение — это 0.

Определение 3.6. Пусть $\psi \in \mathcal{L}(V)$ — нильпотентный оператор. Подпространство $U \leqslant V$ называется *циклическим* относительно ψ, если оно инвариантно относительно ψ и существует вектор $\overline{v} \in V$ такой, что U — минимальное по включению инвариантное подпространство, содержащее $\overline{v}$.

Замечание. Если подпространство $U \leqslant V$ — циклическое относительно нильпотентного оператора $\psi \in \mathcal{L}(V)$ и вектора $\overline{v} \in V$, то $U = \langle \overline{v}, \psi(\overline{v}), \psi^2(\overline{v}), \dots \rangle$, причем порождающий набор конечен в силу нильпотентности оператора ψ.

Определение 3.7. Пусть $\psi \in \mathcal{L}(V)$ — нильпотентный оператор, $\overline{v} \in V \backslash \{\overline{0}\}$. *Высотой* вектора $\overline{v}$ относительно ψ называется наименьшее $n \in \mathbb{N}$ такое, что $\psi^n(\overline{v}) = \overline{0}$.

Утверждение 3.3. *Пусть $\psi \in \mathcal{L}(V)$ — нильпотентный оператор, и пусть векторы $\overline{v_1}, \dots, \overline{v_k} \in V \backslash \{\overline{0}\}$ имеют попарно различные высоты относительно ψ. Тогда эти векторы образуют линейно независимую систему.*

Доказательство. Предположим, что это не так, тогда существует нетривиальная линейная комбинация с коэффициентами $\alpha_1, \dots, \alpha_k \in F$, равная нулю:

$$\alpha_1 \overline{v_1} + \dots + \alpha_k \overline{v_k} = \overline{0}$$

Пусть $\overline{v_i}$ — вектор с наибольшей высотой n_i, коэффициент при котором не равен нулю. Применяя к данному равенству $\psi^{n_i - 1}$, получим, что $\alpha_i \psi^{n_i - 1}(\overline{v_i}) = \overline{0}$. Значит, $\alpha_i = 0$, что противоречит нашему предположению. □

Следствие. *Пусть $\psi \in \mathcal{L}(V)$ — нильпотентный оператор, и пусть $U \leqslant V$ — циклическое подпространство, порожденное вектором $\overline{v} \in V \backslash \{\overline{0}\}$ высоты n. Тогда система $(\overline{v}, \psi(\overline{v}), \dots, \psi^{n-1}(\overline{v}))$ образует базис в U.*

Доказательство. Как уже было отмечено, $U = \langle \overline{v}, \psi(\overline{v}), \psi^2(\overline{v}), \dots \rangle$, тогда, поскольку высота веткора $\overline{v}$ равна n, имеем $U = \langle \overline{v}, \psi(\overline{v}), \dots, \psi^{n-1}(\overline{v}) \rangle$. Кроме того, система $(\overline{v}, \psi(\overline{v}), \dots, \psi^{n-1}(\overline{v}))$ линейно независима по утверждению выше. □

Замечание. Матрица оператора ψ в базисе $(\psi^{n-1}(\overline{v}), \dots, \psi(\overline{v}), \overline{v})$ имеет вид жордановой клетки $J_n \in M_n(F)$.

Утверждение 3.4. *Пусть $\psi \in \mathcal{L}(V)$ — нильпотентный оператор, $\overline{v} \in V \backslash \{\overline{0}\}$ — вектор наибольшей высоты n в пространстве V, и $U \leqslant V$ — циклическое подпространство, порожденное $\overline{v}$. Тогда существует $W \leqslant V$ такое, что W инвариантно относительно ψ и $V = U \oplus V$.*

Доказательство. Отметим сначала, что вектор $\overline{v}$ из условия определен корректно, поскольку все векторы из $V \backslash \{\overline{0}\}$ имеют конечную высоту, ограниченную сверху величиной $\dim V$. Выберем инвариантное подпространство $W \leqslant V$ наибольшей размерности такое, что $W \cap U = \{\overline{0}\}$. Такое подпространство точно существует, потому что по меньшей мере

$\{\overline{0}\} \leqslant V$ удовлетворяет условию. Если $U \oplus W = V$, то утверждение доказано. Если же $U \oplus W \neq V$, то выберем $\overline{x} \notin U \oplus W$. Поскольку в наборе $\overline{x}, \psi(\overline{x}), \psi^2(\overline{x}), \ldots, \overline{0}$ первый вектор не лежит в $U \oplus W$, а последний — лежит, то в некоторый момент происходит «скачок» из-за пределов подпространства $U \oplus W$ в $U \oplus W$. Пусть без ограничения общности это происходит на первом шаге, то есть $\psi(\overline{x}) \in U \oplus W$. По свойству прямой суммы, для некоторых скаляров $\alpha_0, \ldots, \alpha_{n-1} \in F$ и вектора $w \in W$ выполнено следующее равенство:

$$\psi(\overline{x}) = \alpha_0 \overline{v} + \cdots + \alpha_{n-1}\psi^{n-1}(\overline{v}) + \overline{w}$$

Поскольку n — наибольшая высота в V, то $\overline{0} = \psi^n(\overline{x}) = \psi^{n-1}(\psi(\overline{x}))$. Тогда, применив оператор ψ^{n-1} к обеим частям равенства, получим следующее:

$$\overline{0} = \alpha_0 \psi^{n-1}(\overline{v}) + \psi^{n-1}(\overline{w})$$

Поскольку сумма $U \oplus W$ — прямая, то оба слагаемых в правой части равенства равны нулю, то есть $\alpha_0 = 0$ и $\psi^{n-1}(\overline{w}) = \overline{0}$. Положим теперь $\overline{x'} := \overline{x} - \alpha_1\overline{v} - \cdots - \alpha_{n-1}\psi^{n-1}(\overline{v})$ и заметим, что $\overline{x'} \notin U \oplus W$, поскольку $\overline{x} \notin U \oplus W$. Кроме того, выполнено следующее:

$$\psi(\overline{x'}) = \psi(\overline{x}) - \alpha_1\overline{v_1} - \cdots - \alpha_{n-1}\overline{v_{n-1}} = \overline{w}$$

Значит, пространство $W \oplus \langle \overline{x'} \rangle$ — тоже инвариантное, причем $(W \oplus \langle \overline{x'} \rangle) \cap U = \{\overline{0}\}$. Получено противоречие с максимальностью размерности подпространства W. $\square$

Следствие. *Пусть $\psi \in \mathcal{L}(V)$ — нильпотентный оператор. Тогда V раскладывается в прямую сумму циклических относительно ψ подпространств.*

Доказательство. Проведем индукцию по размерности пространства V. База, $\dim V = 0$, тривиальна, докажем переход. Выберем в V вектор $\overline{v}$ наибольшей высоты и порожденное им циклическое подпространство U, а также $W \leqslant V$ такое, что W инвариантно относительно ψ и $V = U \oplus W$. Для подпространства W и оператора $\psi|_W \in \mathcal{L}(W)$ применимо предположение индукции. $\square$

Замечание. Размерности циклических подпространств, построенных в доказательстве выше, образуют невозрастающую последовательность.

Теорема 3.1 (о существовании жордановой нормальной формы)**.** *Пусть $\varphi \in \mathcal{L}(V)$, и χ_φ имеет вид $(\star)$. Тогда у оператора φ есть жорданова нормальная форма.*

Доказательство. Представим V в виде прямой суммы корневых подпространств:

$$V = V^{\lambda_1} \oplus \cdots \oplus V^{\lambda_k}$$

Для любого $i \in \{1, \ldots, k\}$ оператор $\varphi_{\lambda_i}|_{V^{\lambda_i}} \in \mathcal{L}(V^{\lambda_i})$ — нильпотентный, поэтому он раскладывается в прямую сумму циклических подпространств и, как следствие, имеет жорданов базис e_i. В этом базисе оператор $\varphi_{\lambda_i}|_{V^{\lambda_i}}$ имеет жорданову нормальную форму с нулями на главной диагонали, то есть для некоторых $k_1, \ldots, k_m \in \mathbb{N}$ выполнено следующее:

$$\varphi_{\lambda_i}|_{V^{\lambda_i}} \leftrightarrow_{e_i} \begin{pmatrix} \boxed{J_{k_1}} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \boxed{J_{k_m}} \end{pmatrix}$$

В этом же базисе e_i оператор $\varphi|_{V^{\lambda_i}} \in \mathcal{L}(V^{\lambda_i})$ имеет жорданову нормальную форму следующего вида:

$$\varphi|_{V^{\lambda_i}} \leftrightarrow_{e_i} \begin{pmatrix} \boxed{J_{k_1}(\lambda_i)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \boxed{J_{k_m}(\lambda_i)} \end{pmatrix}$$

Объединение жордановых базисов в подпространствах $V^{\lambda_1}, \dots, V^{\lambda_k}$ дает искомый жорданов базис в V. □

Замечание. Существует и более конструктивный подход к получению жорданова базиса для нильпотентного оператора. Опишем его ниже.

Пусть $\psi \in \mathcal{L}(V)$ — нильпотентный, k — наибольшая высота вектора в V относительно ψ. Для каждого $i \in \{1, \dots, k\}$ положим $V_i := \operatorname{Ker} \psi^i$ — пространство векторов высоты, не превосходящей i. Выберем U_k — прямое дополнение подпространства V_{k-1} в V_k, тогда все ненулевые векторы в U_k имеют высоту k, поэтому $\psi(U_k) \leqslant V_{k-1}$, причем все ненулевые векторы в $\psi(U_k)$ имеют высоту $k-1$, откуда $\psi(U_k) \cap V_{k-2} = \{\overline{0}\}$. Значит, можно также выбрать U_{k-1} — такое прямое дополнение подпространства V_{k-2} в V_{k-1}, что $\psi(U_k) \leqslant U_{k-1}$. Продолжая процесс, получим $U_k, \dots, U_1 \leqslant V$ такие, что для каждого $i \in \{1, \dots, k-1\}$ выполнено $\psi(U_{i+1}) \leqslant U_i$. Для каждого $i \in \{1, \dots, k-1\}$ также выберем W_i — прямое дополнение подпространства $\psi(U_{i+1})$ в U_i. Тогда пространство V примет следующий вид:

$$\begin{array}{r|c|c|c|c|}
\cline{2-2}
U_k \{ & U_k & \multicolumn{3}{c}{} \\
\cline{2-3}
U_{k-1} \{ & \psi(U_k) & W_{k-1} & \multicolumn{2}{c}{} \\
\cline{2-4}
U_{k-2} \{ & \multicolumn{2}{c|}{\psi(U_{k-1})} & W_{k-2} & \multicolumn{1}{c}{} \\
\cline{2-4}
\multicolumn{1}{r}{} & \multicolumn{1}{c}{\vdots} & \multicolumn{1}{c}{\vdots} & \multicolumn{1}{c}{\vdots} & \multicolumn{1}{c}{\ddots} \\
\cline{2-5}
U_1 \{ & \multicolumn{3}{c|}{\psi(U_2)} & W_1 \\
\cline{2-5}
\end{array}$$

Заметим теперь, что для любого $i \in \{1, \dots, k-1\}$ линейно независимая система $(\overline{v_1}, \dots, \overline{v_t})$ векторов из U_{i+1} под действием ψ переходит в линейно независимую систему $(\psi(\overline{v_1}), \dots, \psi(\overline{v_t}))$ векторов из U_i. Действительно, любая нетривиальная линейная комбинация системы имеет высоту i и потому не обращается в ноль под действием ψ. Значит, если на каждой «ступеньке» U_i выбрать базис e_i, то образ этого базиса $\psi(e_i)$ будет базисом в $\psi(U_i)$, который можно будет дополнить до базиса e_{i-1} в U_{i-1}. Тогда система $e_k \cup \dots \cup e_1$ будет искомым жордановым базисом в V, а каждая вертикальная «цепочка» вида $\overline{v}, \psi(\overline{v}), \dots$ будет порождать очередное циклическое подпространство C, и сумма таких циклических подпространств будет прямой и равной V.

Замечание. Диагональный вид матрицы также является жордановым видом: каждый элемент главной диагонали — это жорданова клетка размера 1.

Теорема 3.2 (о единственности жордановой нормальной формы)**.** *Пусть $\varphi \in \mathcal{L}(V)$, и χ_φ имеет вид $(\star)$. Тогда жорданова нормальная форма оператора φ единственна с точностью до перестановки клеток.*

Доказательство. Пусть $\lambda_0 \in F$ — собственное значение оператора φ. Выберем жорданов базис $e = (\overline{e_1}, \dots, \overline{e_n})$ такой, что $\varphi \leftrightarrow_e A \in M_n(F)$, где A — жорданова нормальная форма, в которой все клетки со значением λ_0 стоят в начале и имеют суммарный размер $d \leqslant n$, тогда этим клеткам соответствует начальный фрагмент базиса $(\overline{e_1}, \dots, \overline{e_d})$. Обозначим размеры этих клеток через $k_1, \dots, k_s$, тогда $\sum_{j=1}^{s} k_j = d$. Достаточно показать, что набор $\{k_1, \ldots, k_s\}$ определен однозначно, поскольку для клеток с другими собственными значениями рассуждение будет аналогичным.

Пусть α_0 — алгебраическая кратность значения λ_0, тогда выполнено равенство $d = \alpha_0$. Рассмотрим оператор φ_{λ_0} и заметим, что $\varphi_{\lambda_0} \leftrightarrow_e A_{\lambda_0} = A - \lambda_0 E$, то есть первые d элементов на главной диагонали A_{λ_0} равны нулю, а остальные — отличны от нуля. При возведении матрицы A_{λ_0} в некоторую степень каждая клетка возводится в степень независимо, причем ранг вырожденной клетки в каждой следующей степени уменьшается на один, пока клетка не станет нулевой, а невырожденные клетки остаются невырожденными. Значит, $\operatorname{rk}(A_{\lambda_0})^d = n - d$, и выполнены следующие равенства:

$$\dim V^{\lambda_0} = \dim \operatorname{Ker}(\varphi_{\lambda_0})^d = n - \operatorname{rk}(A_{\lambda_0})^d = d$$

Кроме того, поскольку V^{λ_0} — это пространство всех векторов, обнуляемых оператором $(\varphi_{\lambda_0})^d$, то $\langle \overline{e_1}, \ldots, \overline{e_d} \rangle = V^{\lambda_0}$. Исследуем нильпотентный оператор $\psi := \varphi_{\lambda_0}|_{V^{\lambda_0}} \in \mathcal{L}(V^{\lambda_0})$. Его матрица в базисе $e' := (\overline{e_1}, \ldots, \overline{e_d})$ имеет следующий вид:

$$\psi \leftrightarrow_{e'} B := \begin{pmatrix} \boxed{J_{k_1}} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \boxed{J_{k_s}} \end{pmatrix}$$

Пусть n_1 — число клеток размера $\geqslant 1$, n_2 — число клеток размера $\geqslant 2$, и так далее. Число клеток размера $j \in \{1, \ldots, d\}$ равно $n_j - n_{j+1}$, поэтому для определения числа клеток каждого размера достаточно найти все числа $n_1, \ldots, n_d$. Для каждого $i \in \{1, \ldots, d\}$ положим $V_i := \operatorname{Ker} \psi^i$, тогда $V_{\lambda_0} = V_1 \leqslant V_2 \leqslant \ldots \leqslant V_d = V^{\lambda_0}$. Чтобы определить величины $\dim V_1, \ldots, \dim V_d$, снова воспользуемся замечанием о том, что возведение клетки J_k в каждую следующую степень уменьшает ее ранг на один, пока клетка не станет нулевой:

$$\begin{aligned}
&\dim V_1 = \dim \operatorname{Ker} \psi = d - \operatorname{rk} B = n_1 \\
&\dim V_2 = \dim \operatorname{Ker} \psi^2 = (d - \operatorname{rk} B) + (\operatorname{rk} B - \operatorname{rk} B^2) = n_1 + n_2 \\
&\dim V_3 = \dim \operatorname{Ker} \psi^3 = (d - \operatorname{rk} B^2) + (\operatorname{rk} B^2 - \operatorname{rk} B^3) = n_1 + n_2 + n_3 \\
&\qquad\ldots \\
&\dim V_d = \dim \operatorname{Ker} \psi^d = (d - \operatorname{rk} B^{d-1}) + (\operatorname{rk} B^{d-1} - \operatorname{rk} B^d) = \sum_{j=1}^{d} n_j
\end{aligned}$$

Таким образом, числа $n_1, \ldots, n_d$ выражаются через величины $\dim V_1, \ldots, \dim V_d$ вне зависимости от выбора базиса, и по ним однозначно определяется набор $\{k_1, \ldots, k_s\}$. Таким образом, жорданова нормальная форма оператора φ определена однозначно с точностью до перестановки клеток задается свойствами оператора, не зависящими от выбора базиса, что и означает ее единственность. $\square$

3.2 Простейшие приложения жордановой нормальной формы

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$.

Замечание. Пусть $A, B \in M_n(F)$ — матрицы, характеристические многочлены которых имеют вид $(\star)$. Тогда матрицы A, B подобны $\Leftrightarrow$ их жордановы формы A' и B' совпадают с точностью до перестановки клеток. Это верно потому, что жорданова нормальная форма оператора, а значит и задающей его матрицы, единственна с точностью до перестановки клеток.

Утверждение 3.5. *Пусть $\varphi \in \mathcal{L}(V)$, и χ_φ имеет вид $(\star)$. Тогда его минимальный многочлен имеет вид $\mu_\varphi = \prod_{i=1}^{k}(\lambda_i - \lambda)^{\beta_i}$, где β_i — наибольший размер клетки с собственным значением λ_i в жордановой нормальной форме оператора φ для каждого $i \in \{1, \dots, k\}$.*

Доказательство. Уже было доказано, что многочлен μ_φ делит многочлен χ_φ. Заметим, что наблюдение о независимом возведении жордановых клеток в степень можно обобщить на случай произвольного многочлена $p \in F[x]$:

$$p\begin{pmatrix} \boxed{J_{k_1}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \boxed{J_{k_s}} \end{pmatrix} = \begin{pmatrix} \boxed{p(J_{k_1})} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \boxed{p(J_{k_s})} \end{pmatrix}$$

Значит, многочлен p — аннулирующий $\Leftrightarrow$ p обнуляет каждую клетку жордановой нормальной формы. Пусть s_j — наибольший размер клетки, соответствующий значению λ_j в жордановой нормальной форме оператора φ, тогда:

$$0 = \mu_\varphi(J_{s_j}(\lambda_j)) = \prod_{i=1}^{k}(\lambda_i E - J_{s_j}(\lambda_j))^{\beta_i} = (-J_{s_j})^{\beta_j}\prod_{i \neq j}(\lambda_i E - J_{s_j}(\lambda_j))^{\beta_i}$$

Поскольку все матрицы в произведении кроме первой — невырожденные, то их произведение — тоже невырожденная матрица, поэтому выполнено равенство $(J_{s_j})^{\beta_j} = 0$, откуда $\beta_j \geqslant s_j$. С другой стороны, степени $\beta_j = s_j$ достаточно, чтобы обнулить все клетки, соответствующие значению λ_j. □

Утверждение 3.6. *Если минимальный многочлен μ_φ оператора $\varphi \in \mathcal{L}(V)$ раскладывается на линейные сомножители, то и χ_φ раскладывается на линейные сомножители, то есть имеет вид $(\star)$.*

Доказательство. Минимальный многочлен, конечно, является аннулирующим, причем $\mu_\varphi = \prod_{i=1}^{k}(\lambda_i - \lambda)^{\gamma_i}$, поэтому, как уже было доказано, $V = V_1 \oplus \dots \oplus V_k$, где $V_i = \mathrm{Ker}\,(\varphi_{\lambda_i})^{\gamma_i}$ для каждого $i \in \{1, \dots, k\}$. Тогда справедливы дальнейшие рассуждения из доказательства существования жордановой нормальной формы оператора φ. Но если оператор φ имеет жорданову нормальную форму, то многочлен χ_φ раскладывается на линейные сомножители. □

Замечание. Жорданова нормальная форма позволяет быстро возводить в степень матрицы $A \in M_n(F)$ такие, что χ_A имеет вид $(\star)$. Действительно, если матрица A имеет жорданову нормальную форму B, то для некоторой матрицы $S \in \mathrm{GL}_n(F)$ выполнено $A = S^{-1}BS$, откуда $A^n = S^{-1}B^nS$. Поскольку жордановы клетки возводятся в степень независимо, достаточно уметь находить степень каждой из них. Для этого заметим, что для любой клетки $J_k(\lambda_0)$ и любого $m \in \mathbb{N}$ выполнено следующее:

$$(J_k(\lambda_0))^m = (\lambda_0 E + J_k)^m = \sum_{i=0}^{m} C_m^i \lambda_0^{m-i} J_k^i = \sum_{i=0}^{k-1} C_m^i \lambda_0^{m-i} J_k^i$$

Формула бинома Ньютона для матриц $\lambda_0 E$ и J_k справедлива потому, что эти матрицы коммутируют.

Замечание. Можно показать, что верно более сильное утверждение: если $\operatorname{char} F = 0$, то для любого многочлена $p \in F[x]$ выполнено равенство $p(A) = S^{-1}p(B)S$, причем для каждой жордановой клетки $J_k(\lambda)$ в B справедлива следующая формула:

$$p(J_k(\lambda)) = p(\lambda)E + p'(\lambda)J_k + \frac{p''(\lambda)}{2}J_k^2 + \dots + \frac{p^{(k-1)}(\lambda)}{(k-1)!}J_k^{k-1}$$

3.3 Линейные рекурренты

До конца раздела зафиксируем поле F.

Определение 3.8. Пусть F — поле. Обозначим через F^∞ *линейное пространство последовательностей* вида $(a_0, a_1, \dots)$, с элементами из F, с покоординатными операциями сложения и умножения на скаляр из F.

Определение 3.9. Последовательность $(a_i) \in F^\infty$ называется *линейной рекуррентой* с *характеристическим многочленом* $p = x^n + p_{n-1}x^{n-1} + \dots + p_0 \in F[x]$, если выполнено следующее условие:

$$\forall k \in \mathbb{N} \cup \{0\} : a_{k+n} + p_{n-1}a_{k+n-1} + \dots + p_0 a_k = 0 \qquad (\dagger)$$

Обозначим множество рекуррент с характеристическим многочленом p через V_p.

Замечание. Если $p_0 = 0$, то элемент a_0 никак не влияет на все остальные элементы последовательности и может принимать произвольные значения, а остальные элементы образуют линейную рекурренту с характеристическим многочленом $\frac{p}{x} \in F[x]$. Поэтому далее будем считать, что $p_0 \neq 0$.

Утверждение 3.7. *V_p — линейное пространство над F, причем $\dim V_p = \deg p = n$.*

Доказательство. Нетривиальна только вторая часть утверждения. Заметим, что любая последовательность $(a_i) \in V_p$ однозначно задается первыми своими n членами $a_0, a_1, \dots, a_{n-1}$, и рассмотрим в V_p следующие последовательности:

$$\overline{e_0} = (\overset{0}{1}, \overset{1}{0}, \dots, 0, 0, \overset{n}{-p_0}, \dots)$$
$$\dots$$
$$\overline{e_{n-1}} = (0, 0 \dots, 0, 1, -p_{n-1}, \dots)$$

Эти последовательности, очевидно, образуют линейно независимую систему, и любая последовательность $(a_i) \in V_p$ выражается через них как $\sum_{i=0}^{n-1} a_i\overline{e_i}$. ☐

Определение 3.10. *Оператором левого сдвига* на F^∞ называется оператор $\varphi : F^\infty \to F^\infty$, заданный для каждой последовательности $(a_i) \in F^\infty$ как $\varphi((a_i)) := (a_{i+1})$.

Утверждение 3.8. *Пусть $p \in F[x]$, φ — оператор левого сдвига. Тогда $V_p = \operatorname{Ker} p(\varphi)$.*

Доказательство. Для любой последовательности $A = (a_i) \in F^\infty$ выполнена равносильность $A \in \operatorname{Ker} p(\varphi) \Leftrightarrow p(\varphi)(A) = (0)$. Заметим, что для любого $k \in \mathbb{N} \cup \{0\}$ выполнено равенство $[p(\varphi)(A)]_k = a_{k+n} + a_{k+n-1}p_{n-1} + \dots + a_k p_0$, поэтому $A \in \operatorname{Ker} p(\varphi) \Leftrightarrow A$ удовлетворяет $(\dagger)$. ☐

Замечание. В частности, из доказанного следует, что V_p инвариантно относительно φ. С этого момента рассматривать оператор $\psi := \varphi|_{V_p} \in \mathcal{L}(V_p)$ для произвольного $p \in F[x]$.

Утверждение 3.9. *Минимальный многочлен μ_ψ оператора ψ равен p с точностью до ассоциированности.*

Доказательство. С одной стороны, $V_p = \operatorname{Ker} p(\varphi)$, поэтому $\operatorname{Im} p(\psi) = p(\varphi)(V_p) = \{\overline{0}\}$, значит, $\mu_\psi \mid p$. С другой стороны, $\forall A \in V_p : \mu_\psi(\psi)(A) = (0) \Leftrightarrow \mu_\psi(\varphi)(A) = (0)$, тогда $A \in \operatorname{Ker} \mu_\psi(\psi) \Leftrightarrow A \in V_{\mu_\psi}$. Значит, $V_p \leqslant V_{\mu_\psi}$, откуда $\deg p = \dim V_p \leqslant \dim V_{\mu_\psi} = \deg \mu_\psi$. Таким образом, $p = \mu_\psi$ с точностью до ассоциированности. □

Утверждение 3.10. *Характеристический многочлен χ_ψ оператора ψ равен $(-1)^n p$.*

Доказательство. В использовавшемся ранее базисе $e = (\overline{e_0}, \ldots, \overline{e_{n-1}})$ матрица преобразования ψ имеет следующий вид:

$$\psi \leftrightarrow_e A_p = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ -p_0 & -p_1 & \ldots & -p_{n-1} \end{pmatrix}$$

Докажем индукцией по n, что $|A_p - \lambda E| = (-1)^n p(\lambda)$. База, $n = 1$, тривиальна. Докажем переход, используя разложение определителя по первому столбцу:

$$\begin{vmatrix} -\lambda & 1 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ -p_0 & -p_1 & \ldots & -p_{n-1} - \lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & 1 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ -p_1 & -p_2 & \ldots & -p_{n-1} - \lambda \end{vmatrix} + (-1)^{n+1}(-p_0) =$$
$$= (-\lambda)(-1)^{n-1}(p_1 + p_2\lambda + \cdots + p_{n-1}\lambda^{n-2}) + (-1)^n p_0 = (-1)^n p(\lambda)$$

Таким образом, $\chi_\psi(\lambda) = (-1)^n p(\lambda)$. □

Замечание. Матрица $A_p \in M_n(F)$ называется *сопутствующей матрицей* многочлена $p \in F[x]$.

Теорема 3.3. *Пусть $p \in F[x]$ имеет вид $(\star)$. Тогда каждое решение линейной рекурренты с характеристическим многочленом p имеет следующий вид для некоторого набора коэффициентов $\{\beta_{it}\} \subset F$:*

$$(a_m) = \left(\sum_{i=1}^{k} \sum_{t=1}^{\alpha_i} \beta_{it} C_m^{t-1} \lambda_i^{m+1-t} \right)$$

Доказательство. Поскольку $\mu_\psi = \chi_\psi = (-1)^n p$, то для каждого собственного значения λ в жордановой нормальной форме оператора ψ есть ровно одна жорданова клетка с таким значением, и она имеет размер α. Найдем последовательности $A_1, A_2, \ldots, A_\alpha$ такие, что:

$$\varphi_\lambda(A_1) = (0)$$
$$\varphi_\lambda(A_2) = A_1$$
$$\ldots$$
$$\varphi_\lambda(A_\alpha) = A_{\alpha-1}$$

Все эти последовательности обнуляются оператором $(\varphi_\lambda)^\alpha$, и, как следствие, оператором $p(\varphi)$, поэтому они лежат в V_p. Именно они образуют жорданов базис в V^λ. Итак, пусть $A_1 = (1, \lambda, \lambda^2, \dots)$, тогда $\psi_\lambda(A) = (0)$. Теперь для каждого $t \in \{1, \dots, \alpha\}$ положим $A_t := (f_t(m)\lambda^{m+1-t})$ для некоторой функции $f : \mathbb{N}\cup\{0\} \to F$, причем $f_1 \equiv 1$ по построению. Считая A_t уже найденным, найдем A_{t+1}:

$$\begin{gathered}\varphi_\lambda(A_{t+1}) = A_t \\ f_{t+1}(m+1)\lambda^{m+1-t} - \lambda f_{t+1}(m)\lambda^{m-t} = f_t(m)\lambda^{m+1-t} \\ f_{t+1}(m+1) - f_{t+1}(m) = f_t(m)\end{gathered}$$

Базе $f_1 \equiv 1$ и рекуррентному соотношению выше удовлетворяет семейство функций $\{f_t\}_{t=1}^{\alpha}$ такое, что $f_t(m) = C_m^{t-1}$ для любых $t \in \{1, \dots, \alpha\}$ и $m \in \mathbb{N} \cup \{0\}$. Таким образом, $A_t = (C_m^{t-1}\lambda^{m+1-t})$ для каждого $t \in \{1, \dots, \alpha\}$. Объединение жордановых базисов клеток и будет искомым базисом в V_p, который позволяет представить каждый элемент из V_p в требуемом виде. $\square$

Замечание. Полученная формула справедлива в поле любой характеристики, поскольку число сочетаний — это целое число.

3.4 Поле разложения

Теорема 3.4. *Пусть F — поле, $p \in F[x]$ — неприводимый над F многочлен, $n := \deg p > 1$. Тогда существует такое поле $K \supset F$, в котором p имеет корень.*

Доказательство. Рассмотрим A_p — сопутствующую матрицу многочлена p и докажем, что условию удовлетворяет множество $K := F[A_p] = \{f(A_p) \mid f \in F[x]\} \subset M_n(F)$.

1. Очевидно, что K — подкольцо в $M_n(F)$. Более того, если $f \in F[x]$ — константа, то $f(A_p) = fE$. Матрицы вида fE образуют поле, изоморфное полю F, и лежат в K.

2. Умножение в K коммутативно, поскольку кольцо многочленов коммутативно, а подстановка матрицы A_p в $f \in F[x]$ — это гомоморфизм.

3. Пусть $f(A_p) \in K\backslash\{0\}$, тогда $p = \mu_{A_p} \nmid f$. Но многочлен p неприводим, поэтому НОД$(f, p) = 1$, и существуют многочлены $u, v \in F[x]$ такие, что $uf + vp = 1$. Подставляя в данное равенство A_p, получаем, что $u(A_p)f(A_p) = E$, то есть элемент $f(A_p)$ обратим. Значит, K является полем.

4. В поле K многочлен p имеет корень $A_p \in K$, поскольку $p(A_p) = 0$. $\square$

Следствие. *Пусть F — поле, $p \in F[x]$, $n := \deg p > 1$. Тогда существует такое поле $K \supset F$, над которым p раскладывается на линейные сомножители.*

Доказательство. Пусть p раскладывается на k неприводимых сомножителей над F. Проведем «дедукцию» по k. База, $k = n$, тривиальна, докажем переход. Если $k < n$, то в разложении многочлена p есть неприводимый сомножитель q, $\deg q > 1$. Тогда существует поле $F_1 \subset F$, в котором q имеет корень, и, следовательно, над F_1 многочлен p раскладывается на хотя $k+1$ неприводимых сомножителей, и применимо предположение «дедукции». $\square$

Замечание. Поле $\mathbb{C}$ было получено из $\mathbb{R}$ такой же процедурой: мы расширяли поле $\mathbb{R}$ корнями многочлена $p(x) := x^2 + 1 \in \mathbb{R}[x]$ с сопутствующей матрицей следующего вида:

$$A_p = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$$

Так как уже во второй степени эта матрица дает $-E$, можно считать, что все многочлены имеют степень не выше первой:

$$\mathbb{R}\left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right] = \left\{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} : a, b \in \mathbb{R}\right\}$$

Замечание. Теперь мы можем доказать теорему Гамильтона-Кэли в общем случае. Пусть $\varphi \in \mathcal{L}(V)$, V — линейное пространство над F, $\varphi \leftrightarrow_e A$. Рассмотрим K — надполе F, над которым χ_A раскладывается на линейные сомножители. Тогда, считая A матрицей над K, можно утверждать, что $\chi_A(A) = 0$, причем, поскольку все элементы A и все коэффициенты χ_A лежат в F, то в поле F вычисление $\chi_A(A)$ происходит аналогично, поэтому и над F данное утверждение верно.

Замечание. В прошлом семестре доказывалось, что если F — конечное поле, $\operatorname{char} F = p$, то $|F| = p^n$. Применяя теорему Лагранжа к группе F^*, получим, что $\forall a \in F^* : a^{p^n-1} = 1$, тогда $\forall a \in F : a^{p^n} = a$, то есть все элементы поля являются корнями многочлена $x^{p^n} - x$.

Теорема 3.5. *Пусть p — простое число, $n \in \mathbb{N}$. Тогда существует поле F такое, что $|F| = p^n$.*

Доказательство. Рассмотрим поле $\mathbb{Z}_p$ и найдем его надполе K, над которым многочлен $P := x^{p^n} - x$ раскладывается на линейные сомножители. Пусть $F \subset K$ — множество корней многочлена $P(x)$. Его производная $P'(x) = p^n x^{p^n-1} - 1 = -1$ не имеет корней, поэтому все корни P — простые, то есть $|F| = p^n$. Докажем, что $F = \{a \in K : a^{p^n} = a\}$ — поле.

1. Если $a, b \in F$, то $(ab)^{p^n} = a^{p^n} b^{p^n} = ab$, то есть $ab \in F$.

2. Если $a \in F$, то $(a^{-1})^{p^n} = (a^{p^n})^{-1} = a^{-1}$, то есть $a^{-1} \in F$.

3. Если $a, b \in F$, то $(a + b)^p = \sum_{i=0}^{p} C_p^i a^i b^{p-i} = a^p + b^p + pC = a^p + b^p$, следовательно, $(a + b)^{p^n} = (a^p + b^p)^{p^{n-1}} = \dots = a^{p^n} + b^{p^n} = a + b$, то есть $a + b \in F$.

4. Если $a \in F$, то $(-a)^{p^n} = (-1)^{p^n} a = -a$, то есть $-a \in F$.

Таким образом, F — подмножество в K, замкнутое относительно всех операций, поэтому F — подполе в K. $\square$

Замечание. Пусть F — поле, $P \in F[x]$, K — надполе F, над которым P раскладывается на линейные сомножители. K называется *полем разложения* P, если K — единственное подполе K, содержащее поле F и все корни многочлена P. Можно доказать, что поле разложения P единственно с точностью до изоморфизма. Значит, и поле из p^n элементов единственно с точностью до изоморфизма. Его обозначают через $\mathbb{F}_{p^n}$.

4 Билинейные и квадратичные формы

4.1 Билинейные формы

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$.

Определение 4.1. *Билинейной формой* на V называется функция $b : V \times V \to F$, линейная по обоим аргументам:

- $\forall \overline{x_1}, \overline{x_2}, \overline{y} \in V : b(\overline{x_1} + \overline{x_2}, \overline{y}) = b(\overline{x_1}, \overline{y}) + b(\overline{x_2}, \overline{y})$
- $\forall \overline{x}, \overline{y} \in V : \forall \alpha \in F : b(\alpha\overline{x}, \overline{y}) = \alpha b(\overline{x}, \overline{y})$
- $\forall \overline{x}, \overline{y_1}, \overline{y_2} \in V : b(\overline{x}, \overline{y_1} + \overline{y_2}) = b(\overline{x}, \overline{y_1}) + b(\overline{x}, \overline{y_2})$
- $\forall \overline{x}, \overline{y} \in V : \forall \alpha \in F : b(\overline{x}, \alpha\overline{y}) = \alpha b(\overline{x}, \overline{y})$

Множество всех билинейных форм на V обозначается через $\mathcal{B}(V)$.

Пример. Рассмотрим несколько примеров билинейных форм:

- Скалярное произведение в пространстве векторов V_n, $n \in \{1, 2, 3\}$, является билинейной формой на V_n
- Умножение в поле F является билинейной формой на F
- Функция b, заданная для произвольных $x, y \in F^2$ как $b(\overline{x}, \overline{y}) = |\overline{x}\ \overline{y}|$, является билинейной формой на F^2
- Функция b, заданная для произвольных $x, y \in F^3$ как $b(\overline{x}, \overline{y}) = |\overline{x}\ \overline{y}\ \overline{c}|$, где $\overline{c} \in F^3$ — фиксированный столбец, является билинейной формой на F^3

Определение 4.2. *Матрицей формы* $b \in \mathcal{B}(V)$ в базисе $(\overline{e_1}, \ldots, \overline{e_n}) =: e$ называется следующая матрица B:

$$B = (b(\overline{e_i}, \overline{e_j})) = \begin{pmatrix} b(\overline{e_1}, \overline{e_1}) & \ldots & b(\overline{e_1}, \overline{e_n}) \\ \vdots & \ddots & \vdots \\ b(\overline{e_n}, \overline{e_1}) & \ldots & b(\overline{e_n}, \overline{e_n}) \end{pmatrix} \in M_n(F)$$

Обозначение — $b \leftrightarrow_e B$.

Утверждение 4.1. *Пусть $b \in \mathcal{B}(V)$, e — базис в V, $b \leftrightarrow_e B$, $\overline{u}, \overline{v} \in V$, $\overline{u} \leftrightarrow_e x$, $\overline{v} \leftrightarrow_e y$. Тогда выполнено равенство $b(\overline{u}, \overline{v}) = x^T By$.*

Доказательство. Выполнены следующие равенства:

$$b(\overline{u}, \overline{v}) = b\left(\sum_{i=1}^n x_i\overline{e_i}, \sum_{j=1}^n y_j\overline{e_j}\right) = \sum_{i=1}^n x_i \left(\sum_{j=1}^n y_j b(\overline{e_i}, \overline{e_j})\right) = x^T By$$

Получено требуемое. □

Замечание. Множество $\mathcal{B}(V)$ образует линейное пространство над F с операциями сложения и умножения на скаляр, определенными следующим образом:

- $\triangleright$ $\forall b_1, b_2 \in \mathcal{B}(V) : \forall \overline{u}, \overline{v} \in V : (b_1 + b_2)(\overline{u}, \overline{v}) := b_1(\overline{u}, \overline{v}) + b_2(\overline{u}, \overline{v})$
- $\triangleright$ $\forall b \in \mathcal{B}(V) : \forall \alpha \in F : \forall \overline{u}, \overline{v} \in V : (\alpha b)(\overline{u}, \overline{v}) = \alpha b(\overline{u}, \overline{v})$

Теорема 4.1. *Пусть e — базис в V. Тогда соответствие между $\mathcal{B}(V)$ и $M_n(F)$ вида $b \leftrightarrow_e B$ осуществляет изоморфизм линейных пространств $\mathcal{B}(V)$ и $M_n(F)$.*

Доказательство.

1. Если $b_1, b_2 \in \mathcal{B}(V)$, $b_1 \leftrightarrow_e B_1$, $b_2 \leftrightarrow_e B_2$, то $b_1 + b_2 \leftrightarrow_e B_1 + B_2$, $\alpha b_1 \leftrightarrow_e \alpha B_1$. Значит, описанное в теореме отображение $\varphi : \mathcal{B}(V) \to M_n(F)$ линейно.
2. Отображение φ инъективно, поскольку если $\varphi(b_1) = \varphi(b_2) = B$, то $\forall \overline{u}, \overline{v} \in V$, $\overline{u} \leftrightarrow_e x$, $\overline{v} \leftrightarrow_e y$: $b_1(\overline{u}, \overline{v}) = b_2(\overline{u}, \overline{v})$, то есть $b_1 = b_2$.
3. Отображение φ сюръективно, поскольку для каждой матрицы $B \in M_n(F)$ можно определить билинейную форму b как $b(\overline{u}, \overline{v}) = x^T B y$, где $\overline{u} \leftrightarrow_e x, \overline{v} \leftrightarrow_e y$. Тогда если $B = (b_{ij})$, то $b(\overline{e_i}, \overline{e_j}) = b_{ij}$, то есть $b \leftrightarrow_e B$.

Таким образом, φ — это изоморфизм линейных пространств. □

Следствие. *Выполнено равенство* $\dim \mathcal{B}(V) = n^2$.

Теорема 4.2. *Пусть $b \in \mathcal{B}(V)$, e и e' — два базиса в V, $e' = eS$, и $b \leftrightarrow_e B$, $b \leftrightarrow_{e'} B'$. Тогда выполнено равенство $B' = S^T B S$.*

Доказательство. Перепишем S в виде $(s_1 | \ldots | s_n)$. Тогда:

$$B' = \begin{pmatrix} b(\overline{e'_1}, \overline{e'_1}) & \ldots & b(\overline{e'_1}, \overline{e'_n}) \\ \vdots & \ddots & \vdots \\ b(\overline{e'_n}, \overline{e'_1}) & \ldots & b(\overline{e'_n}, \overline{e'_n}) \end{pmatrix} = \begin{pmatrix} s_1^T B s_1 & \ldots & s_1^T B s_n \\ \vdots & \ddots & \vdots \\ s_n^T B s_1 & \ldots & s_n^T B s_n \end{pmatrix} = S^T B S$$

Получено требуемое. □

Следствие. *Пусть $b \in \mathcal{B}(V)$, e и e' — два базиса в V, и $b \leftrightarrow_e B$, $b \leftrightarrow_{e'} B'$. Тогда выполнено равенство* $\operatorname{rk} B = \operatorname{rk} B'$.

Определение 4.3. *Рангом билинейной формы $b \in \mathcal{B}(V)$* называется ранг ее матрицы в произвольном базисе. Обозначение — $\operatorname{rk} b$.

Следствие. *Пусть $F = \mathbb{R}$, $b \in \mathcal{B}(V)$, e и e' — два базиса в V, $e' = eS$, и $b \leftrightarrow_e B$, $b \leftrightarrow_{e'} B'$. Тогда* $\det B$ *и* $\det B'$ *имеют один и тот же знак.*

Доказательство. $\det B' = \det S^T B S = (\det S)^2 \det B$, что и означает требуемое в силу положительности числа $\det S$. □

Определение 4.4. Пусть $b \in \mathcal{B}(V)$. Форма b называется *симметрической*, если для всех $\overline{u}, \overline{v} \in V$ выполнено $b(\overline{u}, \overline{v}) = b(\overline{v}, \overline{u})$. Пространство симметрических форм на V обозначается через $\mathcal{B}^+(V)$.

Определение 4.5. Пусть $b \in \mathcal{B}(V)$. Форма b называется *кососимметрической*, если выполнены следующие условия:

1. $\forall \overline{u}, \overline{v} \in V : b(\overline{u}, \overline{v}) = -b(\overline{v}, \overline{u})$
2. $\forall \overline{u} \in V : b(\overline{u}, \overline{u}) = 0$

Пространство кососимметрических форм на V обозначается через $\mathcal{B}^-(V)$.

Замечание. В определении выше из условия (2) следует условие (1). Действительно, пусть выполенно условие (2), тогда:

$$0 = b(\overline{u}+\overline{v}, \overline{u}+\overline{v}) = b(\overline{u}, \overline{u}) + b(\overline{u}, \overline{v}) + b(\overline{v}, \overline{u}) + b(\overline{v}, \overline{v}) = b(\overline{u}, \overline{v}) + b(\overline{v}, \overline{u})$$

Значит, $b(\overline{u}, \overline{v}) = -b(\overline{v}, \overline{u})$. При этом из условия (1) следует условие (2) лишь в том случае, когда $\operatorname{char} F \neq 2$.

Утверждение 4.2. *Пусть e — базис в V, $b \in \mathcal{B}(V)$, и $b \leftrightarrow_e B$. Тогда:*

1. *$b \in \mathcal{B}^+(V) \Leftrightarrow B^T = B$*
2. *$b \in \mathcal{B}^-(V) \Leftrightarrow B^T = -B$ и на главной диагонали B стоят нули*

Доказательство.

$\Rightarrow$ Поскольку $B = (b_{ij}) = (b(\overline{e_i}, \overline{e_j}))$, то непосредственная подстановка базисных векторов позволяет убедиться, что в первом случае $B^T = B$, а во втором — $B^T = -B$, и на главной диагонали матрицы B стоят нули.

$\Leftarrow$ Пусть $\overline{u}, \overline{v} \in V$, $\overline{u} \leftrightarrow_e x$, $\overline{v} \leftrightarrow_e y$. Тогда выполнены следующие цепочки равенств:

1. $b(\overline{u}, \overline{v}) = x^T B y = (x^T B y)^T = y^T B x = b(\overline{v}, \overline{u})$
2. $b(\overline{u}, \overline{u}) = x^T B x = \sum\limits_{i=1}^n \sum\limits_{j=1}^n x_i b_{ij} x_j = \sum\limits_{i=1}^n x_i b_{ii} x_i + \sum\limits_{1 \leqslant i < j \leqslant n} x_i (b_{ij} + b_{ji}) x_j = 0$

В обоих случаях получено требуемое. □

4.2 Симметрические билинейные и квадратичные формы

До конца раздела зафиксируем линейное пространство V над полем F и положим $n := \dim V$.

Теорема 4.3. *Пусть $\operatorname{char} F \neq 2$. Тогда $\mathcal{B}(V) = \mathcal{B}^+(V) \oplus \mathcal{B}^-(V)$.*

Доказательство. Рассмотрим произвольную форму $b \in \mathcal{B}(V)$ и зададим $b^+ \in \mathcal{B}^+$, $b^- \in \mathcal{B}^-$ на произвольных $\overline{u}, \overline{v} \in V$ следующим образом:

$$b^+(\overline{u}, \overline{v}) := \frac{b(\overline{u}, \overline{v}) + b(\overline{v}, \overline{u})}{2}, \; b^-(\overline{u}, \overline{v}) := \frac{b(\overline{u}, \overline{v}) - b(\overline{v}, \overline{u})}{2}$$

Тогда $b = b^+ + b^-$, и, в силу произвольности выбора формы $b \in \mathcal{B}(V)$, получено равенство $\mathcal{B}(V) = \mathcal{B}^+(V) + \mathcal{B}^-(V)$. Проверим теперь, что $\mathcal{B}^+(V) \cap \mathcal{B}^-(V) = \{0\}$. Действительно, если $b \in \mathcal{B}^+(V) \cap \mathcal{B}^-(V)$, то для любых $\overline{u}, \overline{v} \in V$ выполнено следующее:

$$b(\overline{u}, \overline{v}) = b(\overline{v}, \overline{u}) = -b(\overline{u}, \overline{v}) \Rightarrow b(\overline{u}, \overline{v}) = 0$$

Значит, $b = 0$, поэтому сумма $\mathcal{B}^+(V) + \mathcal{B}^-(V)$ — прямая. □

Замечание. Если $\operatorname{char} F = 2$, то теорема выше уже неверна, поскольку выполнено включение $\mathcal{B}^-(V) \leqslant \mathcal{B}^+(V)$.

Определение 4.6. Пусть $b \in \mathcal{B}^{\pm}(V)$. *Ядром* формы b называется подпространство $\operatorname{Ker} b := \{\overline{v} \in V : \forall \overline{u} \in V : b(\overline{u}, \overline{v}) = 0\} = \{\overline{u} \in V : \forall \overline{v} \in V : b(\overline{u}, \overline{v}) = 0\} \leqslant V$.

Замечание. Для произвольной билинейной формы $b \in \mathcal{B}(V)$ два множества выше необязательно совпадают и называются *правым* и *левым ядром* соответственно.

Теорема 4.4. *Для любой формы $b \in \mathcal{B}^{\pm}(V)$ выполнено равенство $\dim \operatorname{Ker} b = \dim V - \operatorname{rk} b$.*

Доказательство. Пусть e — произвольный базис в V, $b \leftrightarrow_e B$. Тогда для произвольного вектора $\overline{v} \in V$, $\overline{v} \leftrightarrow_e x$, выполнены следующие равносильности:

$$\overline{v} \in \operatorname{Ker} b \Leftrightarrow \forall \overline{u} \in V : b(\overline{u}, \overline{v}) = 0 \Leftrightarrow \forall i \in \{1, \ldots, n\} : b(\overline{e_i}, \overline{v}) = 0$$

Последнее из условий выше равносильно тому, что $Bx = 0$. Размерность пространства решений полученной однородной системы равна $\dim V - \operatorname{rk} B = \dim V - \operatorname{rk} b$. □

Определение 4.7. Пусть $b \in \mathcal{B}^{\pm}(V)$.

- ▷ Векторы $\overline{u}, \overline{v} \in V$ называются *ортогональными относительно* b, если $b(\overline{u}, \overline{v}) = 0$.
- ▷ *Ортогональным дополнением подпространства $U \leqslant V$ относительно* b называется подпространство $U^{\perp} := \{\overline{v} \in V : \forall \overline{u} \in U : b(\overline{u}, \overline{v}) = 0\} \leqslant V$.

Пример. Справедливы следующие утверждения:

1. Если b — скалярное произведение в пространстве V_n, $n \in \{1, 2, 3\}$, то ортогональность относительно b соответствует ортогональности геометрических векторов. Тогда, например, ортогональное дополнение к плоскости в V_3 относительно b — это прямая, перпендикулярная ей.

2. Если $b = 0$, то для любого подпространства $U \leqslant V$ выполнено равенство $U^{\perp} = V$.

Определение 4.8. Форма $b \in \mathcal{B}^{\pm}(V)$ называется *невырожденной*, если $\operatorname{rk} b = \dim V$.

Теорема 4.5. *Пусть $b \in \mathcal{B}^{\pm}(V)$, $U \leqslant V$. Тогда:*

1. $\dim U^{\perp} \geqslant \dim V - \dim U$

2. *Если форма b — невырожденная, то* $\dim U^{\perp} = \dim V - \dim U$

Доказательство. Пусть $n := \dim V$, $k := \dim U$. Дополним базис $(\overline{e_1}, \ldots, \overline{e_k})$ в U до базиса $e := (\overline{e_1}, \ldots, \overline{e_n})$ в V. Тогда, если $\overline{v} \in V$, $\overline{v} \leftrightarrow_e x$, то $\overline{v} \in U^{\perp} \Leftrightarrow \forall i \in \{1, \ldots, k\} : b(\overline{e_i}, \overline{v}) = 0 \Leftrightarrow B'x = 0$, где B' — матрица из первых k строк матрицы B. Тогда, так как $\operatorname{rk} B' \leqslant k$, то $\dim U^{\perp} \geqslant \dim V - \dim U$. Более того, если форма b невырожденна, то матрица B тоже невырожденна, откуда $\operatorname{rk} B' = k$ и $\dim U^{\perp} = \dim V - \dim U$. □

Определение 4.9. Подпространство $U \leqslant V$ называется *невырожденным относительно* $b \in \mathcal{B}^{\pm}(V)$, если ограничение $b|_U \in \mathcal{B}^{\pm}(U)$ невырожденно.

Теорема 4.6. *Пусть $b \in \mathcal{B}^{\pm}(V)$. Тогда подпространство $U \leqslant V$ невырожденно относительно $b \Leftrightarrow V = U \oplus U^{\perp}$.*

Доказательство. Пусть $n := \dim V$, $k := \dim U$.

$\Rightarrow$ Построим базис e аналогично предыдущей теореме. Если $b \leftrightarrow_e B$, то матрица $b|_U$ в базисе $(\overline{e_1}, \ldots, \overline{e_k})$ — это левый верхний угол B' матрицы B. Поскольку U невырожденно относительно b, то $\operatorname{rk} B' = k$, следовательно, первые k строк B линейно независимы. Значит, $\dim U^\perp = n - k$. Кроме того, $\operatorname{Ker} b|_U = \{\overline{0}\}$, так как $\dim \operatorname{Ker} b|_U = 0$. Значит, $\forall \overline{v} \in U, \overline{v} \neq \overline{0} : \exists \overline{u} \in U : b(\overline{u}, \overline{v}) \neq 0$. Это значит, что $U \cap U^\perp = \{\overline{0}\}$, то есть сумма $U + U^\perp$ — прямая, тогда $\dim(U \oplus U^\perp) = n$ и потому $U \oplus U^\perp = V$.

$\Leftarrow$ Если сумма $U \oplus U^\perp$ прямая, то $U \cap U^\perp = \{\overline{0}\}$, тогда $\operatorname{Ker} b|_U = \{\overline{0}\}$, откуда $\operatorname{rk} B' = k$ и U невырожденно относительно b. $\square$

Пример. Зафиксируем базис $e := (\overline{e_1}, \overline{e_2})$ в пространстве F^2, и рассмотрим следующие матрицы:

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 1 \\ \pm 1 & 0 \end{pmatrix}$$

Тогда верны следующие утверждения:

1. Форма $b \leftrightarrow_e B$ вырожденна, но $\langle \overline{e_1} \rangle$ невырожденно относительно b.

2. Форма $b \leftrightarrow_e C$ невырожденна, но $\langle \overline{e_1} \rangle$ вырожденно относительно b.

Утверждение 4.3. *Пусть $b \in \mathcal{B}^\pm(V)$, $U \leqslant V$ и $V = U \oplus U^\perp$. Выберем базисы $(\overline{e_1}, \ldots, \overline{e_k})$ в U и $(\overline{e_{k+1}}, \ldots, \overline{e_n})$ в $U^\perp$. Тогда в базисе $(\overline{e_1}, \ldots, \overline{e_n})$ в V форма b имеет матрицу:*

$$B = \left(\begin{array}{c|c} B_1 & 0 \\ \hline 0 & B_2 \end{array}\right),$$

где B_1, B_2 — матрицы форм $b|_U$ и $b|_{U^\perp}$

Доказательство. Если $i \in \{1, \ldots, k\}$ и $j \in \{k+1, \ldots, n\}$, то по определению ортогонального дополнения $b(\overline{e_i}, \overline{e_j}) = 0$. $\square$

Определение 4.10. *Квадратичной формой*, соответствующей форме $b \in \mathcal{B}(V)$, называется функция $h : V \to F$ такая, что $\forall \overline{v} \in V : h(\overline{v}) = b(\overline{v}, \overline{v})$. Квадратичные формы на V образуют линейное пространство над F, обозначаемое через $\mathcal{Q}(V)$.

Замечание. Разным билинейным формам могут соответствовать одинаковые квадратичные. Например, любой кососимметрической форме соответствует нулевая квадратичная.

Теорема 4.7. *Пусть* $\operatorname{char} F \neq 2$. *Тогда для любой квадратичной формы h на V существует единственная форма $b \in \mathcal{B}^+(V)$, соответствующая ей.*

Доказательство. Пусть $b \in \mathcal{B}(V)$, $h(\overline{v}) = b(\overline{v}, \overline{v})$. Как уже было доказано, b представима в виде $b^+ + b^-$, где $b^\pm \in \mathcal{B}^\pm(V)$, тогда $h(\overline{v}) = b(\overline{v}, \overline{v}) = b^+(\overline{v}, \overline{v})$. Более того, по h можно однозначно восстановить $b^+ \in \mathcal{B}^+(V)$ следующим образом: $b^+(\overline{u}, \overline{v}) = \frac{h(\overline{u}+\overline{v}) - h(\overline{u}) - h(\overline{v})}{2}$. $\square$

Замечание. Согласно теореме выше, если $\operatorname{char} F \neq 2$, то $\mathcal{Q}(V) \cong \mathcal{B}^+(V)$, и изоморфизм осуществляется описанным выше образом. Если же $\operatorname{char} F = 2$, то теорема неверна. Рассмотрим следующую билинейную форму:

$$b \leftrightarrow_e B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Соответствующая данной билинейной форме квадратичная форма не выражается никакой симметрической билинейной формой.

Определение 4.11. Пусть $\operatorname{char} F \neq 2$, $h \in \mathcal{Q}(V)$. Симметрическая билинейная форма $b \in \mathcal{B}^+(V)$ называется *полярной к* h, если $\forall \overline{v} \in V : h(\overline{v}) = b(\overline{v}, \overline{v})$. *Матрицей квадратичной формы* h в базисе e называется матрица B полярной к ней формы b в базисе e. Обозначение — $h \leftrightarrow_e B$.

Замечание. Матрица квадратичной формы всегда симметрична.

Теорема 4.8. *Пусть* $\operatorname{char} F \neq 2$, $h \in \mathcal{Q}(V)$. *Тогда в пространстве* V *существует такой базис* e, *что* h *в этом базисе имеет диагональную матрицу.*

Доказательство. Проведем индукцию по $n := \dim V$. База, $n = 1$, тривиальна. Пусть теперь $n > 1$, $h \in \mathcal{Q}(V)$, $b \in \mathcal{B}^+(V)$ — полярная к h форма. Если $h = 0$, то $b = 0$, поэтому у h нулевая матрица. Если же $h \neq 0$, то $\exists \overline{e_1} \in V : h(\overline{e_1}) \neq 0$. Тогда $\langle \overline{e_1} \rangle$ невырожденно относительно b, откуда $V = \langle \overline{e_1} \rangle \oplus \langle \overline{e_1} \rangle^\perp$. По предположению индукции, в $U := \langle \overline{e_1} \rangle^\perp$ существует базис, в котором матрица $h|_U$ диагональна, и объединение этого базиса с $\overline{e_1}$ и дает искомый базис в V. □

Замечание. Диагональные элементы разных диагональных матриц одной квадратичной формы могут быть различными. Например, в базисе $(2\overline{e_1}, \overline{e_2}, \ldots, \overline{e_n})$ элемент в левом верхнем углу в 4 раза больше, чем в $(\overline{e_1}, \overline{e_2}, \ldots, \overline{e_n})$, так как он равен $h(2\overline{e_1}) = b(2\overline{e_1}, 2\overline{e_1}) = 4b(\overline{e_1}, \overline{e_1}) = 4h(\overline{e_1})$

Следствие. *Пусть* $F = \mathbb{R}$, $h \in \mathcal{Q}(V)$. *Тогда в пространстве* V *существует такой базис* e, *что* h *в этом базисе имеет диагональную матрицу с числами* 0 *и* ± 1 *на главной диагонали.*

Доказательство. Пусть $f = (\overline{f_1}, \ldots, \overline{f_n})$ — базис, в котором матрица h диагональна. Тогда искомым базисом является базис $e = (\overline{e_1}, \ldots, \overline{e_n})$ такой, что для каждого $i \in \{1, \ldots, n\}$ вектор $\overline{e_i}$ имеет вид:

$$\overline{e_i} := \begin{cases} \overline{f_i}, \text{ если } h(\overline{f_i}) = 0 \\ \dfrac{1}{\sqrt{|h(\overline{f_i})|}} \overline{f_i}, \text{ если } h(\overline{f_i}) \neq 0 \end{cases}$$

В базисе e на главной диагонали матрицы будут только числа 0 и ± 1, а элементы вне диагонали останутся нулевыми. □

4.3 Положительная и отрицательная определенность

В данном разделе будем считать, что V — линейное пространство над $\mathbb{R}$.

Определение 4.12. Пусть $h \in \mathcal{Q}(V)$. Базис, в котором матрица h диагональна с числами 0 и ± 1 на главной диагонали, называется *нормальным базисом*, а матрица h в этом базисе — *нормальной формой* h.

Определение 4.13. Пусть $h \in \mathcal{Q}(V)$. Тогда h называется:

▷ *положительно определенной*, если $\forall \overline{v} \in V, \overline{v} \neq \overline{0} : h(\overline{v}) > 0$.

▷ *положительно полуопределенной*, если $\forall \overline{v} \in V : h(\overline{v}) \geqslant 0$.

▷ *отрицательно определенной*, если $\forall \overline{v} \in V, \overline{v} \neq \overline{0} : h(\overline{v}) < 0$.

▷ *отрицательно полуопределенной*, если $\forall \overline{v} \in V : h(\overline{v}) \leqslant 0$.

Форма $b \in \mathcal{B}^+(V)$, полярная к h, приобретает те же названия.

Утверждение 4.4. *Пусть $h \in \mathcal{Q}(V)$, B — нормальная форма h в нормальном базисе e. Тогда:*

1. *h положительно определена $\Leftrightarrow B = E$.*

2. *h положительно полуопределена $\Leftrightarrow$ на диагонали B стоят только нули и единицы.*

Доказательство.

$\Leftarrow$ Пусть $\overline{v} \in V, v \neq 0, \overline{v} \leftrightarrow_e x$. Если $B = E$, то $h(\overline{v}) = x^T Bx = x_1^2 + \dots + x_n^2 > 0$. Если на диагонали B расположены только нули и единицы, то $h(\overline{v})$ — это сумма квадратов некоторых координат вектора $\overline{v}$, откуда $h(\overline{v}) \geqslant 0$.

$\Rightarrow$ i-й диагональный элемент матрицы B — это $h(\overline{e_i})$. Если h положительно определена, то $\forall i \in \{1, \dots, n\} : h(\overline{e_i}) > 0$, то есть все диагональные элементы — единицы. Если же h положительно полуопределена, то $\forall i \in \{1, \dots, n\} : h(\overline{e_i}) \geqslant 0$, то есть диагональные элементы — это нули и единицы. □

Замечание. Для случая отрицательной определенности и полуопределенности утверждение аналогично при замене единиц на минус единицы. Более того, h положительно определена (полуопределена) $\Leftrightarrow -h$ отрицательно определена (полуопределена).

Определение 4.14. Пусть $h \in \mathcal{Q}(V)$. Ее *положительным индексом инерции* $\sigma_+(h)$ называется наибольшая размерность подпространства $U \leqslant V$ такого, что $h|_U$ положительно определена, *отрицательным индексом инерции* $\sigma_-(h)$ — наибольшая размерность подпространства $U \leqslant V$ такого, что $h|_U$ отрицательно определена.

Замечание. Множества $\{\overline{v} \in V \mid h(\overline{v}) > 0\}$, $\{\overline{v} \in V \mid h(\overline{v}) \geqslant 0\}$ не всегда является подпространствами в V. Например, это неверно для следующей квадратичной формы:

$$h \leftrightarrow_e B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Теорема 4.9. *Пусть $h \in \mathcal{Q}(V)$, B — нормальная форма h в нормальном базисе e. Тогда на диагонали матрицы B стоит ровно $\sigma_+(h)$ единиц и ровно $\sigma_-(h)$ минус единиц.*

Доказательство. Пусть $n := \dim V$. Без ограничения общности можно считать, что нормальная форма h имеет следующий вид:

$$B = \begin{pmatrix} \boxed{E_k} & 0 & 0 \\ 0 & \boxed{0_l} & 0 \\ 0 & 0 & \boxed{-E_m} \end{pmatrix}, \; k + l + m = n$$

Пусть $U := \langle \overline{e_1}, \dots, \overline{e_k} \rangle$, $W := \langle \overline{e_{k+1}}, \dots, \overline{e_n} \rangle$. Сужение $h|_U$ положительно определено, откуда $\sigma_+(h) \geqslant k$. С другой стороны, сужение $h|_W$ отрицательно полуопределено. Пусть теперь $U' \leqslant V$ — такое, что $h|_{U'}$ положительно определено, тогда $U' \cap W = \{\overline{0}\}$, поскольку $\forall \overline{w} \in W : h(\overline{w}) \leqslant 0$. Значит, $\sigma_+(h) = k$. Аналогично, $\sigma_-(h) = m$. □

Следствие (закон инерции)**.** *Нормальный вид квадратичной формы $h \in \mathcal{Q}(V)$ определен однозначно с точностью до перестановки диагональных элементов.*

Определение 4.15. Пусть $B \in M_n(\mathbb{R})$ — симметричная матрица. B называется *положительно* или *отрицательно определенной (полуопределенной)*, если она задает квадратичную форму, обладающую этим свойством.

Утверждение 4.5. *$B \in M_n(\mathbb{R})$ положительно определена $\Leftrightarrow \exists A \in \mathrm{GL}_n(\mathbb{R})$: $B = A^T A$.*

Доказательство. Квадратичная форма $h \leftrightarrow_e B$ положительно определена $\Leftrightarrow$ h имеет нормальный вид E $\Leftrightarrow$ для некоторой матрицы $S \in \mathrm{GL}_n(\mathbb{R})$ выполнено $E = S^T B S$ $\Leftrightarrow$ для некоторой матрицы $A \in \mathrm{GL}_n(\mathbb{R})$ выполнено $B = A^T A$. □

Утверждение 4.6. *$B \in M_n(\mathbb{R})$ положительно полуопределена $\Leftrightarrow \exists A \in M_n(\mathbb{R})$: $B = A^T A$.*

Доказательство. Аналогично утверждению выше с заменой нормального вида E на нормальный вид $E' = \mathrm{diag}(1, \ldots, 1, 0, \ldots, 0)$. □

Определение 4.16. Пусть $B \in M_n(\mathbb{R})$ — симметричная матрица. Ее *главным минором порядка i* называется $\Delta_i(B)$ — определитель подматрицы размера $i \times i$, расположенной в левом верхнем углу B.

Теорема 4.10 (Метод Якоби)**.** *Пусть $h \in \mathcal{Q}(V)$, $h \leftrightarrow_e B$, причем все главные миноры матрицы B отличны от нуля. Тогда существует такой базис $e' = eS$, что матрица перехода S — верхнетреугольная с единицами на главной диагонали, $h \leftrightarrow_{e'} B'$ и B' диагональна. Более того, тогда $B' = \mathrm{diag}(\Delta_1(B), \frac{\Delta_2(B)}{\Delta_1(B)}, \ldots, \frac{\Delta_n(B)}{\Delta_{n-1}(B)})$.*

Доказательство. Докажем индукцией по $n := \dim V$, что матрица формы h приводится к диагональному виду в базисе с матрицей перехода из условия. База, $n = 1$, тривиальна: подходит исходный базис e. Пусть теперь $n > 1$, тогда $U := \langle \overline{e_1}, \ldots, \overline{e_{n-1}} \rangle$ невырожденно относительно формы b, полярной к h, так как $\Delta_{n-1}(B) \neq 0$. Значит, $V = U \oplus U^\perp$. Представим $\overline{e_n}$ в виде $\overline{e_n} = \overline{u} + \overline{e'_n}$, где $\overline{u} \in U$, $\overline{e'_n} \in U^\perp$, причем $\overline{e'_n} \neq \overline{0}$. По предположению индукции, в U можно выбрать подходящий базис $(\overline{e'_1}, \ldots, \overline{e'_{n-1}})$, тогда его объединение с $\overline{e'_n}$ будет искомым. Матрица перехода S действительно будет верхнетреугольной с единицами на главной диагонали: для первых $n-1$ столбцов это верно по предположению индукции, для последнего столбца — в силу того, что $\overline{e'_n} = \overline{e_n} - u$.

Вычислим теперь значения диагональных элементов d_i, $i \in \{1, \ldots, n\}$. Заметим, что, поскольку базис e' получен описанным выше образом, $\forall i \in \{1, \ldots, n\} : \overline{e'_i} \in \langle \overline{e_1}, \ldots, \overline{e_i} \rangle$ и $\langle \overline{e_1}, \ldots, \overline{e_i} \rangle = \langle \overline{e'_1}, \ldots, \overline{e'_i} \rangle$. Пусть B_i — подматрица B в левом верхнем углу, а B'_i — аналогичная подматрица B'. Тогда $B'_i = S_i^T B_i S_i$, где S_i — соответствующая подматрица S, также являющаяся верхнетреугольной с единицами на диагонали, поэтому $\Delta_i(B') = |B'_i| = |S_i^T B_i S_i| = |B_i| = \Delta_i(B)$. Значит, $\forall i \in \{1, \ldots, n\} : \Delta_i(B) = \Delta_i(B') = d_1 \ldots d_i$, откуда $B' = \mathrm{diag}(\Delta_1(B), \frac{\Delta_2(B)}{\Delta_1(B)}, \ldots, \frac{\Delta_n(B)}{\Delta_{n-1}(B)})$. □

Теорема 4.11 (Критерий Сильвестра)**.** *Пусть $h \in \mathcal{Q}(V)$, $h \leftrightarrow_e B$. Тогда h положительно определена $\Leftrightarrow \forall i \in \{1, \ldots, n\} : \Delta_i(B) > 0$.*

Доказательство. Пусть $n := \dim V$.

- $\Rightarrow$ Если h положительно определена, то $B = A^T A$ для некоторой $A \in \mathrm{GL}_n(\mathbb{R})$. Тогда $\Delta_n(B) = |B| = |A|^2 > 0$. Поскольку главному минору порядка $i \in \{1, \ldots, n-1\}$ соответствует ограничение h на $U := \langle \overline{e_1}, \ldots, \overline{e_i} \rangle$, которое тоже положительно определено, то, аналогично, $\Delta_i(B) > 0$.

$\Leftarrow$ Согласно методу Якоби, существует базис e' в V такой, что матрица h в нем диагональна, причем $h \leftrightarrow_{e'} \operatorname{diag}(\Delta_1(B), \frac{\Delta_2(B)}{\Delta_1(B)}, \ldots, \frac{\Delta_n(B)}{\Delta_{n-1}(B)})$. Все элементы на главной диагонали положительны, поэтому h положительно определена. $\square$

Замечание. Если $\forall i \in \{1, \ldots, n\} : \Delta_i(B) \neq 0$, то, согласно методу Якоби, $\sigma_-(h)$ — это число перемен знака в последовательности $(1, \Delta_1(B), \ldots, \Delta_n(B))$, а $\sigma_+(h)$ — число сохранений знака в этой же последовательности.

Замечание. Прямой аналог критерия Сильвестра для положительной полуопределенности неверен. Например, пусть B имеет следующий вид:

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

Тогда $\Delta_1(B) = 1, \Delta_2(B) = \Delta_3(B) = 0$, но при этом B не является положительно полуопределенной. Тем не менее, можно показать, что $B \in M_n(\mathbb{R})$ положительно полуопределена $\Leftrightarrow$ все ее симметричные относительно главной диагонали миноры неотрицательны.

Теорема 4.12. *Пусть $b \in \mathcal{B}^-(V)$. Тогда в V существует базис e, в котором матрица b имеею следующий вид:*

$$b \leftrightarrow_e B = \begin{pmatrix} \boxed{B_1} & 0 & \ldots & 0 \\ 0 & \boxed{B_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \boxed{B_m} \end{pmatrix},$$

где $\forall i \in \{1, \ldots, m\} : B_i = (0)$ или $B_i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Доказательство. Докажем данное утверждение индукцией по $n := \dim V$. База, $n = 1$, и случай, когда $b = 0$, тривиальны. Пусть теперь $\exists \overline{e_1}, \overline{e_2} \in V : b(\overline{e_1}, \overline{e_2}) \neq 0$. Тогда эти векторы линейно независимы в силу кососимметричности формы b, и без ограничения общности можно считать, что $b(\overline{e_1}, \overline{e_2}) = 1$. Рассмотрим ограничение b на $U := \langle \overline{e_1}, \overline{e_2} \rangle$, тогда в базисе $e' := (\overline{e_1}, \overline{e_2})$ матрица ограничения имеет вид:

$$b|_U \leftrightarrow_{e'} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Заметим, что U невырожденно относительно b, откуда $V = U \oplus U^\perp$, и к $U^\perp$ применимо предположение индукции. $\square$

Следствие. *Если $b \in \mathcal{B}^-(V)$, то $\operatorname{rk} b$ — четное число.*

Замечание. Полученные матрица и базис называются *нормальными* для кососимметричной формы b. Перестановкой базисных векторов нормальную форму можно также преобразовать к следующему виду:

$$B = \begin{pmatrix} 0 & \boxed{E_k} & 0 \\ \boxed{-E_k} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4.4 Эрмитовы и эрмитовы квадратичные формы

В данном разделе будем считать, что V — линейное пространство над $\mathbb{C}$.

Определение 4.17. *Полуторалинейной формой* на V называется функция $b : V \times V \to \mathbb{C}$ такая, что:

1. b линейна по первому аргументу:
   - ▷ $\forall \overline{u_1}, \overline{u_2}, \overline{v} \in V : b(\overline{u_1} + \overline{u_2}, \overline{v}) = b(\overline{u_1}, \overline{v}) + b(\overline{u_2}, \overline{v})$
   - ▷ $\forall \overline{u}, \overline{v} \in V : \forall \lambda \in \mathbb{C} : b(\lambda \overline{u}, \overline{v}) = \lambda b(\overline{u}, \overline{v})$
2. b *сопряженно-линейна* по второму аргументу:
   - ▷ $\forall \overline{u}, \overline{v_1}, \overline{v_2} \in V : b(\overline{u}, \overline{v_1} + \overline{v_2}) = b(\overline{u}, \overline{v_1}) + b(\overline{u}, \overline{v_2})$
   - ▷ $\forall \overline{u}, \overline{v} \in V : \forall \lambda \in \mathbb{C} : b(\overline{u}, \lambda \overline{v}) = \overline{\lambda} b(\overline{u}, \overline{v})$

Полуторалинейные формы на V образуют линейное пространство над F, обозначаемое через $\mathcal{S}(V)$.

Определение 4.18. *Матрицей формы* $b \in \mathcal{S}(V)$ в базисе $(\overline{e_1}, \ldots, \overline{e_n}) =: e$ называется следующая матрица B:

$$B = (b(\overline{e_i}, \overline{e_j})) = \begin{pmatrix} b(\overline{e_1}, \overline{e_1}) & \ldots & b(\overline{e_1}, \overline{e_n}) \\ \vdots & \ddots & \vdots \\ b(\overline{e_n}, \overline{e_1}) & \ldots & b(\overline{e_n}, \overline{e_n}) \end{pmatrix} \in M_n(F)$$

Обозначение — $b \leftrightarrow_e B$.

Замечание. Аналогично билинейному случаю, для любых $\overline{u}, \overline{v} \in V$, $\overline{u} \leftrightarrow_e x, \overline{v} \leftrightarrow_e y$, выполнено $b(\overline{u}, \overline{v}) = x^T B \overline{y}$.

Утверждение 4.7. *$\mathcal{S}(V) \cong M_n(\mathbb{C})$.*

Доказательство. Доказательство аналогично билинейному случаю. □

Теорема 4.13. *Пусть $b \in \mathcal{S}(V)$, e и e' — два базиса в V, $e' = eS$. Если $b \leftrightarrow_e B$ и $b \leftrightarrow_{e'} B'$, то $B' = S^T B \overline{S}$.*

Доказательство. Доказательство аналогично билинейному случаю. □

Следствие. *Пусть $b \in \mathcal{S}(V)$, e и e' — два базиса в V. Тогда если $b \leftrightarrow_e B$, $b \leftrightarrow_{e'} B'$, то $\operatorname{rk} B = \operatorname{rk} B'$.*

Определение 4.19. *Рангом полуторалинейной формы* $b \in \mathcal{S}(V)$ называется ранг ее матрицы в произвольном базисе. Обозначение — $\operatorname{rk} b$.

Следствие. *Пусть $b \in \mathcal{S}(V)$, e и e' — два базиса в V. Тогда если $b \leftrightarrow_e B$, $b \leftrightarrow_{e'} B'$, то $\arg |B| = \arg |B'|$.*

Доказательство. Поскольку $B' = S^T B \overline{S}$, то $|B'| = |S^T||B||\overline{S}| = |B||\det S|^2$. Значит, $|B|$ и $|B'|$ отличаются друг от друга умножением на положительное вещественное число, откуда $\arg |B| = \arg |B'|$. □

Определение 4.20. Пусть $b \in \mathcal{S}(V)$. Форма b называется *эрмитовой*, если для всех $\overline{u}, \overline{v} \in V$ выполнено $b(\overline{u}, \overline{v}) = \overline{b(\overline{v}, \overline{u})}$. Матрица $B \in M_n(\mathbb{C})$ называется *эрмитовой*, если $B^T = \overline{B}$, или $B = B^*$, где $B^* := \overline{B^T}$ — *эрмитово сопряженная* к B матрица.

Теорема 4.14. *Пусть e — базис в V, $b \in \mathcal{S}(V)$, $b \leftrightarrow_e B$. Тогда форма b — эрмитова $\Leftrightarrow$ $\Leftrightarrow$ матрица B — эрмитова.*

Доказательство.

$\Rightarrow$ Утверждение доказывается непосредственной проверкой.

$\Leftarrow$ Пусть $B^T = \overline{B}$. Тогда $\overline{b(\overline{v}, \overline{u})} = \overline{y^T B \overline{x}} = \overline{\overline{x}^T B^T y} = x^T \overline{B^T} \overline{y} = x^T B \overline{y} = b(\overline{u}, \overline{v})$. ☐

Определение 4.21. *Эрмитовой квадратичной формой*, соответствующей эрмитовой форме $b \in \mathcal{S}(V)$, называется функция $h : V \to \mathbb{C}$ такая, что $\forall \overline{v} \in V : h(\overline{v}) = b(\overline{v}, \overline{v})$. Форма b называется *полярной* к h.

Утверждение 4.8. *Пусть $b \in \mathcal{S}(V)$ — эрмитова форма, тогда:*

1. $\forall \overline{v} \in V : b(\overline{v}, \overline{v}) \in \mathbb{R}$
2. *Если e — базис в V и $b \leftrightarrow_e B$, то $\det B \in \mathbb{R}$*

Доказательство.

1. $\forall \overline{v} \in V : b(\overline{v}, \overline{v}) = \overline{b(\overline{v}, \overline{v})} \Rightarrow b(\overline{v}, \overline{v}) \in \mathbb{R}$.
2. $B^T = \overline{B} \Rightarrow \det B = \det \overline{B} = \overline{\det B} \Rightarrow \det B \in \mathbb{R}$. ☐

Следствие. *Эрмитова квадратичная форма принимает только вещественные значения.*

Утверждение 4.9. *Если $b_1, b_2 \in \mathcal{S}(V)$ — различные эрмитовы формы, то соответствующие им квадратичные формы также различны.*

Доказательство. Пусть h — эрмитова квадратичная форма. Восстановим эрмитову форму $b \in \mathcal{S}(V)$, полярную к h:

$$\operatorname{Re} b(\overline{u}, \overline{v}) = \frac{h(\overline{u} + \overline{v}) - h(\overline{u}) - h(\overline{v})}{2}$$
$$\operatorname{Im} b(\overline{u}, \overline{v}) = \operatorname{Re}(-ib(\overline{u}, \overline{v})) = \operatorname{Re} b(\overline{u}, i\overline{v})$$

Итак, получено взаимно однозначное соответствие между эрмитовыми квадратичными и эрмитовыми формами. Более того, это соответствие $\mathbb{R}$-линейно. ☐

Следствие. *Эрмитовы и эрмитовы квадратичные формы на V образуют линейные вещественные пространства, изоморфные друг другу.*

Определение 4.22. Пусть b — эрмитова форма. *Ядром* формы b называется подпространство $\operatorname{Ker} b = \{\overline{u} \mid \forall \overline{v} \in V : b(\overline{u}, \overline{v}) = 0\} = \{\overline{v} \mid \forall \overline{u} \in V : b(\overline{u}, \overline{v}) = 0\} \leqslant V$.

Замечание. Аналогично билинейному случаю, $\dim \operatorname{Ker} b = \dim V - \operatorname{rk} b$.

Определение 4.23. Пусть $b \in \mathcal{S}(V)$ — эрмитова форма, $U \leqslant V$. *Ортогональным дополнением к U относительно b* называется $U^\perp = \{\overline{v} \mid \forall \overline{u} \in U : b(\overline{u}, \overline{v}) = 0\}$.

Определение 4.24. Эрмитова форма $b \in \mathcal{S}(V)$ называется *невырожденной*, если $\operatorname{rk} b = \dim V$. Подпространство $U \leqslant V$ называется *невырожденным относительно* b, если ограничение $b|_U \in \mathcal{S}(U)$ невырожденно.

Замечание. Аналогично билинейному случаю, $\dim U^\perp \geqslant \dim V - \dim U$, причем равенство достигается в случае, когда форма b невырожденна. Более того, подпространство $U \leqslant V$ невырожденно относительно $b \Leftrightarrow V = U \oplus U^\perp$.

Утверждение 4.10. *Пусть h — эрмитова квадратичная форма. Тогда в пространстве V существует такой базис e, что h в этом базисе имеет диагональную матрицу с числами 0 и ± 1 на главной диагонали.*

Доказательство. Доказательство аналогично билинейному случаю. □

Определение 4.25. Пусть h — эрмитова квадратичная форма. Тогда h называется:

- ▷ *положительно определенной*, если $\forall \overline{v} \in V, \overline{v} \neq \overline{0} : h(\overline{v}) > 0$.
- ▷ *положительно полуопределенной*, если $\forall \overline{v} \in V : h(\overline{v}) \geqslant 0$.
- ▷ *отрицательно определенной*, если $\forall \overline{v} \in V, \overline{v} \neq \overline{0} : h(\overline{v}) < 0$.
- ▷ *отрицательно полуопределенной*, если $\forall \overline{v} \in V : h(\overline{v}) \leqslant 0$.

Эрмитова форма $b \in \mathcal{S}(V)$, полярная к h, приобретает те же названия.

Замечание. Аналогично билинейному случаю, для эрмитовых квадратичных форм можно определить положительный и отрицательный индексы инерции и доказать аналоги закона инерции, метода Якоби и критерия Сильвестра.

5 Операторы в евклидовых и эрмитовых пространствах

5.1 Евклидовы и эрмитовы пространства

Определение 5.1. *Евклидовым пространством* называется линейное пространство V над $\mathbb{R}$, на котором определена положительно определенная симметрическая билинейная форма $(\overline{u}, \overline{v})$ — *скалярное произведение.*

Определение 5.2. *Эрмитовым пространством* называется линейное пространство V над $\mathbb{C}$, на котором определена положительно определенная эрмитова форма $(\overline{u}, \overline{v})$ — *эрмитово скалярное произведение.*

В данном разделе зафиксируем евклидово (эрмитово) пространство V.

Определение 5.3. *Длиной* вектора $\overline{v} \in V$ называется $||\overline{v}|| := \sqrt{(\overline{v}, \overline{v})}$.

Замечание. $||\overline{v}|| = 0 \Leftrightarrow \overline{v} = \overline{0}$ в силу положительной определенности.

Определение 5.4. Пусть $(\overline{v_1}, \ldots, \overline{v_k})$ — система векторов из V. *Матрицей Грама* этой системы называется следующая матрица:

$$\Gamma = ((\overline{v_i}, \overline{v_j})) = \begin{pmatrix} (\overline{v_1}, \overline{v_1}) & \ldots & (\overline{v_1}, \overline{v_k}) \\ \vdots & \ddots & \vdots \\ (\overline{v_k}, \overline{v_1}) & \ldots & (\overline{v_k}, \overline{v_k}) \end{pmatrix}$$

Замечание. В евклидовом случае матрица Γ симметрична, в эрмитовом — эрмитова. Более того, очевидно, для любых векторов $\overline{u_1} = (\overline{v_1}, \ldots, \overline{v_k})x$, $\overline{u_2} = (\overline{v_1}, \ldots, \overline{v_k})y$ выполнено $(\overline{u_1}, \overline{u_2}) = x^T\Gamma\overline{y}$. В частности, матрица Γ положительно полуопределена.

Теорема 5.1. *Система $(\overline{v_1}, \ldots, \overline{v_k})$ линейно независима $\Leftrightarrow$ ее матрица Грама Γ положительно определена $\Leftrightarrow$ $\det\Gamma > 0$.*

Доказательство. Если система $(\overline{v_1}, \ldots, \overline{v_k})$ линейно зависима, то тогда существует столбец $x \neq \overline{0}$ такой, что $(\overline{v_1}, \ldots, \overline{v_k})x = \overline{0}$. Тогда $\forall i \in \{1, \ldots, k\} : (0, \ldots, 1_i, 0, \ldots, 0)\Gamma x = 0 \Rightarrow$ $\Rightarrow E\Gamma x = \Gamma x = 0 \Rightarrow \Gamma$ вырожденна, откуда $\det\Gamma = 0$ и Γ не положительно определена. Если же система $(\overline{v_1}, \ldots, \overline{v_k})$ линейно независима, то Γ — это матрица ограничения скалярного произведения на $\langle\overline{v_1}, \ldots, \overline{v_k}\rangle$, тогда Γ положительно определена в силу положительной определенности скалярного произведения и, в частности, $\det\Gamma > 0$. $\square$

Теорема 5.2 (неравенство Коши-Буняковского-Шварца)**.** *Для любых векторов $\overline{u}, \overline{v} \in V$ выполнено $|(\overline{u}, \overline{v})| \leqslant ||\overline{u}|| \cdot ||\overline{v}||$, причем равенство достигается тогда и только тогда, когда $\overline{u}$ и $\overline{v}$ коллинеарны.*

Доказательство. Обозначим через Γ матрицу Грама системы векторов $(\overline{u}, \overline{v})$, тогда выполнено $0 \geqslant \det\Gamma = ||\overline{u}||^2 \cdot ||\overline{v}||^2 - |(\overline{u}, \overline{v})|^2$, откуда $|(\overline{u}, \overline{v})| = \sqrt{||\overline{u}||^2 \cdot ||\overline{v}||^2 - \det\Gamma}$, причем $\det\Gamma = 0 \Leftrightarrow \overline{u}$ и $\overline{v}$ коллинеарны. $\square$

Теорема 5.3 (неравенство треугольника)**.** *Для любых векторов $\overline{u}, \overline{v} \in V$ выполнено $||\overline{u} + \overline{v}|| \leqslant ||\overline{u}|| + ||\overline{v}||$.*

Доказательство. Воспользуемся неравенством Коши-Буняковского-Шварца:

$$||\overline{u} + \overline{v}||^2 = (\overline{u} + \overline{v}, \overline{u} + \overline{v}) = ||\overline{u}||^2 + ||\overline{v}||^2 + 2\,\mathrm{Re}\,(\overline{u}, \overline{v}) \leqslant$$
$$\leqslant ||\overline{u}||^2 + ||\overline{v}||^2 + 2|(\overline{u}, \overline{v})| \leqslant ||\overline{u}||^2 + ||\overline{v}||^2 + 2||\overline{u}|| \cdot ||\overline{v}|| = (||\overline{u}|| + ||\overline{v}||)^2$$

Значит, $||\overline{u} + \overline{v}|| \leqslant ||\overline{u}|| + ||\overline{v}||$. $\square$

Замечание. Равенство в теореме выше достигается тогда и только тогда, когда выполнены равенства $\mathrm{Re}\,(\overline{u}, \overline{v}) = |(\overline{u}, \overline{v})| = ||\overline{u}|| \cdot ||\overline{v}||$, то есть $\overline{u}$ и $\overline{v}$ коллинеарны с вещественным коэффициентом.

Определение 5.5 (только для евклидова пространства)**.** Пусть $\overline{u}, \overline{v} \in V \backslash \{\overline{0}\}$. *Углом между векторами $\overline{u}$ и $\overline{v}$* называется величина $\arccos\frac{(\overline{u}, \overline{v})}{||\overline{u}|| \cdot ||\overline{v}||}$.

Определение 5.6. Векторы $\overline{u}, \overline{v} \in V$ называются *ортогональными*, если $(\overline{u}, \overline{v}) = 0$. Обозначение — $\overline{u} \perp \overline{v}$. Система векторов $(\overline{v_1}, \ldots, \overline{v_k})$ из V называется *ортогональной*, если все векторы системы попарно ортогональны, и *ортонормированной*, если она ортогональна и $\forall i \in \{1, \ldots, k\} : ||\overline{v_i}|| = 1$.

Определение 5.7. Подпространства $U_1, U_2 \leqslant V$ называются *ортогональными*, если $\forall \overline{u_1} \in U_1 : \forall \overline{u_2} \in U_2 : \overline{u_1} \perp \overline{u_2}$. Аналогично определяется ортогональность системы подпространств.

Утверждение 5.1. *Пусть $(U_1, \ldots, U_k)$ — ортогональная система подпространств в V. Тогда сумма $\sum_{i=1}^{k} U_i$ — прямая.*

Доказательство. Пусть $\overline{0} = \overline{u_1} + \dots + \overline{u_k}$, где $\forall i \in \{1, \dots, k\} : \overline{u_i} \in U_i$. Достаточно проверить, что $\forall j \in \{1, \dots, k\} : \overline{u_j} = \overline{0}$. Умножая равенство выше скалярно на $\overline{u_j}$, получаем, что $||\overline{u_j}|| = 0$, откуда $\overline{u_j} = \overline{0}$. □

Следствие. *Если $(v_1, \dots, v_k)$ — ортогональная система ненулевых векторов из V, то она линейно независима.*

Утверждение 5.2. *В пространстве V существует ортонормированный базис.*

Доказательство. Приведем скалярное произведение к нормальному виду E. Нормальный базис и будет искомым ортонормированным базисом. □

Определение 5.8. Пусть V_1 и V_2 — евклидовы (эрмитовы) пространства. Отображение $\varphi : V_1 \to V_2$ называется *изоморфизмом евклидовых (эрмитовых) пространств*, если:

1. φ — изоморфизм линейных пространств V_1 и V_2
2. $\forall \overline{u}, \overline{v} \in V_1 : (\overline{u}, \overline{v}) = (\varphi(\overline{u}), \varphi(\overline{v}))$

Теорема 5.4. *Пусть V_1 и V_2 — евклидовы (эрмитовы) пространства. Тогда $V_1 \cong V_2 \Leftrightarrow$ $\Leftrightarrow \dim V_1 = \dim V_2$.*

Доказательство.

- $\Leftarrow$ Пусть e_1, e_2 — ортонормированные базисы в V_1 и V_2, φ — линейное отображение такое, что $\varphi(e_1) = e_2$. Тогда φ — изоморфизм линейных пространств, причем для любых $\overline{u}, \overline{v} \in V_1$, $\overline{u} \leftrightarrow_{e_1} x, \overline{v} \leftrightarrow_{e_1} y$, выполнено $(\overline{u}, \overline{v}) = x^T E \overline{y} = x^T \overline{y} = (\varphi(\overline{u}), \varphi(\overline{v}))$.
- $\Rightarrow$ Поскольку $V_1 \cong V_2$, то они в частности изоморфны как линейные пространства, откуда $\dim V_1 = \dim V_2$. □

Замечание. Если V — евклидово пространство и $\overline{u}, \overline{v} \in V$ образуют линейно независимую систему, то $\langle \overline{u}, \overline{v} \rangle \cong V_2$.

Замечание. Далее в курсе будет показано, что если для отображения $\varphi : V_1 \to V_2$ выполнено $\forall \overline{u}, \overline{v} \in V_1 : (\varphi(\overline{u}), \varphi(\overline{v})) = (\overline{u}, \overline{v})$, то $\varphi \in \mathcal{L}(V_1, V_2)$. Это означает, что для изоморфизма евклидовых пространств достаточно проверять биективность и сохранение скалярного произведения.

Замечание. Скалярное произведение — положительно определенная форма, поэтому любое подпространство $U \leqslant V$ невырожденно относительно скалярного произведения, тогда, по уже доказанной теореме, $V = U \oplus U^\perp$.

Определение 5.9. Пусть $U \leqslant V$, $\overline{v} \in V$. Вектор $\overline{v}$ единственным образом представляется в виде суммы $\overline{u} + \overline{u'}$, $\overline{u} \in U$, $\overline{u'} \in U^\perp$. Вектор $\overline{u}$ называется *ортогональной проекцией $\overline{v}$ на подпространство U*. Обозначение — $\mathrm{pr}_U \overline{v}$.

Утверждение 5.3. *Пусть $e = (\overline{e_1}, \dots, \overline{e_k})$ — это ортогональный базис в U. Тогда для любого вектора $\overline{v} \in V$ выполнено равенство:*

$$\mathrm{pr}_U \overline{v} = \sum_{i=1}^{k} \frac{(\overline{v}, \overline{e_i})}{||\overline{e_i}||^2} \overline{e_i}$$

В частности, если базис e — ортонормированный, то $\mathrm{pr}_U \overline{v} = \sum_{i=1}^{k} (\overline{v}, \overline{e_i}) \overline{e_i}$.

Доказательство. Представим $\overline{v}$ в виде $\overline{v} = \overline{u} + \overline{u'}$, $\overline{u} \in U$, $\overline{u'} \in U^{\perp}$, тогда $\overline{u} = \sum_{i=1}^{k} \alpha_i \overline{e_i}$. Заметим теперь, что $\forall j \in \{1, \ldots, k\} : (\overline{v}, \overline{e_j}) = (\overline{u} + \overline{u'}, \overline{e_j}) = (\overline{u}, \overline{e_j}) = \alpha_j ||\overline{e_j}||^2$, откуда получаем требуемое. $\square$

Утверждение 5.4 (теорема Пифагора). *Если $(\overline{v_1}, \ldots, \overline{v_k})$ — ортогональная система векторов из V, то $||\overline{v_1} + \cdots + \overline{v_k}||^2 = ||\overline{v_1}||^2 + \cdots + ||\overline{v_k}||^2$.*

Доказательство. Раскрывая произведение $(\overline{v_1} + \cdots + \overline{v_k}, \overline{v_1} + \cdots + \overline{v_k})$ по линейности, получаем требуемое. $\square$

Утверждение 5.5. *Пусть $U \leqslant V$, $\overline{v} \in V$, $\overline{u} = \operatorname{pr}_U \overline{v}$. Тогда для любого вектора $\overline{u_1} \in U$, отличного от $\overline{u}$, выполнено $||\overline{v} - \overline{u_1}|| > ||\overline{v} - \overline{u}||$.*

Доказательство. Пусть $\overline{v} = \overline{u} + \overline{u'}, \overline{u'} \in U^{\perp}$. Тогда $\overline{v} - \overline{u_1} = (\overline{v} - \overline{u}) + (\overline{u} - \overline{u_1}) = \overline{u'} + (\overline{u} - \overline{u_1})$. Тогда, по теореме Пифагора, $||\overline{v} - \overline{u_1}||^2 = ||\overline{u'}||^2 + ||\overline{u} - \overline{u_1}||^2 > ||\overline{u'}||^2$. $\square$

Определение 5.10. Пусть V — евклидово (эрмитово) пространство, $U \leqslant V$, $\overline{v} \in V$. Тогда *расстоянием от $\overline{v}$ до U* называется $\rho(\overline{v}, U) := \inf_{\overline{u} \in U} ||\overline{v} - \overline{u}|| = ||\overline{v} - \operatorname{pr}_U \overline{v}|| = ||\operatorname{pr}_{U^{\perp}} \overline{v}||$.

Теорема 5.5 (метод Грама-Шмидта). *Пусть $(\overline{f_1}, \ldots, \overline{f_n})$ — базис в V. Тогда в V существует ортогональный базис $(\overline{e_1}, \ldots, \overline{e_n})$ такой, что $\forall k \in \{1, \ldots, n\} : \langle \overline{e_1}, \ldots, \overline{e_k} \rangle = \langle \overline{f_1}, \ldots, \overline{f_k} \rangle$, причем матрица перехода S — верхнетреугольная с единицами на главной диагонали.*

Доказательство. Положим $\overline{e_1} := \overline{f_1}$ и $\overline{e_k} := \overline{f_i} - \operatorname{pr}_{\langle \overline{f_1}, \ldots, \overline{f_{k-1}} \rangle} \overline{f_k}$ при всех $k \in \{2, \ldots, n\}$. Тогда матрица перехода S — верхнетреугольная с единицами на главной диагонали, поэтому $(\overline{e_1}, \ldots, \overline{e_n})$ является базисом в V. Проверим равенство $\langle \overline{e_1}, \ldots, \overline{e_k} \rangle = \langle \overline{f_1}, \ldots, \overline{f_k} \rangle$ индукцией по k. База, $i = 1$, тривиальна. Пусть теперь $i > 1$, тогда: $\langle \overline{e_1}, \ldots, \overline{e_{k-1}}, \overline{e_k} \rangle = \langle \overline{e_1}, \ldots, \overline{e_{k-1}}, \overline{f_k} \rangle = \langle \overline{f_1}, \ldots, \overline{f_{k-1}}, \overline{f_i} \rangle$. $\square$

Замечание. Получим явную формулу для $\overline{e_k}$ при всех $k \in \{2, \ldots, n\}$:

$$\overline{e_k} = \overline{f_k} - \operatorname{pr}_{\langle \overline{f_1}, \ldots, \overline{f_{k-1}} \rangle} \overline{f_k} = \overline{f_k} - \operatorname{pr}_{\langle \overline{e_1}, \ldots, \overline{e_{k-1}} \rangle} \overline{f_k} = \overline{f_k} - \sum_{j=1}^{k-1} \frac{(\overline{f_k}, \overline{e_j})}{||\overline{e_j}||^2} \overline{e_j}$$

Следствие. *Пусть $(\overline{e_1}, \ldots, \overline{e_k})$ — ортогональная система ненулевых векторов из V. Тогда в V существует ортогональный базис $(\overline{e_1}, \ldots, \overline{e_n}) \supset (\overline{e_1}, \ldots, \overline{e_k})$.*

Доказательство. Дополним систему $(\overline{e_1}, \ldots, \overline{e_k})$ до произвольного базиса и применим метод Грама-Шмидта. Тогда базис станет ортогональным, при этом первые k векторов в нем не изменятся, поскольку $\forall i \in \{1, \ldots, k\} : \overline{e_i} \mapsto \overline{e_i} - \operatorname{pr}_{\langle \overline{e_1}, \ldots, \overline{e_{i-1}} \rangle} \overline{e_i} = \overline{e_i}$. $\square$

Замечание. Из ортогонального базиса $(\overline{e_1}, \ldots, \overline{e_n})$ в V легко получить ортонормированный базис $(\frac{\overline{e_1}}{||\overline{e_1}||}, \ldots, \frac{\overline{e_n}}{||\overline{e_n}||})$.

Определение 5.11. Матрица $S \in M_n(\mathbb{R})$ называется *ортогональной*, если $S^T S = E$. Матрица $S \in M_n(\mathbb{C})$ называется *унитарной*, если $\overline{S^T} S = E$, или $S^T \overline{S} = E$.

Теорема 5.6. *Пусть e — ортонормированный базис в V, $e' = eS$ — произвольный базис в V. Тогда e' — ортонормированный $\Leftrightarrow$ S — ортогональная (унитарная).*

Доказательство. По условию, $\Gamma(e) = E$, и $e' = eS$. Тогда e' — ортонормированный $\Leftrightarrow$ $\Leftrightarrow \Gamma(e') = E \Leftrightarrow S^T\Gamma(e)\overline{S} = E \Leftrightarrow S^T\overline{S} = E$. $\square$

Утверждение 5.6. *Ортогональные (унитарные) матрицы порядка n образуют группу по умножению.*

Доказательство. Достаточно проверить замкнутость соответствующих множеств матриц относительно умножения и взятия обратного элемента. Пусть S, T — ортогональные (унитарные) матрицы порядка n. Рассмотрим e — ортонормированный базис в V, тогда базисы $e' = eS$ и $e'' = e'T$ — тоже ортонормированные. Но $e'' = eST$ и $e = e'S^{-1}$, поэтому матрицы ST и S^{-1} ортогональны (унитарны). $\square$

Определение 5.12. Группа ортогональных матриц порядка n обозначается через $\mathcal{O}(n)$, группа унитарных матриц порядка n — через $\mathcal{U}(n)$.

5.2 Сопряженное пространство

В данном разделе будем считать, что V — евклидово пространство.

Определение 5.13. *Сопряженным к V пространством* называется пространство линейных функционалов на V. Обозначение — V^*.

Теорема 5.7. *Для каждого $\overline{v} \in V$ положим $f_{\overline{v}}(\overline{u}) := (\overline{v}, \overline{u})$. Тогда сопоставление $\overline{v} \mapsto f_{\overline{v}}$ осуществляет изоморфизм между V и V^*.*

Доказательство. Проверим, что заданное сопоставление линейно:

$$f_{\overline{v_1}+\overline{v_2}}(\overline{u}) = (\overline{v_1} + \overline{v_2}, \overline{u}) = (\overline{v_1}, \overline{u}) + (\overline{v_2}, \overline{u}) = f_{\overline{v_1}}(\overline{u}) + f_{\overline{v_2}}(\overline{u})$$
$$f_{\alpha\overline{v}}(\overline{u}) = (\alpha\overline{v}, \overline{u}) = \alpha(\overline{v}, \overline{u}) = \alpha f_{\overline{v}}(\overline{u})$$

Поскольку $\dim V = \dim V^*$ и отображение линейно, то нам достаточно проверить его инъективность, что эквивалентно условию $\forall \overline{v} \in V, \overline{v} \neq \overline{0} : f_{\overline{v}} \neq 0$. Но это условие выполнено в силу положительной определенности скалярного произведения: $\forall \overline{v} \in V, \overline{v} \neq \overline{0} : f_{\overline{v}}(\overline{v}) = = (\overline{v}, \overline{v}) > 0$. $\square$

Следствие. *Пусть $U, W \leqslant V$. Тогда:*

1. $(U^\perp)^\perp = U$
2. $(U + W)^\perp = U^\perp \cap W^\perp$
3. $(U \cap W)^\perp = U^\perp + W^\perp$

Доказательство. Рассмотрим изоморфизм $\Theta : V \to V^*$ из предыдущей теоремы и заметим, что $\Theta(U^\perp) = \{f_{\overline{v}} \in V^* \mid \overline{v} \in U^\perp\} = \{f \in V^* \mid \forall \overline{u} \in U : f(\overline{u}) = 0\} = U^0$ для любого подпространства $U \leqslant V$.

1. Докажем данную формулу непосредственно. Так как $\dim U^\perp = \dim V - \dim U$ и $\dim (U^\perp)^\perp = \dim V - \dim U^\perp$, то $\dim U = \dim (U^\perp)^\perp$. При этом, с другой стороны, $U \leqslant (U^\perp)^\perp$, поскольку $\forall \overline{u} \in U : \forall \overline{v} \in U^\perp : \overline{u} \perp \overline{v}$. Значит, $U = (U^\perp)^\perp$.

2. Поскольку $(U + W)^0 = U^0 \cap W^0$, то, применяя Θ к обеим частям равенства, получим $(U + W)^\perp = U^\perp \cap W^\perp$.

3. Поскольку $(U\cap W)^0 = U^0 + W^0$, то, применяя Θ к обеим частям равенства, получим $(U\cap W)^\perp = U^\perp + W^\perp$. □

Замечание. В эрмитовом пространстве ситуация почти аналогична: если рассмотреть сопоставление $\overline{v} \mapsto f_{\overline{v}}$, то функционал $f_{\overline{v}}$ уже не линеен, а сопряженно-линеен. Если же положить $g_{\overline{v}}(\overline{u}) := (\overline{u}, \overline{v})$, то функционал $g_{\overline{v}}$ линеен, но сопоставление $\overline{v} \mapsto g_{\overline{v}}$ не линейно, а сопряженно-линейно. Оно не является изоморфизмом, но является сопряженно-линейной биекцией, называемой *антиизоморфизмом.*

Легко проверить, что при антиизоморфизме Θ остается справедливым равенство $\Theta(U^\perp) = U^0$. Следовательно, в эрмитовом пространстве V остается справедливым, что для любых подпространств $U, W \leqslant V$ выполнены равенства $(U^\perp)^\perp = U$, $(U+W)^\perp = U^\perp \cap W^\perp$ и $(U\cap W)^\perp = U^\perp + W^\perp$.

5.3 Объем в евклидовых пространствах

В данном разделе будем считать, что V — евклидово пространство.

Определение 5.14. *k-мерным объемом* системы $(\overline{v_1}, \ldots, \overline{v_k})$ векторов из V называется величина $V_k(\overline{v_1}, \ldots, \overline{v_k})$, определяемая индуктивно:

- ▷ Если $k = 1$, то $V_1(\overline{v_1}) := ||\overline{v_1}||$
- ▷ Если $k \geqslant 2$, то $V_k(\overline{v_1}, \ldots, \overline{v_k}) := V_{k-1}(\overline{v_1}, \ldots, \overline{v_{k-1}})\rho(\overline{v_k}, \langle\overline{v_1}, \ldots, \overline{v_{k-1}}\rangle)$

Теорема 5.8. *Пусть $\overline{v_1}, \ldots, \overline{v_k} \in V$, $\Gamma := \Gamma(\overline{v_1}, \ldots, \overline{v_k})$. Тогда $V_k(\overline{v_1}, \ldots, \overline{v_k}) = \sqrt{\det \Gamma}$.*

Доказательство. Если система $(\overline{v_1}, \ldots, \overline{v_k})$ линейно зависима, то $\det\Gamma = 0$, но при этом $\exists i \in \{1, \ldots, k\} : \overline{v_i} \in \langle\overline{v_1}, \ldots, \overline{v_{i-1}}\rangle$, тогда $V_i(\overline{v_1}, \ldots, \overline{v_i}) = 0$, и потому все последующие объемы также равны нулю. Если же система $(\overline{v_1}, \ldots, \overline{v_k})$ линейно независима, то она образует базис в $U := \langle\overline{v_1}, \ldots, \overline{v_k}\rangle$. Применим метод Грама-Шмидта к данному базису и получим ортогональный базис $(\overline{e_1}, \ldots, \overline{e_k})$ такой, что $(\overline{e_1}, \ldots, \overline{e_k}) = (\overline{v_1}, \ldots, \overline{v_k})S$, где матрица перехода S — верхнетреугольная с единицами на главной диагонали. Тогда $\det\Gamma(\overline{e_1}, \ldots, \overline{e_k}) = \det\Gamma(\det S)^2 = \det\Gamma$, откуда:

$$V_k(\overline{e_1}, \ldots, \overline{e_k}) = V_1(\overline{e_1})\rho(\overline{e_2}, \langle\overline{e_1}\rangle)\ldots\rho(\overline{e_k}, \langle\overline{e_1}, \ldots, \overline{e_{k-1}}\rangle) = \\ = ||\overline{e_1}||\ldots||\overline{e_k}|| = \sqrt{\det\Gamma(\overline{e_1}, \ldots, \overline{e_k})} = \sqrt{\det\Gamma}$$

Остается проверить по индукции, что $\forall i \in \{1, \ldots, k\} : V_i(\overline{e_1}, \ldots, \overline{e_i}) = V_i(\overline{v_1}, \ldots, \overline{v_i})$. База, $i = 1$, тривиальна. Пусть теперь $i \geqslant 2$, тогда:

$$V_i(\overline{e_1}, \ldots, \overline{e_i}) = V_{i-1}(\overline{e_1}, \ldots, \overline{e_{i-1}})\rho(\overline{e_i}, \langle\overline{e_1}, \ldots, \overline{e_{i-1}}\rangle) = \\ = V_{i-1}(\overline{v_1}, \ldots, \overline{v_{i-1}})\rho\left(\overline{v_i}, \langle\overline{e_1}, \ldots, \overline{e_{i-1}}\rangle\right) = V_i(\overline{v_1}, \ldots, \overline{v_i})$$

Таким образом, переход доказан, и $V_k(\overline{v_1}, \ldots, \overline{v_k}) = \sqrt{\det\Gamma}$. □

Следствие. *k-мерный объем системы векторов не зависит от перестановки векторов.*

Следствие. *Пусть $\overline{v_1}, \ldots, \overline{v_k} \in V$, $(\overline{u_1}, \ldots, \overline{u_k}) = (\overline{v_1}, \ldots, \overline{v_k})S$. Тогда выполнено равенство $V_k(\overline{u_1}, \ldots, \overline{u_k}) = V_k(\overline{v_1}, \ldots, \overline{v_k})|\det S|$.*

Доказательство. Извлекая из равенства $\det\Gamma(\overline{u_1},\dots,\overline{u_k}) = \det\Gamma(\overline{v_1},\dots,\overline{v_k})(\det S)^2$, получаем требуемое по теореме выше. $\square$

Следствие. *Пусть система $(\overline{e_1},\dots,\overline{e_n})$ векторов из V образует ортонормированный базис в $\langle\overline{e_1},\dots,\overline{e_n}\rangle$, $(\overline{v_1},\dots,\overline{v_n}) = (\overline{e_1},\dots,\overline{e_k})S$. Тогда $V_n(\overline{v_1},\dots,\overline{v_n}) = |\det S|$.*

Следствие. *Пусть система $(\overline{u_1},\dots,\overline{u_k})$ векторов из V линейно независима, $\overline{v}\in V$. Тогда:*

$$\rho(\overline{v},\langle\overline{u_1},\dots,\overline{u_k}\rangle) = \frac{V_{k+1}(\overline{u_1},\dots,\overline{u_k},\overline{v})}{V_k(\overline{u_1},\dots,\overline{u_k})} = \sqrt{\frac{\det\Gamma(\overline{u_1},\dots,\overline{u_k},\overline{v})}{\det\Gamma(\overline{u_1},\dots,\overline{u_k})}}$$

Определение 5.15. Евклидово пространство V называется *ориентированным*, если в нем выделен некоторый ортонормированный базис $(\overline{e_1},\dots,\overline{e_n})$. Тогда базис $(\overline{f_1},\dots,\overline{f_n}) = (\overline{e_1},\dots,\overline{e_n})S$ называется *положительно ориентированным*, если $\det S > 0$, и *отрицательно ориентированным*, если $\det S < 0$. *Ориентированным объемом* системы $(\overline{f_1},\dots,\overline{f_n}) = (\overline{e_1},\dots,\overline{e_n})S$ называется величина $V_n(\overline{v_1},\dots,\overline{v_n})\cdot\operatorname{sgn}\det S$.

5.4 Сопряженные операторы

В данном разделе зафиксируем евклидово (эрмитово) пространство V. Положим $\theta := 2$, если V — евклидово, и $\theta := \frac{3}{2}$, если V — эрмитово. Пространство θ-линейных форм на V обозначим через $\mathcal{B}_\theta(V)$. В частности, скалярное произведение является θ-линейной формой на V.

Определение 5.16. Пусть $\varphi\in\mathcal{L}(V)$. Для всех $\overline{u},\overline{v}\in V$ положим $f_\varphi(\overline{u},\overline{v}) := (\varphi(\overline{u}),\overline{v})$.

Утверждение 5.7. *Пусть e — ортонормированный базис в V, $\varphi\in\mathcal{L}(V)$ — такой оператор, что $\varphi\leftrightarrow_e A$, $f_\varphi\leftrightarrow_e B$. Тогда $B = A^T$.*

Доказательство. Если $\overline{u}\leftrightarrow_e x$, $\overline{v}\leftrightarrow_e y$, то $f_\varphi(\overline{u},\overline{v}) = (Ax)^T\overline{y} = x^TA^T\overline{y}$, что и означает требуемое в силу биективности сопоставления матриц θ-линейным формам. $\square$

Следствие. *Сопоставление $\varphi\mapsto f_\varphi$ осуществляет изоморфизм линейных пространств $\mathcal{L}(V)$ и $\mathcal{B}_\theta(V)$.*

Доказательство. Сопоставление $\varphi\mapsto f_\varphi$ является композицией изоморфизмов соответствующих линейных пространств вида $\varphi\mapsto_e A\mapsto A^T\mapsto_e f_\varphi$. $\square$

Замечание. Пусть $\varphi\in\mathcal{L}(V)$. Аналогичным образом для всех $\overline{u},\overline{v}\in V$ положим $g_\varphi(\overline{u},\overline{v}) := (\overline{u},\varphi(\overline{v}))$, тогда g_φ — θ-линейная форма, как и f_θ. Если $\varphi\leftrightarrow_e A$, то $g_\varphi\leftrightarrow_e \overline{A}$, причем сопоставление $\varphi\mapsto g_\varphi$ является сопряженно-линейной биекцией, то есть в евклидовом пространстве оно осуществляет изоморфизм, а в эрмитовом — антиизоморфизм.

Определение 5.17. Пусть $\varphi\in\mathcal{L}(V)$. Оператором, *сопряженным к φ*, называется оператор $\varphi^*\in\mathcal{L}(V)$ такой, что $f_\varphi = g_{\varphi^*}$, то есть $\forall\overline{u},\overline{v}\in V : (\varphi(\overline{u}),\overline{v}) = (\overline{u},\varphi^*(\overline{v}))$.

Замечание. Поскольку сопоставления $\varphi\mapsto f_\varphi = g_{\varphi^*}\mapsto\varphi^*$ биективны, то сопряженный оператор φ^* существует и единственен. Более того, сопоставление $\varphi\mapsto\varphi^*$ осуществляет автоморфизм в евклидовом случае и антиавтоморфизм в эрмитовом случае.

Утверждение 5.8. *Пусть e — ортонормированный базис в V, $\varphi = \mathcal{L}(V)$ — такой линейный оператор, что $\varphi\leftrightarrow_e A$. Тогда $\varphi^*\leftrightarrow_e A^*$.*

Доказательство. Поскольку $\varphi \leftrightarrow_e A$, то $f_\varphi = g_{\varphi^*} \leftrightarrow_e A^T$. Значит, $\varphi^* \leftrightarrow_e A^*$. □

Замечание. В неортонормированном базисе e формула получается из аналогичных рассуждений, но вычисления несколько усложняются: если $\varphi \leftrightarrow_e A$, то $f_\varphi = g_{\varphi^*} \leftrightarrow_e A^T\Gamma$, тогда $\varphi^* \leftrightarrow_e \overline{\Gamma^{-1}A^T\Gamma}$.

Утверждение 5.9. *Сопряженные операторы обладают следующими свойствами:*

1. *Сопоставление $\varphi \mapsto \varphi^*$ сопряженно-линейно*
2. $\forall \varphi, \psi \in \mathcal{L}(V) : (\varphi\psi)^* = \psi^*\varphi^*$
3. $\forall \varphi \in \mathcal{L}(V) : \varphi^{**} = \varphi$

Доказательство. Первые свойство уже было отмечено, докажем два последних. Зафиксируем произвольные $\overline{u}, \overline{v} \in V$, тогда:

$$((\varphi\psi)(\overline{u}), \overline{v}) = (\psi(\overline{u}), \varphi^*(\overline{v})) = (\overline{u}, (\psi^*\varphi^*)(\overline{v}))$$
$$(\varphi^*(\overline{u}), \overline{v}) = \overline{(\overline{v}, \varphi^*(\overline{u}))} = \overline{(\varphi(\overline{v}), \overline{u})} = (\overline{u}, \varphi(\overline{v}))$$

В силу единственности сопряженного оператора, получено требуемое. □

Утверждение 5.10. *Пусть $\varphi \in \mathcal{L}(V)$. Тогда $\overline{\chi_\varphi(\lambda)} = \chi_{\varphi^*}(\overline{\lambda})$.*

Доказательство. Пусть A — матрица φ в ортонормированном базисе e. Тогда:

$$\overline{\chi_\varphi(\lambda)} = \overline{|A - \lambda E|} = |\overline{A} - \overline{\lambda}E| = \chi_{\overline{A}}(\overline{\lambda}) = \chi_{A^*}(\overline{\lambda}) = \chi_{\varphi^*}(\overline{\lambda})$$ □

Утверждение 5.11. *Пусть $\varphi \in \mathcal{L}(V)$, и подпространство $U \leqslant V$ инвариантно относительно φ. Тогда $U^\perp$ тоже инвариантно относительно φ^*.*

Доказательство. Пусть $\overline{v} \in U^\perp$. Тогда $\forall \overline{u} \in U : (\overline{u}, \varphi^*(\overline{v})) = (\varphi(\overline{u}), \overline{v}) = (\varphi(\overline{u}), \overline{v}) = 0$ в силу инвариантности U. Значит, $\varphi^*(\overline{v}) \in U^\perp$. □

Замечание. В силу канонического изоморфизма между V и V^* и справедливости соответствующего свойства аннуляторных подпространств, $U_1 \leqslant U_2 \Rightarrow U_1^\perp \geqslant U_2^\perp$.

Теорема 5.9 (Фредгольма)**.** *Пусть $\varphi \in \mathcal{L}(V)$. Тогда* $\operatorname{Ker}\varphi^* = (\operatorname{Im}\varphi)^\perp$.

Доказательство.

- $\subset$ Пусть $\overline{v} \in \operatorname{Ker}\varphi^*$, тогда $\varphi^*(\overline{v}) = \overline{0}$, и $\forall \overline{u} \in V : (\varphi(\overline{u}), \overline{v}) = (\overline{u}, \varphi^*(\overline{v})) = 0 \Rightarrow \overline{v} \in (\operatorname{Im}\varphi)^\perp$.
- $\supset$ Заметим, что $\operatorname{rk}\varphi = \operatorname{rk}\varphi^* = \dim\operatorname{Im}\varphi = \dim\operatorname{Im}\varphi^*$, тогда $\dim\operatorname{Ker}\varphi^* = \dim(\operatorname{Im}\varphi)^\perp$, из чего следует требуемое в силу обратного включения. □

Следствие. *Пусть $\varphi \in \mathcal{L}(V)$. Тогда* $\operatorname{Im}\varphi^* = (\operatorname{Ker}\varphi)^\perp$.

Замечание. В общем случае, когда $\varphi \in \mathcal{L}(U, V)$, *сопряженным к φ отображением* называется такое $\varphi^* \in \mathcal{L}(V^*, U^*)$, что $\forall f \in V^*, \forall \overline{u} \in U : \varphi^*(f)(\overline{u}) = f(\varphi(\overline{u}))$. Свойства такого отображения будут похожи на доказанные выше, например, аналог теоремы Фредгольма имеет вид $\operatorname{Ker}\varphi^* = (\operatorname{Im}\varphi)^0$. Отметим, что в основном рассмотренном случае нам удалось избежать перехода в V^* в силу существования канонического изоморфизма (антиизоморфизма) между V и V^*.

Определение 5.18. Оператор $\varphi \in \mathcal{L}(V)$ называется *самосопряженным*, если $\varphi^* = \varphi$, то есть $\forall \overline{u}, \overline{v} \in V : (\varphi(\overline{u}), \overline{v}) = (\overline{u}, \varphi(\overline{v}))$.

Замечание. Если самосопряженный оператор $\varphi \in \mathcal{L}(V)$ в ортонормированном базисе имеет матрицу A, то $A \leftrightarrow_e \varphi = \varphi^* \leftrightarrow_e A^*$, то есть $A = A^*$ — симметрична в евклидовом случае и эрмитова в эрмитовом случае.

Утверждение 5.12. *Пусть $\varphi \in \mathcal{L}(V)$ — самосопряженный, $U \leqslant V$. Тогда U инвариантно относительно $\varphi \Leftrightarrow U^\perp$ инвариантно относительно φ.*

Доказательство.

$\Rightarrow$ Это свойство уже было доказано.

$\Leftarrow$ $(U^\perp)^\perp = U$, поэтому U инвариантно относительно φ. □

Утверждение 5.13. *Пусть $\varphi \in \mathcal{L}(V)$ — самосопряженный. Тогда его характеристический многочлен χ_φ раскладывается на линейные сомножители над $\mathbb{R}$.*

Доказательство. Пусть сначала V — эрмитово пространство, $\lambda \in \mathbb{C}$ — корень χ_φ. Тогда λ является собственным значением оператора φ с собственным вектором $\overline{v} \in V$, $\overline{v} \neq \overline{0}$, откуда $\lambda ||\overline{v}||^2 = (\varphi(\overline{v}), \overline{v}) = (\overline{v}, \varphi(\overline{v})) = \overline{\lambda} ||\overline{v}||^2$. Значит, $\lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}$.

Пусть теперь V — евклидово пространство с ортонормированным базисом e, тогда $\varphi \leftrightarrow_e A \in M_n(\mathbb{R})$, $A = A^T$. Рассмотрим U — эрмитово пространство той же размерности с ортонормированным базиом $\mathcal{F}$ и оператор $\psi \in \mathcal{L}(U)$, $\psi \leftrightarrow_{\mathcal{F}} A$. Тогда ψ — тоже самосопряженный, поэтому для χ_ψ утверждение выполнено. Остается заметить, что $\chi_\psi = \chi_A = \chi_\varphi$. □

Утверждение 5.14. *Пусть $\varphi \in \mathcal{L}(V)$ — самосопряженный, $\lambda_1, \lambda_2 \in \mathbb{R}$ — два различных собственных значения φ. Тогда $V_{\lambda_1} \perp V_{\lambda_2}$.*

Доказательство. Пусть $\overline{v_1} \in V_{\lambda_1}, \overline{v_2} \in V_{\lambda_2}$. Тогда:

$$\lambda_1(\overline{v_1}, \overline{v_2}) = (\varphi(\overline{v_1}), \overline{v_2}) = (\overline{v_1}, \varphi(\overline{v_2})) = \lambda_2(\overline{v_1}, \overline{v_2}) \Rightarrow (\overline{v_1}, \overline{v_2}) = 0$$ □

Теорема 5.10. *Пусть $\varphi \in \mathcal{L}(V)$ — самосопряженный. Тогда в V существует ортонормированный базис e, в котором матрица оператора φ диагональна.*

Доказательство. Проведем индукцию по $n := \dim V$. База, $n = 1$, тривиальна. Пусть теперь $n > 1$. Поскольку корни χ_φ вещественны, то у φ есть собственное значение $\lambda_0 \in \mathbb{R}$. Пусть $\overline{e_0} \in V$ — соответствующий ему собственный вектор длины 1. Тогда подпространство $U := \langle \overline{e_0} \rangle^\perp$ инвариантно относительно φ, поэтому можно рассмотреть оператор $\varphi|_U \in \mathcal{L}(U)$, который также является самосопряженным. По предположению индукции, в U есть ортонормированный базис из собственных векторов, тогда его объединение с $\overline{e_0}$ дает искомый базис в V. □

Замечание.

1. Пусть φ имеет диагональный вид в некотором ортонормированном базисе. Тогда в евклидовом случае φ — самосопряженный, поскольку имеет симметричную матрицу, а в эрмитовом случае φ — самосопряженный $\Leftrightarrow$ все элементы на диагоналы вещественны.

2. Геометрический смысл самосопряженного оператора φ — это композиция растяжений вдоль взаимно ортогональных осей.

3. На практике при диагонализации самосопряженного оператора φ удобнее искать ортонормированные базисы отдельно в каждом его собственном подпространстве.

5.5 Ортогональные операторы

В данном разделе зафиксируем евклидово (эрмитово) пространство V.

Определение 5.19. Оператор $\varphi \in \mathcal{L}(V)$ называется *ортогональным* (*унитарным*), если $\forall \overline{u}, \overline{v} \in V : (\varphi(\overline{u}), \varphi(\overline{v})) = (\overline{u}, \overline{v})$.

Замечание. Условие в предыдущем определении — это равенство двух θ-линейных форм, симметрических в евклидовом случае и эрмитовых — в эрмитовом. Они однозначно задаются соответствующими им квадратичными формами, поэтому равенство достаточно проверять лишь при $\overline{u} = \overline{v}$, тогда оно принимает вид $\forall \overline{v} \in V : ||\varphi(\overline{v})|| = ||\overline{v}||$.

Утверждение 5.15. *Пусть преобразование $\varphi : V \to V$ евклидова пространства V ортогонально, то есть $\forall \overline{u}, \overline{v} \in V : (\varphi(\overline{u}), \varphi(\overline{v})) = (\overline{u}, \overline{v})$. Тогда $\varphi \in \mathcal{L}(V)$.*

Доказательство. Пусть $\overline{u}, \overline{v} \in V$, $\alpha \in \mathbb{R}$. Пользуясь линейностью скалярного произведения и ортогональностью преобразования φ, получим:

$$||\varphi(\overline{u} + \overline{v}) - \varphi(\overline{u}) - \varphi(\overline{v})||^2 = ||(\overline{u} + \overline{v}) - \overline{u} - \overline{v}||^2 = 0$$
$$||\varphi(\alpha\overline{u}) - \alpha\varphi(\overline{u})||^2 = ||\alpha\overline{u} - \alpha\overline{u}||^2 = 0$$

Значит, $\varphi(\overline{u} + \overline{v}) = \varphi(\overline{u}) + \varphi(\overline{v})$ и $\varphi(\alpha\overline{u}) = \alpha\varphi(\overline{u})$. Таким образом, $\varphi \in \mathcal{L}(V)$. $\square$

Теорема 5.11. *Пусть $\varphi \in \mathcal{L}(V)$. Тогда оператор φ ортогонален (унитарен) $\Leftrightarrow$ φ обратим и $\varphi^{-1} = \varphi^*$.*

Доказательство. По определению, φ ортогональнен (унитарен) $\Leftrightarrow$ для любых векторов $\overline{u}, \overline{v} \in V$ выполнено $(\overline{u}, \overline{v}) = (\varphi(\overline{u}), \varphi(\overline{v})) = (\overline{u}, (\varphi^*\varphi)(\overline{v}))$. В силу единственности сопряженного оператора, это равносильно равенству $\varphi^*\varphi = \mathrm{id}^* = \mathrm{id}$. Это, в свою очередь, равносильно тому, что φ обратим и $\varphi^{-1} = \varphi^*$. $\square$

Замечание. В бесконечномерном случае из того, что $\varphi^*\varphi = \mathrm{id}$, не следует, что φ обратим, поэтому тогда в определении ортогонального оператора приходится добавлять, что он обязательно биективен.

Следствие. *Пусть $\varphi \in \mathcal{L}(V)$, e — ортонормированный базис в V*

1. *Если $\varphi \leftrightarrow_e A$, то оператор φ ортогонален (унитарен) $\Leftrightarrow \exists A^{-1} = A^*$, то есть A ортогональна (унитарна).*

2. *Оператор φ ортогонален (унитарен) $\Leftrightarrow$ $\varphi(e)$ — ортонормированный базис.*

Доказательство.

1. Поскольку $\varphi^* \leftrightarrow_e A^*$, то $\exists \varphi^{-1} = \varphi^* \Leftrightarrow \exists A^{-1} = A^*$.

2. Пусть $\varphi \leftrightarrow_e A$, тогда оператор φ ортогонален (унитарен) $\Leftrightarrow$ матрица A ортогональна (унитарна) $\Leftrightarrow \varphi(e) = eA$ — ортонормированный базис. $\square$

Следствие. *Пусть $\varphi \in \mathcal{L}(V)$ — ортогональный (унитарный). Тогда $|\det\varphi| = 1$.*

Доказательство. Зафиксируем ортонормированный базис e в V. Тогда если $\varphi \leftrightarrow_e A$, то $AA^* = E$. Значит, $|\det A|^2 = \det A\overline{\det A} = \det A \det A^* = 1 \Rightarrow |\det A| = 1$. $\square$

Утверждение 5.16. *Пусть $\varphi \in \mathcal{L}(V)$ — ортогональный (унитарный), $U \leqslant V$. Тогда U инвариантно относительно $\varphi \Leftrightarrow U^\perp$ инвариантно относительно φ.*

Доказательство. Поскольку $(U^\perp)^\perp = U$, то достаточно доказать импликацию $\Rightarrow$. Так как U инвариантно относительно φ, то $U^\perp$ инвариантно относительно $\varphi^* = \varphi^{-1}$, то есть $\varphi^{-1}(U^\perp) \leqslant U^\perp$. Но оператор φ биективен, поэтому $\varphi^{-1}(U^\perp) = U^\perp$ и $\varphi(U^\perp) = U^\perp$, откуда $U^\perp$ инвариантно относительно φ. □

Теорема 5.12. *Пусть V — эрмитово пространство, $\varphi \in \mathcal{L}(V)$ — унитарный. Тогда в V существует ортонормированный базис e, в котором матрица оператора φ диагональна с числами модуля 1 на главной диагонали.*

Доказательство. Докажем диагонализуемость оператора φ в ортонормированном базисе индукцией по $n = \dim V$. База, $n = 1$, тривиальна. Пусть теперь $n > 1$. Поскольку у χ_φ есть корень над $\mathbb{C}$, то у φ есть собственный вектор $\overline{e_0}$ длины 1. Тогда $U := \langle \overline{e_0} \rangle^\perp$ инвариантно относительно φ, поэтому можно расмотреть оператор $\varphi|_U \in \mathcal{L}(V)$, который также является унитарным. По предположению индукции, в U есть ортонормированный базис из собственных векторов, тогда объединение с $\overline{e_0}$ дает искомый базис в V.

Покажем теперь, что все собственные значения оператора φ имеют модуль 1. Действительно, если $\overline{v} \in V$, $\overline{v} \neq 0$ — собственный вектор со значением λ, то $(\overline{v}, \overline{v}) = (\varphi(\overline{v}), \varphi(\overline{v})) = |\lambda|^2(\overline{v}, \overline{v}) \Rightarrow |\lambda| = 1$. □

Замечание. Пусть, напротив, e — ортонормированный базис в V, $\varphi \in \mathcal{L}(V)$, и $\varphi \leftrightarrow_e A$, где A диагональна с числами модуля 1 на диагонали. Тогда $A^*A = E$, поэтому φ унитарен.

Утверждение 5.17. *Пусть V — линейное пространство над $\mathbb{R}$, $\dim V \geqslant 1$, $\varphi \in \mathcal{L}(V)$. Тогда у φ существует одномерное или двумерное инвариантное подпространство.*

Доказательство. По основной теореме алгебры, минимальный многочлен μ_φ имеет следующий вид:

$$\mu_\varphi(x) = \prod_{i=1}^{k}(x - \alpha_i)\prod_{j=1}^{m}(x^2 + \beta_j x + \gamma_j)$$

Поскольку $\mu_\varphi(\varphi) = 0$, то хотя бы один из операторов $\varphi - \alpha_i$, $\varphi^2 + \beta_j\varphi + \gamma_j$ — вырожденный. Более того, все они вырожденные в силу минимальности многочлена μ_φ. Значит, возможны два случая:

1. Если $\varphi - \alpha$ — вырожденный, то $\exists \overline{v} \in V$, $\overline{v} \neq \overline{0}$ — собственный вектор с собственным значением α, и $\langle \overline{v} \rangle \leqslant V$ — искомое подпространство.

2. Если $\varphi^2 + \beta\varphi + \gamma$ — вырожденный, то $\exists \overline{v} \in V$, $\overline{v} \neq \overline{0} : (\varphi^2 + \beta\varphi + \gamma)(\overline{v}) = \overline{0}$. Поскольку $\varphi^2(\overline{v}) = -\beta\varphi(\overline{v}) - \gamma\overline{v}$, то $\langle \overline{v}, \varphi(\overline{v}) \rangle \leqslant V$ — искомое подпространство. □

Теорема 5.13. *Пусть V — евклидово пространство, $\varphi \in \mathcal{L}(V)$ — ортогональный. Тогда в V существует ортонормированный базис e, в котором матрица оператора φ имеет следующий вид:*

$$\varphi \leftrightarrow_e A = \begin{pmatrix} \boxed{B_1} & 0 & \ldots & 0 \\ 0 & \boxed{B_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \boxed{B_m} \end{pmatrix},$$

где $\forall i \in \{1, \ldots, m\} : B_i = (\pm 1)$ или $B_i = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, $\alpha \in [0, 2\pi)$.

Доказательство. Проведем индукцию по $n := \dim V$. База, $n = 0$, тривиальна. Пусть теперь $n \geqslant 1$. Выберем $U \leqslant V$ — одномерное или двумерное инвариантное относительно φ подпространство. Тогда $U^\perp$ тоже инвариантно относительно φ, и в $U^\perp$ есть требуемый базис e' по предположению индукции. Если e'' — некоторый ортонормированный базис в U, то его объединение с e' дает ортонормированный базис в V. Исследуем оператор $\varphi|_U$:

1. Если $\dim U = 1$, то $U = \langle \overline{v} \rangle$, где $\overline{v} \in V$ — собственный вектор длины 1 с собственным значением $\lambda \in \mathbb{R}$, и $\lambda = \pm 1$ аналогично комплексному случаю.

2. Если $\dim U = 2$, то $\varphi_U \leftrightarrow_{e''} C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Тогда:

$$C^T C = E \Leftrightarrow \begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ ab + cd = 0 \end{cases}$$

Выберем $\alpha, \beta \in [0, 2\pi)$ такие, что $a = \cos\alpha$, $b = -\sin\beta$, $c = \sin\alpha$, $d = \cos\beta$, тогда $\sin(\alpha - \beta) = 0$. Значит, либо $\alpha = \beta$, либо $\alpha = \beta \pm \pi$. В первом случае уже получено требуемое, во втором — матрица C имеет следующий вид:

$$C = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

В этом случае $\chi_C(\lambda) = \lambda^2 - 1$, поэтому $U = \langle \overline{u}, \overline{v} \rangle$, где $\overline{u}, \overline{w} \in V$ — собственные векторы длины 1 с собственными значениями 1 и -1. Заметим теперь, что $(\overline{u}, \overline{w}) = (\varphi(\overline{u}), \varphi(\overline{w})) = (\overline{u}, -\overline{w}) = -(\overline{u}, \overline{w})$, откуда $(\overline{u}, \overline{w}) = 0$. □

Замечание. Теорема также утверждает, что геометрический смысл ортогонального оператора — это композиция отражений и двумерных поворотов.

Замечание. Аналогично комплексному случаю, любой оператор, имеющий в ортонормированном базисе матрицу описанного выше вида, ортогонален.

Замечание. Назовем оператор $\varphi \in \mathcal{L}(V)$ *нормальным*, если $\varphi^*\varphi = \varphi\varphi^*$. Тогда аналогичным рассуждением можно показать, что в эрмитовом пространстве V любой нормальный оператор приводится к диагональному виду в ортонормированном базисе, а также что любой диагонализуемый в ортонормированном базисе оператор нормален.

Теорема 5.14. *Пусть $\varphi \in \mathcal{L}(V)$. Тогда существуют $\psi, \Theta \in \mathcal{L}(V)$ такие, что ψ — самосопряженный с неотрицательными собственными значениями, Θ — ортогональный (унитарный), и $\varphi = \psi\Theta$.*

Доказательство. Рассмотрим оператор $\eta := \varphi^*\varphi$, тогда $\eta^* = \varphi^*\varphi = \eta$, то есть η — самосопряженный. Более того, если $\overline{v} \in V \backslash \{\overline{0}\}$ — собственный вектор оператора η с собственным значением $\lambda \in \mathbb{R}$, то $\eta(\overline{v}) = \lambda\overline{v}$, тогда $0 \leqslant (\varphi(\overline{v}), \varphi(\overline{v})) = (\overline{v}, \eta(\overline{v})) = \lambda(\overline{v}, \overline{v}) \Rightarrow \lambda \geqslant 0$.

Пусть $(\overline{e_1}, \ldots, \overline{e_n})$ — ортонормированный базис в V из собственных векторов оператора η с собственными значениями $\lambda_1, \ldots, \lambda_n \geqslant 0$. Положим $\overline{f_i} := \varphi(\overline{e_i})$, $i \in \{1, \ldots, n\}$. Тогда для любых $i, j \in \{1, \ldots, n\}$ выполнено $(\overline{f_i}, \overline{f_j}) = (\varphi(\overline{e_i}), \varphi(\overline{e_j})) = (\overline{e_i}, \eta(\overline{e_j})) = \lambda_j(\overline{e_i}, \overline{e_j})$. Значит, система $(\overline{f_1}, \ldots, \overline{f_n})$ ортогональна, и, более того, для любого $i \in \{1, \ldots, n\}$ выполнено $||\overline{f_i}||^2 = \lambda_i ||\overline{e_i}||^2 = \lambda_i$.

Будем без ограничения общности считать, что $\lambda_1, \ldots, \lambda_k > 0$ и $\lambda_{k+1} = \cdots = \lambda_n = 0$. Положим $\overline{g_i} := \frac{1}{\sqrt{\lambda_i}}\overline{f_i}$, $i \in \{1, \ldots, k\}$, и дополним $(\overline{g_1}, \ldots, \overline{g_k})$ до ортонормированного базиса $(\overline{g_1}, \ldots, \overline{g_n})$. Тогда оператор φ имеет следующий вид: $\overline{e_i} \mapsto \overline{g_i} \mapsto \sqrt{\lambda_i}\overline{g_i} = \overline{f_i}$. Зададим $\psi, \Theta \in \mathcal{L}(V)$ на базисах $(\overline{e_1}, \ldots, \overline{e_n})$ и $(\overline{g_1}, \ldots, \overline{g_n})$ следующим образом:

$$\Theta : \overline{e_i} \mapsto \overline{g_i}$$
$$\psi : \overline{g_i} \mapsto \sqrt{\lambda_i}\overline{g_i} = \overline{f_i}$$

Таким образом, $\psi\Theta = \varphi$. Наконец, Θ переводит ортонормированный базис $(\overline{e_1}, \ldots, \overline{e_n})$ в ортонормированный базис $(\overline{g_1}, \ldots, \overline{g_n})$, поэтому Θ — ортогональный (унитарный), а ψ имеет в ортонормированном базисе $(\overline{g_1}, \ldots, \overline{g_n})$ диагональный вид, поэтому ψ — самосопряженный. $\square$

Замечание. Порядок операторов в композиции несущественен: если $\varphi = \psi\Theta$, то $\varphi^* = \Theta^*\psi^* = \Theta^{-1}\psi$ — теперь ортогональный (унитарный) оператор Θ^{-1} идет перед самосопряженным оператором ψ.

Определение 5.20. Представление $\varphi \in \mathcal{L}(V)$ в виде $\psi\Theta$ (или в виде $\Theta'\psi'$) с соответствующими требованиями из теоремы выше называется *полярным разложением* φ, а базисы $(\overline{e_1}, \ldots, \overline{e_n})$ и $(\overline{g_1}, \ldots, \overline{g_n})$ из доказательства теоремы — *сингулярными базисами* φ, причем эти базисы одинаковы в случаях $\psi\Theta$ и $\Theta'\psi'$.

Замечание. Геометрический смысл полярного разложения — представление оператора φ в виде композиции движения Θ и растяжения ψ (с неотрицательными коэффициентами) вдоль нескольких взаимно ортогональных осей.

Замечание. Можно показать, что если оператор $\varphi \in \mathcal{L}(V)$ — невырожденный, то полярное разложение φ единственно.

5.6 Приведение к главным осям

В евклидовом случае речь пойдет о квадратичных формах, в эрмитовом — об эрмитовых квадратичных формах. В обоих случаях множество квадратичных форм будем обозначать через $\mathcal{Q}(V)$.

Теорема 5.15 (о приведении к главным осям)**.** *Пусть V — евклидово (эрмитово) пространство, $q \in \mathcal{Q}(V)$. Тогда в V существует ортонормированный базис e, в котором q имеет диагональный вид.*

Доказательство. Пусть $b \in \mathcal{B}^+(V)$ — θ-линейная форма, полярная к q. Тогда $\exists \varphi \in \mathcal{L}(V)$ такой, что $b(\overline{u}, \overline{v}) = (\varphi(\overline{u}), \overline{v})$. При этом:

$$(\varphi(\overline{u}), \overline{v}) = b(\overline{u}, \overline{v}) = \overline{b(\overline{v}, \overline{u})} = \overline{(\varphi(\overline{v}), \overline{u})} = (\overline{u}, \varphi(\overline{v}))$$

Значит, φ — самосопряженный, и в V существует ортонормированный базис e, в котором φ диагонализуем. Тогда если $\varphi \leftrightarrow_e A$, то $b \leftrightarrow_e A^T$ и $q \leftrightarrow_e A^T$, поэтому форма q тоже имеет диагональную матрицу в базисе e. $\square$

Замечание. Напротив, если в ортонормированном базисе e матрица формы q диагональна, то и матрица оператора φ диагональна и, следовательно, задана однозначно собственными значениями φ. Значит, диагональный вид q в ортонормированном базисе определен однозначно.

Теорема 5.16. *Пусть V — линейное пространство над $\mathbb{R}$ (над $\mathbb{C}$), $q_1, q_2 \in \mathcal{Q}(V)$, и q_2 положительно определена. Тогда в V существует такой базис e, в котором матрицы форм q_1 и q_2 диагональны.*

Доказательство. Пусть b — θ-линейная форма, полярная к q_2. Тогда b можно объявить b скалярным (эрмитовым скалярным) произведением на V. В полученном евклидовом (эрмитовом) пространстве форма q_1 приводится к главным осям в некотором ортонормированном базисе e. Поскольку базис e — ортонормированный, то в этом же базисе q_2 имеет диагональный вид E. □

Замечание. Требование положительной определенности в теореме существенно. Зафиксируем некоторый базис $e := (\overline{e_1}, \overline{e_2})$ в пространстве V над $\mathbb{R}$ (над $\mathbb{C}$) и рассмотрим формы $q_1, q_2 \in \mathcal{Q}(V)$ следующего вида:

$$q_1 \leftrightarrow_e \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \; q_2 \leftrightarrow_e \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Предположим, что в пространстве V существует такой базис $\mathcal{F}$, в котором обе формы имеют диагональный вид, то есть для некоторых вещественных чисел $\lambda_1, \mu_1, \lambda_2, \mu_2$ выполнено следующее:

$$q_1 \leftrightarrow_{\mathcal{F}} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \; q_2 \leftrightarrow_{\mathcal{F}} \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}$$

Тогда их линейная комбинация $\mu_1 q_1 - \lambda_1 q_2$ вырожденна, но при подсчете в базисе e определитель матрицы формы $\mu_1 q_1 - \lambda_1 q_2$ равен $-\mu_1^2 - \lambda_1^2$. Но обе формы q_1, q_2 невырожденны, поэтому μ_1, λ_1 отличны от нуля и $-\mu_1^2 - \lambda_1^2 \neq 0$ — противоречие.

Замечание. В первом семестре мы уже использовали приведение к главным осям в маломерном случае. Действительно, рассмотрим уравнение кривой второго порядка:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

Здесь выражение $Ax^2 + 2Bxy + Cy^2$ в базисе e задает следующую форму q:

$$q \leftrightarrow_e \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

Для дальнейшего приведения уравнения к каноническому виду мы приводили форму q к виду $A'x'^2 + C'y'^2$ в базисе e'. По только что доказанной теореме, тот же метод работает в многомерном случае. Рассмотрим уравнение следующего вида:

$$\sum_{1 \leqslant i,j \leqslant n} b_{ij} x_i x_j + \sum_{1 \leqslant i \leqslant n} c_i x_i + d = 0$$

Зафиксируем базис e в n-мерном евклидовом (эрмитовом) пространстве V. Пусть квадратичная форма h в базисе e имеет вид $h(\overline{x}) = \sum_{1 \leqslant i,j \leqslant n} b_{ij} x_i x_j$. Тогда h можно привести к главным осям, и в некотором базисе e' она примет вид $\sum_{1 \leqslant i \leqslant n} b'_i x_i'^2$. Тогда исходное уравнение примет следующий вид:

$$\sum_{1 \leqslant i \leqslant n} b'_i x_i'^2 + \sum_{1 \leqslant i \leqslant n} c'_i x'_i + d = 0$$

Далее можно применить аналогичную двумерному случаю процедуру избавления от линейных членов $c_i x'_i$, если соответствующие квадратичные члены $b'_i x'^2_i$ отличны от нуля, и затем привести уравнение к окончательному виду с не более чем одним линейным членом.

6 Тензоры

6.1 Тензор и его координатная запись

В данном разделе зафиксируем линейное пространство V размерности n над полем F и сопряженное к нему пространство $V^* = \mathcal{L}(V, F)$. Для любых $s, t \in \mathbb{N} \cup \{0\}$ положим $V^s := \underbrace{V \times \dots \times V}_{s}$, $(V^*)^t := \underbrace{V^* \times \dots \times V^*}_{t}$.

Определение 6.1. *Тензором типа* (p, q), или *p раз контравариантным и q раз ковариантным тензором* называется полилинейное отображение $t : (V^*)^p \times V^q \to F$. Все тензоры типа (p, q) образуют линейное пространство над F, обозначение — $\mathbb{T}^p_q(V)$ или $\mathcal{L}(\underbrace{V^*, \dots, V^*}_{p}, \underbrace{V, \dots, V}_{q}; F)$.

Замечание. Тензор задается однозначно своими значениями на всевозможных комбинациях аргументов из базиса в V и базиса в V^*, то есть на n^{p+q} наборах векторов.

Пример. Рассмотрим несколько тензоров различных типов:

1. Тензор типа $(0, 1)$ — это линейный функционал на V, поэтому $\mathbb{T}^0_1 = V^*$.

2. Тензор типа $(1, 0)$ — это элемент пространства $V^{**} \cong V$, поэтому $\mathbb{T}^1_0 \cong V$, причем эти пространства можно отождествить в силу канонического изоморфизма.

3. Тензор типа $(0, 2)$ — это билинейная форма на V, поэтому $\mathbb{T}^0_2 = \mathcal{B}(V)$.

4. Тензор типа $(1, 1)$ — это билинейное отображение $t : V^* \times V \to F$. Зафиксируем $\overline{v} \in V$, тогда $t_{\overline{v}}(f) := t(f, \overline{v})$ — линейный функционал на V^*, то есть $t_{\overline{v}} = \overline{u} \in V$. Тензору t можно поставить в соответствие линейный оператор $\varphi \in \mathcal{L}(V)$, $\varphi(\overline{v}) = t_{\overline{v}} = \overline{u}$. Это соответствие линейно, поскольку t линеен по второму аргументу, и обратимо: $\forall \varphi \in \mathcal{L}(V) : \varphi \mapsto t$, где $t \in \mathbb{T}^1_1$ — такой, что $t(f, \overline{v}) = f(\varphi(\overline{v}))$. Значит, $\mathbb{T}^1_1 \cong \mathcal{L}(V)$, причем эти пространства можно отождествить в силу канонического изоморфизма.

5. Пусть A — алгебра над F. Тогда умножение $\cdot : A \times A \to A$ — это билинейное отображение, $\cdot \in \mathcal{L}(A, A; A)$, и ему соответствует тензор $t \in \mathbb{T}^1_2(A)$ следующего вида:

$$t(f, \overline{a_1}, \overline{a_2}) := f(\overline{a_1} \cdot \overline{a_2})$$

   Аналогично прошлому примеру, соответствие $\cdot \mapsto t$ линейно и обратимо, поэтому $\mathbb{T}^1_2(A) \cong \mathcal{L}(A \times A, A)$, причем эти пространства можно отождествить в силу канонического изоморфизма.

6. Один из тензоров типа $(0, n)$, $n \in \mathbb{N}$, — это определитель.

Определение 6.2. Пусть $t \in \mathbb{T}^p_q$, $t' \in \mathbb{T}^{p'}_{q'}$. Тогда *тензорным произведением* тензоров t и t' называется тензор $t \otimes t' \in \mathbb{T}^{p+p'}_{q+q'}$ следующего вида:

$$t \otimes t'(f_1, \ldots, f_{p+p'}, \overline{v_1}, \ldots, \overline{v_{q+q'}}) := t(f_1, \ldots, f_p, \overline{v_1}, \ldots, \overline{v_q})t'(f_{p+1}, \ldots, f_{p+p'}, \overline{v_{q+1}}, \ldots, \overline{v_{q+q'}})$$

Пример. Рассмотрим несколько тензорных произведений:

1. Пусть $f_1, f_2 \in \mathbb{T}^0_1 = V^*$. Тогда $f_1 \otimes f_2 \in \mathbb{T}^0_2 = \mathcal{B}(V)$, причем $f_1 \otimes f_2(\overline{v_1}, \overline{v_2}) = f_1(\overline{v_1})f_2(\overline{v_2})$ и легко видеть, что $\operatorname{rk} f_1 \otimes f_2 \leqslant 1$.

2. Пусть $g \in V^*$, $\overline{u} \in V$. Тогда $g \otimes \overline{u} \in \mathbb{T}^1_1 = \mathcal{L}(V)$, и данному тензору соответствует оператор $\varphi \in \mathcal{L}(V)$ такой, что $\varphi(\overline{v}) = g(\overline{v})\overline{u}$. В частности, $\operatorname{rk} \varphi \leqslant 1$.

Утверждение 6.1. *Тензорное произведение обладает следующими свойствами:*

1. $\otimes$ *линейно по обоим аргументам.*

2. $\otimes$ *ассоциативно, но необязательно коммутативно.*

Доказательство. Оба свойства следуют непосредственно из формулы в определении тензорного произведения. В то же время, если, например, $t_1, t_2 \in T^0_1$, то:

$$t_1 \otimes t_2(\overline{v_1}, \overline{v_2}) = t_1(\overline{v_1})t_2(\overline{v_2})$$
$$t_2 \otimes t_1(\overline{v_1}, \overline{v_2}) = t_2(\overline{v_1})t_1(\overline{v_2})$$

Видно, что при $\dim V > 0$ можно подобрать такие тензоры и такие векторы, на которых значения выражений выше будут отличаться. $\square$

Замечание. Далее в записях будут применяться нижние и верхние индексы, не означающие возведение в степень. Они нужны исключительно для упрощения формул.

Определение 6.3. Пусть $e = (e_1, \ldots, e_n)$ — базис в V. *Взаимным (биортогональным)* к e базисом называется базис $e^* = (e^1, \ldots, e^n)$ в V^* такой, что:

$$\forall i, j \in \{1, \ldots, n\} : e^j(e_i) = e_i(e^j) = \delta^j_i = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$$

Будем обозначать через $v^j := e^j(\overline{v})$ j-ую координату вектора $\overline{v}$ в базисе e, а через $f_i := e_i(f)$ — i-ую координату функционала f в базисе e^*.

Замечание (соглашение Эйнштейна)**.** Если в некотором выражении встречается один и тот же индекс сверху и снизу, будем считать, что по этому индексу происходит суммирование, как в примере ниже:

$$\forall \overline{v} \in V : \overline{v} = \sum_{i=1}^{n} v^i e_i =: v^i e_i$$

Определение 6.4. Пусть e и e^* — взаимные базисы в V и V^*, $t \in \mathbb{T}^p_q$. *Координатами тензора t в базисе e* называется набор из следующих величин:

$$t^{i_1, \ldots, i_p}_{j_1, \ldots, j_q} = t(e^{i_1}, \ldots, e^{i_p}, e_{j_1}, \ldots, e_{j_q}),\ i_1, \ldots, i_p, j_1, \ldots, j_q \in \{1, \ldots, n\}$$

Замечание. Как уже было отмечено, тензор $t \in \mathbb{T}^p_q$ однозначно задается своими координатами в неотором базисе. Заметим, что тензор $t = e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q} \in \mathbb{T}^p_q$ имеет координаты следующего вида:

$$t^{i'_1,\dots,i'_p}_{j'_1,\dots,j'_q} = \delta^{i'_1}_{i_1} \dots \delta^{i'_p}_{i_p} \delta^{j_1}_{j'_1} \dots \delta^{j_q}_{j'_q}$$

Значит, произвольный тензор $t \in \mathbb{T}^p_q$ можно записать в таком виде:

$$t = t^{i_1,\dots,i_p}_{j_1,\dots,j_q} e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q}$$

Равенство выше справедливо потому, что значения тензоров в левой и правой части совпадают на всех наборах вида $(e^{i_1}, \dots, e^{i_p}, e_{j_1}, \dots, e_{j_q})$.

Замечание. Как уже было отмечено, тензоры вида $e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q}$ — это порождающая система в $\mathbb{T}^p_q$. Более того, она линейно независима, поскольку для каждого тензора вида $e_{i_1} \otimes \dots \otimes e_{i_p} \otimes e^{j_1} \otimes \dots \otimes e^{j_q}$ можно выбрать такой набор $(e^{i_1}, \dots, e^{i_p}, e_{j_1}, \dots, e_{j_q})$, который обнулит все тензоры системы кроме данного. Значит, эта система образует базис в пространстве $\mathbb{T}^p_q$.

Пример. Пусть $t \in \mathbb{T}^1_2(V)$, e и e^* — взаимные базисы в V и V^*, $f \in V^*$, $\overline{u}, \overline{v} \in V$. Координаты тензора t — это набор величин вида $t^i_{jk} = t(e^i, e_j, e_k)$, тогда:

$$t(f, \overline{u}, \overline{v}) = t(f_i e^i, u^j e_j, v^k e_k) = t^i_{jk} f_i u^j v^k$$

Пример. Пусть e, e^* и e', e'^* — две пары взаимных базисов в V и V^* таких, что $e'_j = a^i_j e_i$, $e'^k = b^k_i e^i$, то есть выполнены следующие равенства:

$$e' = (e'_1, e'_2, \dots, e'_n) = (e_1, e_2, \dots, e_n) \begin{pmatrix} a^1_1 & a^1_2 & \dots & a^1_n \\ a^2_1 & a^2_2 & \dots & a^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a^n_1 & a^n_2 & \dots & a^n_n \end{pmatrix} =: eS$$

$$e'^* = \begin{pmatrix} e'^1 \\ e'^2 \\ \vdots \\ e'^n \end{pmatrix} = \begin{pmatrix} b^1_1 & b^1_2 & \dots & b^1_n \\ b^2_1 & b^2_2 & \dots & b^2_n \\ \vdots & \vdots & \ddots & \vdots \\ b^n_1 & b^n_2 & \dots & b^n_n \end{pmatrix} \begin{pmatrix} e^1 \\ e^2 \\ \vdots \\ e^n \end{pmatrix} =: Te^*$$

Тогда $\forall j, k \in \{1, \dots, n\} : \delta^k_j = e'^k(e'_j) = b^k_i e^i(a^l_j e_l) = b^k_i a^l_j \delta^i_l = b^k_i a^i_j$, откуда $ST = E$. Следовательно, $e^* = T^{-1} e'^* = S e'^*$, то есть $e^k = a^k_i e'^i$.

Теорема 6.1. *Пусть e, e' — базисы в V такие, что $e'_j = a^i_j e_i$, $e^k = a^k_i e'^i$. Тогда преобразование координат тензора $t \in \mathbb{T}^p_q$ при замене базиса имеет следующий вид:*

$$t^{i_1,\dots,i_p}_{j_1,\dots,j_q} = a^{i_1}_{i'_1} \dots a^{i_p}_{i'_p} b^{j'_1}_{j_1} \dots b^{j'_q}_{j_q} t'^{i'_1,\dots,i'_p}_{j'_1,\dots,j'_q}$$

Доказательство. Для простоты выполним проверку в случае, когда $t \in \mathbb{T}^1_1$, поскольку в общем случае рассуждение аналогично:

$$t = t^i_j e_i \otimes e^j = t^{i'}_{j'} e'_{i'} \otimes e'^{j'} = t^{i'}_{j'} (a^i_{i'} e_i) \otimes (b^{j'}_j e^j) = t^{i'}_{j'} a^i_{i'} b^{j'}_j e_i \otimes e^j$$

Получено разложение тензора t по базису e двумя способами, поэтому $t^i_j = t'^{i'}_{j'} a^i_{i'} b^{j'}_j$. $\square$

Замечание. Из определения координат тензора, координаты тензорного произведения образованы произведением соответствующих координат сомножителей. Например, если $t, s \in \mathbb{T}^1_1(V)$, то:

$$(t \otimes s)^{i_1,i_2}_{j_1,j_2} = t^{i_1}_{j_1} s^{i_2}_{j_2}$$

Замечание. Мы определили тензор как полилинейное отображение и получили следующую формулу замены координат:

$$t^{i_1,\dots,i_p}_{j_1,\dots,j_q} = a^{i_1}_{i'_1} \dots a^{i_p}_{i'_p} b^{j'_1}_{j_1} \dots b^{j'_q}_{j_q} t'^{i'_1,\dots,i'_p}_{j'_1,\dots,j'_q}$$

1. Из формулы выше ясен смысл контравариантности верхних и ковариантности нижних индексов: при переходе от e' к e координаты тензора умножаются как на элементы матрицы перехода от e к e', то есть контравариантно, так и на элементы матрицы перехода от e' к e, то есть ковариантно.

2. Тензор можно определить иначе. Можно считать, что тензор типа (p, q) задан в том случае, когда для любого базиса e в V определен набор скаляров $t^{i_1,\dots,i_p}_{j_1,\dots,j_q}$, изменяющийся по указанной выше формуле перехода. Мы доказали, что это определение эквивалентно исходному.

3. Полученная формула перехода позволяет еще раз понять, что $\mathbb{T}^1_1(V) = \mathcal{L}(V)$. Действительно, легко видеть, что если e — базис в V, в котором $\varphi \leftrightarrow_e A = (a_{ij})$, то $A = (\varphi^i_j)$, где φ^i_j — координаты тензора, соответствующего φ. Тогда формула замены координат для φ имеет вид $(\varphi^i_j) = S(\varphi^{i'}_{j'})S^{-1}$, и в ней один множитель контравариантен и один ковариантен.

6.2 Тензорное произведение пространств

Определение 6.5. Пусть U, V — линейные пространства над полем F. *Тензорным произведением пространств* U *и* V называется пространство T над F вместе с билинейным отображением $f : U \times V \to T$ таким, что для любых базисов $e = (e_i)$ в U и $\mathcal{G} = (g_j)$ в V векторы $t_{ij} := f(e_i, g_j)$ образуют базис в T. Обозначение — $U \otimes V$.

Замечание. По определению, если $U \otimes V$ существует, то $\dim T = \dim V \dim U$.

Утверждение 6.2. *Пусть U, V, W — линейные пространства над F, $e = (e_1, \dots, e_k)$ и $\mathcal{G} = (g_1, \dots, g_l)$ — базисы в U и V соответственно, и $f : U \times V \to W$ — билинейное отображение. Тогда эквивалентны следующие условия:*

1. *$t_{ij} := f(e_i, g_j)$ образуют базис в W*

2. *$\forall \overline{w} \in W : \exists!\, \overline{v_1}, \dots, \overline{v_k} \in V : \overline{w} = \sum_{i=1}^k f(e_i, \overline{v_i})$*

3. *$\forall \overline{w} \in W : \exists!\, \overline{u_1}, \dots, \overline{u_l} \in U : \overline{w} = \sum_{j=1}^l f(\overline{u_j}, g_j)$*

Доказательство. Докажем равносильность $1 \Leftrightarrow 2$, поскольку равносильность $1 \Leftrightarrow 3$ доказывается аналогично. Действительно, t_{ij} образуют базис в $W \Leftrightarrow \forall \overline{w} \in W : \exists!\, \alpha_{ij} \in F : \overline{w} = \sum_{i=1}^k \sum_{j=1}^l \alpha_{ij} f(e_i, g_j) = \sum_{i=1}^k f(e_i, \sum_{j=1}^l \alpha_{ij} g_j)$. Но поскольку $\mathcal{G}$ — базис в V, это эквивалентно тому, что $\forall \overline{w} \in W : \exists!\, \overline{v_1}, \dots, \overline{v_k} \in V : \sum_{i=1}^k f(e_i, \overline{v_i})$. $\square$

Замечание. Доказанное выше означает, что свойство (1) не зависит от выбора базисов: если при фиксированном базисе в одном из пространств это свойство выполняется, то базис в другом пространстве можно выбирать произвольно. Следовательно, если свойство (1) выполнено хотя бы для одной пары базисов, то оно выполнено и для всех пар базисов, тогда W и f задают тензорное произведение $U \otimes V$.

Следствие. *Для любых линейных пространств U и V над F существует их тензорное произведение $U \otimes V$.*

Доказательство. В силу утверждения выше, достаточно взять пространство T размерности $\dim U \dim V$, выбрать в нем базис (t_{ij}), и для фиксированных базисов (e_i) в U и (g_j) в V положить $f(e_i, g_j) := t_{ij}$. □

Теорема 6.2. *Пусть U и V — линейные пространства над F, T и f задают $U \otimes V$. Тогда для любого билинейного отображения $b : U \times V \to W$ существует единственное линейное отображение $\varphi_b : T \to W$ такое, что $\varphi_b \circ f = b$. Более того, соответствие между b и φ_b осуществляет изоморфизм между $\mathcal{B}(U, V; W)$ и $\mathcal{L}(T, W)$.*

Доказательство. Пусть $e = (e_i)$ и $\mathcal{G} = (g_j)$ — базисы в U и V соответственно. Тогда на базисе $(t_{ij}) = f(e_i, g_j)$ искомое отображение φ_b однозначно задается как $\varphi_b(t_{ij}) := b(e_i, g_j)$. Сопоставление $\varphi_b \mapsto b$ линейно, и, более того, оно обратимо: действительно, по любому отображению $\varphi_b \in \mathcal{L}(T, W)$ можно восстановить $b \in \mathcal{B}(U, V; W)$, задав его на базисах $e, \mathcal{G}$ как $b(e_i, g_j) := \varphi_b(t_{ij})$. □

Следствие. *Пусть U и V — линейные пространства над F, (T_1, f_1) и (T_2, f_2) — два их тензорных произведения. Тогда существует изоморфизм $\varphi : T_2 \to T_1$ такой, что $\varphi \circ f_2 = f_1$.*

Доказательство. Применим теорему выше, считая (T_2, f_2) тензорным произведением, а f_1 — некоторым билинейным отображением, и получим отображение $\varphi \in \mathcal{L}(T_2, T_1)$ такое, что $\varphi \circ f_2 = f_1$. При этом $\operatorname{Im} \varphi \supset \operatorname{Im} f_1$, поэтому $\operatorname{Im} \varphi$ содержит базис пространства T_1, откуда $\operatorname{Im} \varphi = T_1$. Поскольку φ линеен и $\dim T_1 = \dim T_2$, то он биективен и потому осуществляет изоморфизм между T_1 и T_2. □

Замечание. С точностью до такого изоморфизма тензорное произведение единственно, поэтому можно зафиксировать произвольное тензорные произведение (T, f), опуская при этом отображение f, и для любых $\overline{u} \in U$, $\overline{v} \in V$ обозначать $f(\overline{u}, \overline{v})$ через $\overline{u} \otimes \overline{v} \in T$.

Замечание. $\operatorname{Im} f$ в доказательстве выше не обязан быть подпространством в T, поскольку f — билинейное, а не линейное отображение. Не любой вектор из T представляется в виде $\overline{u} \otimes \overline{v}$. Например, если оба пространства U, V хотя бы двумерные, то $e_1 \otimes g_1 + e_2 \otimes g_2 \notin \operatorname{Im} f$.

Пример. Пусть U, V — линейные пространства над F. Рассмотрим несколько тензорных произведений пространств:

1. $U^* \otimes V^* = \mathcal{B}(U, V; F)$. Для любых $c \in U^*$, $d \in V^*$ положим $f(c, d) := c \otimes d \in \mathcal{B}(U, V; F)$, где $\otimes$ означает тензорное произведение тензоров. Если выбрать пары взаимных базисов e, e^* в U и $\mathcal{G}, \mathcal{G}^*$ в V, то $f(e^i, g^j) = e^i \otimes g^j$ — это билинейная форма, принимающая значение 1 на паре (e_i, e_j) и 0 — на других парах базисных векторов. Значит, f задает тензорное произведение, потому что переводит пару базисов в U^* и V^* в базис в $\mathcal{B}(U, V; F)$.

2. $U \otimes V = \mathcal{B}(U^*, V^*; F)$ в силу пункта (1) и двойственности пространств V и V^*.

3. $U^* \otimes V = \mathcal{L}(U, V)$. Для любых $c \in U^*$, $\overline{v} \in V$ положим $f(c, \overline{v})(\overline{u}) := c(\overline{u})\overline{v} \in \mathcal{L}(U, V)$. Аналогично выберем пары взаимных базисов U и V, тогда матрица каждого оператора вида $f(e^i, g_j)$ будет состоять из одной единицы и остальных нулей, поэтому такие операторы образуют базис в $\mathcal{L}(U, V)$, и f задает тензорное произведение.

Утверждение 6.3. *Пусть U, V, W — линейные пространства над полем F. Тогда выполнены следующие свойства тензорного произведения:*

1. *$\otimes$ ассоциативно, то есть $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$*

2. *$\otimes$ коммутативно, то есть $U \otimes V \cong V \otimes U$*

Доказательство. Имеет место естественные изоморфизмы $\overline{u} \otimes (\overline{v} \otimes \overline{w}) \mapsto (\overline{u} \otimes \overline{v}) \otimes \overline{w}$ и $\overline{u} \otimes \overline{v} \mapsto \overline{v} \otimes \overline{u}$. □

Замечание. Примеры выше с учетом доказанных свойств тензорного произведения позволяют сделать следующий вывод:

$$\mathbb{T}^p_q = \{t : (V^*)^p \times V^q \to F\} = \underbrace{V \otimes \cdots \otimes V}_{p} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q} =: V^{\otimes p} \otimes (V^*)^{\otimes q}$$

Из этого, в частности, следует, что $\mathbb{T}^p_q \otimes \mathbb{T}^{p'}_{q'} \cong \mathbb{T}^{p+p'}_{q+q'}$.

Пример. Пусть U, V — линейные пространства над F. Тогда, по свойствам тензорного произведения, $\mathcal{L}(U, V) = U^* \otimes V \cong V \otimes U^* = \mathcal{L}(V^*, U^*)$. Можно проверить, что этот изоморфизм задает сопоставление $\varphi \mapsto \varphi^*$, где $\varphi \in \mathcal{L}(U, V)$, а $\varphi^* \in \mathcal{L}(V^*, U^*)$ — сопряженное к φ отображение.

6.3 Свертка тензора

В данном разделе зафиксируем линейное пространство V над полем F.

Пример. Рассмотрим тензор $\varphi \in \mathbb{T}^1_1$. Пусть в некотором базисе он задается координатами φ^i_j. При переходе к новому базису с матрицей перехода (a^i_j) и обратной матрицей перехода (b^i_j) координаты тензора меняются по закону $\varphi^i_j = a^i_k \varphi'^k_l b^l_j$. Тогда $\varphi^i_i = a^i_k \varphi'^k_l b^l_i = \delta^l_k \varphi'^k_l = \varphi'^k_k$. Таким образом, мы доказали тензорным способом инвариантность следа φ^i_i относительно замены координат.

Определение 6.6. *Сверткой* тензора $t \in \mathbb{T}^p_q$ по индексам i_p, j_q называется тензор $t' \in \mathbb{T}^{p-1}_{q-1}$ с координатами следующего вида:

$$\widetilde{t}^{i_1,\ldots,i_{p-1}}_{j_1,\ldots,j_{q-1}} = t^{i_1,\ldots,i_{p-1},i}_{j_1,\ldots,j_{q-1},i}$$

Свертка по другим парам из верхнего и нижнего индексов определяется аналогично.

Замечание. Аналогично проверке для следа оператора, можно показать, что полученный объект действительно является тензором, поскольку величина, получаемая при суммировании по i и фиксации всех остальных индексов, не зависит от выбора базиса.

Замечание. Зафиксируем базис e в V. Тогда свертка по индексам i_p, j_q действует на базисных векторах пространства $\mathbb{T}^p_q$ следующим образом:

$$e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q} \mapsto e^{j_q}(e_{i_p}) e_{i_1} \otimes \ldots \otimes e_{i_{p-1}} \otimes e^{j_1} \otimes \cdots \otimes e^{j_{q-1}}$$

Значит, по полилинейности, для произвольных $\overline{v_1}, \ldots, \overline{v_p} \in V$, $u^1, \ldots, u^q \in V^*$:

$$\overline{v_1} \otimes \ldots \otimes \overline{v_p} \otimes u^1 \otimes \cdots \otimes u^q \mapsto u^q(\overline{v_p}) \overline{v_1} \otimes \ldots \otimes \overline{v_{p-1}} \otimes u^1 \otimes \cdots \otimes u^{q-1}$$

Пример. Рассмотрим несколько примеров свертки:

1. Пусть $\overline{v} \in V$, $u \in V^*$. Тогда свертка тензора $u \otimes \overline{v}$ — это скаляр $u(\overline{v})$.

2. Пусть $b \in \mathcal{B}(V)$ — тензор с координатами b_{ij}, $\overline{u}, \overline{v} \in V$. Тогда скаляр $b(\overline{u}, \overline{v}) = u^i b_{ij} v^j$ получается как *двойная*, или *полная*, свертка тензора $\overline{u} \otimes b \otimes \overline{v}$.

3. Пусть $\varphi \in \mathcal{L}(V)$ — тензор с координатами φ^i_j, $\overline{v} \in V$. Тогда вектор $\varphi(\overline{v})$ имеет координаты $\varphi^i_j v^j$.

4. Пусть $\varphi, \psi \in \mathcal{L}(V)$ — тензоры с координатами φ^i_j, ψ^k_l. Тогда тензор $\varphi \circ \psi$ имеет координаты $\varphi^i_j \psi^j_k$.

5. Пусть V — евклидово пространство, в нем введено скалярное произведение, или *метрический тензор*, с координатами g_{ij}. Тогда канонический изоморфизм между V и V^* осуществляется сопоставлением $v^i \mapsto v^i g_{ij}$, называемым *опусканием индекса*. На V^* тоже можно задать скалярное произведение как тензор с координатами g^{ij}, тоже называемый метрическим тензором, позволяющий, наоборот, поднимать индексы. Можно также показать, что $g_{ij} g^{ik} = \delta^k_j$.

6. Пусть $\varphi \in \mathcal{L}(V)$ — тензор с координатами φ^i_j. Если в пространстве V задано скалярное произведение с координатами g_{ij}, то сопоставление $\varphi^i_j \mapsto \varphi^i_j g_{ik}$ осуществляет это изоморфизм между $\mathcal{L}(V)$ и $\mathcal{B}(V)$.

6.4 Тензорная алгебра

Определение 6.7. Пусть $V_1, \ldots, V_n$ — линейные пространства над полем F. Тогда их *(внешней) прямой суммой* называется пространство $V_1 \times \cdots \times V_n$ с операциями сложения и умножения на скаляр, введенными покомпонентно. Обозначение — $V_1 \oplus \cdots \oplus V_n$.

Замечание. Для любого $i \in \{1, \ldots, n\}$ в $V_1 \oplus \cdots \oplus V_n$ есть подпространство, канонически изоморфное V_i, вида $\{\overline{0}\} \times \cdots \times \{\overline{0}\} \times V_i \times \{\overline{0}\} \times \cdots \times \{\overline{0}\}$. Кроме того, каждый вектор $\overline{v} \in V_1 \oplus \cdots \oplus V_n$ единственным образом раскладывается в сумму векторов из таких подпространств, поэтому $V_1 \oplus \cdots \oplus V_n$ можно считать внутренней суммой таких подпространств, каждое из которых можно отождествить с соответствующим V_i.

Следствие. *Все свойства прямой суммы переносятся на внешнюю прямую сумму, в частности, базис в $V_1 \oplus \cdots \oplus V_n$ — это объединение базисов в $V_1, \ldots, V_n$.*

Замечание. В случае бесконечной прямой суммы $V := \bigoplus_{i=1}^{\infty} V_i$ нужно дополнительно требовать, чтобы в каждом наборе $(\overline{v_1}, \overline{v_2}, \ldots) \in V$ было лишь конечное число ненулевых векторов, чтобы сохранить связь с внутренней прямой суммой. Если разрешить бесконечное количество ненулевых векторов в наборах, то объединение базисов в $V_1, V_2, \dots$ уже не будет порождать V. Такая конструкция отличается от прямой суммы и называется *прямым произведением.*

До конца раздела зафиксируем линейное пространство V над полем F.

Определение 6.8. $\mathbb{T} := \bigoplus_{p=0}^{\infty} V^{\oplus p}$ называется *тензорной алгеброй* пространства V.

Замечание. В определении выше и везде далее считается, что $V^{\oplus 0} = F$.

Пример. Элемент алгебры $\mathbb{T}$ может, например, иметь вид $\alpha + \overline{v} + \overline{u_1} \otimes \overline{u_2} \otimes \overline{u_3}$, где $\alpha \in F$, $\overline{v}, \overline{u_1}, \overline{u_2}, \overline{u_3} \in V$.

Замечание. Умножение в $\mathbb{T}$ задается как тензорное произведение тензоров на базисных тензорах и продолжается на все пространство $\mathbb{T}$ по билинейности.

Замечание. Если $e = (e_1, \dots, e_n)$ — базис в V, то базис в $\mathbb{T}$ может быть получен как объединение систем (1), $(e_1, \dots, e_n)$, $(e_1 \otimes e_1, e_1 \otimes e_2, \dots, e_n \otimes e_n)$, и так далее.

Утверждение 6.4. *Умножение в $\mathbb{T}$ ассоциативно, но необязательно коммутативно.*

Доказательство. Данные свойства следуют из соответствующих свойств тензорного произведения. □

Определение 6.9. Пусть $t \in \mathbb{T}^p_q$. Тензор t называется *симметричным по первым двум координатам*, если для любых функционалов $f_1, \dots, f_p \in V^*$ и векторов $\overline{v_1}, \dots, \overline{v_q} \in V$ выполнено $t(f_1, f_2, \dots, f_p, \overline{v_1}, \dots, \overline{v_q}) = t(f_2, f_1, \dots, f_p, \overline{v_1}, \dots, \overline{v_q})$.

Замечание. Легко видеть, что t симметричен по первым двум верхним индексам $\Leftrightarrow$ его координаты симметричны по первым двум верхним индексам. Симметричность по другим наборам координат одного типа определяется аналогично.

Определение 6.10. Пусть $t \in \mathbb{T}^p_0$, $\sigma \in S_p$. Будем обозначать через $g_\sigma(t)$ такой тензор $g \in \mathbb{T}^p_0$, что $\forall f_1, \dots, f_p \in V^* : g(f_1, \dots, f_p) = t(f_{\sigma(1)}, \dots, f_{\sigma(p)})$.

Замечание. Пусть e — базис в V. Если t имеет в базисе e координаты $t^{i_1, \dots, i_p}$, то $g_\sigma(t)$ в этом же базисе имеет координаты $t^{i_{\sigma(1)}, \dots, i_{\sigma(p)}}$.

Определение 6.11. Тензор $t \in \mathbb{T}^p_0$ называется *симметричным*, если $\forall \sigma \in S_p : g_\sigma(t) = t$. Такие тензоры образуют подпространство в $\mathbb{T}^p_0$, обозначаемое через $\mathbb{ST}^p$.

Замечание. Равенство $g_\sigma(t) = t$ достаточно проверять только для набора перестановок $\sigma \in S_p$, порождающего S_p, например, для всех транспозиций соседних элементов. Иными словами, тензор $t \in \mathbb{T}^p$ симметричен $\Leftrightarrow$ t симметричен по любой паре соседних индексов.

Определение 6.12. *Симметризацией* тензора $t \in \mathbb{T}^p_0$ называется следующий тензор:

$$s(t) := \frac{1}{p!} \sum_{\sigma \in S_p} g_\sigma(t) \in \mathbb{T}^p_0$$

Симметризация определена, если $\operatorname{char} F \nmid p$.

Утверждение 6.5. *Симметризация обладает следующими свойствами:*

1. *Для любого тензора* $t \in \mathbb{T}_0^p$ *выполнено* $s(t) \in \mathbb{ST}^p$.
2. *Если* $t \in \mathbb{ST}^p$, *то* $s(t) = t$.
3. $\operatorname{Im} s = \mathbb{ST}^p$.

Доказательство.

1. Пусть $\tau \in S_p$. Тогда:

$$g_\tau(s(t)) = g_\tau\left(\frac{1}{p!}\sum_{\sigma\in S_p} g_\sigma(t)\right) = \frac{1}{p!}\sum_{\sigma\in S_p} g_{\tau\sigma}(t) = \frac{1}{p!}\sum_{\widetilde{\tau}\in S_p} g_{\widetilde{\tau}}(t) = s(t)$$

2. Если $t \in \mathbb{ST}^p$, то $\forall \sigma \in S_p : g_\sigma(t) = t$, из чего и следует требуемое.
3. Равенство $\operatorname{Im} s = \mathbb{ST}^p$ выполнено в силу пункта (2). □

Замечание. Конечно, *частичная симметризация* возможна и для произвольных тензоров типа (p, q), в этом случае суммирование производится по всевозможным перестановкам того набора индексов, по которому производится симметризация.

Утверждение 6.6. *Для произвольных тензоров* $t_1 \in \mathbb{T}^{p_1}(V)$, $t_2 \in \mathbb{T}^{p_2}(V)$ *выполнены равенства* $s(t_1 \otimes t_2) = s(s(t_1) \otimes t_2) = s(t_1 \otimes s(t_2))$.

Доказательство. Положим $p := p_1 + p_2$. Тогда:

$$s(s(t_1)\otimes t_2) = \frac{1}{p!}\sum_{\sigma\in S_p} g_\sigma(s(t_1)\otimes t_2) = \frac{1}{p!}\sum_{\sigma\in S_p} g_\sigma\left(\left(\frac{1}{p_1!}\sum_{\tau\in S_{p_1}} g_\tau(t_1)\right)\otimes t_2\right)$$

Теперь для каждой перестановки $\tau \in S_{p_1}$ определим перестановку $\widetilde{\tau} \in S_p$ следующим образом: $\widetilde{\tau}|_{\{1,\ldots,p_1\}} = \tau$, $\widetilde{\tau}|_{\{p_1+1,\ldots,p_1+p_2\}} = \mathrm{id}$. Тогда:

$$s(s(t_1)\otimes t_2) = \frac{1}{p_1!}\sum_{\tau\in S_{p_1}} \frac{1}{p!}\sum_{\sigma\in S_p} g_{\sigma\widetilde{\tau}}(t_1\otimes t_2) = \frac{1}{p_1!}\sum_{\tau\in S_{p_1}} s(t_1\otimes t_2) = s(t_1\otimes t_2)$$ □

Равенство $s(t_1 \otimes s(t_2)) = s(t_1 \otimes t_2)$ доказывается аналогично.

Определение 6.13. Для произвольных тензоров $t_1 \in \mathbb{ST}^{p_1}$, $t_2 \in \mathbb{ST}^{p_2}$ будем обозначать через $t_1 \vee t_2$ тензор $s(t_1 \otimes t_2) \in \mathbb{ST}^{p_1+p_2}$.

Утверждение 6.7. *Пусть* $t_1 \in \mathbb{ST}^{p_1}, t_2 \in \mathbb{ST}^{p_2}, t_3 \in \mathbb{ST}^{p_3}$. *Тогда выполнены следующие равенства:*

1. $(t_1 \vee t_2) \vee t_3 = t_1 \vee (t_2 \vee t_3)$
2. $t_1 \vee t_2 = t_2 \vee t_1$

Доказательство.

1. $(t_1 \vee t_2) \vee t_3 = s(s(t_1 \otimes t_2) \otimes t_3) = s(t_1 \otimes t_2 \otimes t_3) = s(t_1 \otimes s(t_2 \otimes t_3)) = t_1 \vee (t_2 \vee t_3)$.

2. Положим $p := p_1 + p_2$. Заметим, что $\exists \tau \in S_p : t_1 \otimes t_2 = g_\tau(t_2 \otimes t_1)$, тогда:

$$s(t_1 \otimes t_2) = \frac{1}{p!} \sum_{\sigma \in S_p} g_\sigma(t_1 \otimes t_2) = \frac{1}{p!} \sum_{\sigma \in S_p} g_{\sigma\tau}(t_2 \otimes t_1) = s(t_2 \otimes t_1) \qquad \square$$

Замечание. Пусть $e = (e_1, \ldots, e_n)$ — базис в V. Тогда $\mathbb{ST}^p$ порождается тензорами вида $e_1^{\vee\alpha_1} \vee \cdots \vee e_n^{\vee\alpha_n}$, где $\alpha_1 + \cdots + \alpha_n = p$. Легко видеть, что эти тензоры линейно независимы, поэтому $\dim \mathbb{ST}^p = \overline{C_n^p}$

Определение 6.14. Алгебра $\mathbb{S} := \bigoplus_{p=0}^{\infty} \mathbb{ST}^p$ называется *симметрической алгеброй* пространства V.

Замечание. В отличие от тензорной алгебры, симметрическая алгебра коммутативна. Более того, нетрудно показать, что $\mathbb{S} \cong F[x_1, \ldots, x_k]$, где $k := \dim V$.

Определение 6.15. Тензор $t \in \mathbb{T}_0^p$ называется *кососимметричным*, если $\forall \sigma \in S_p : g_\sigma(t) = \operatorname{sgn} \sigma \cdot t$. Такие тензоры образуют подпространство в $\mathbb{T}_0^p$, обозначаемое через Λ^p.

Замечание. Как и в симметричном случае, равенство $g_\sigma(t) = \operatorname{sgn} \sigma \cdot t$ достаточно проверять только для набора перестановок $\sigma \in S_p$, порождающего S_p, например, для всех транспозиций соседних элементов.

Определение 6.16. *Альтернированием* тензора $t \in \mathbb{T}_0^p$ называется следующий тензор:

$$a(t) := \frac{1}{p!} \sum_{\sigma \in S_p} \operatorname{sgn} \sigma \cdot g_\sigma(t) \in \mathbb{T}_0^p$$

Альтернирование определено, если $\operatorname{char} F \nmid p$.

Замечание. Как и в симметричном случае, *частичное альтернирование* возможно и для произвольных тензоров типа (p, q).

Утверждение 6.8. *Альтернирование обладает следующими свойствами:*

1. *Для любого тензора $t \in \mathbb{T}_0^p$ выполнено $a(t) \in \Lambda^p$.*
2. *Если $t \in \Lambda^p$, то $a(t) = t$.*
3. $\operatorname{Im} a = \Lambda^p$.

Доказательство. Доказательство аналогично симметричному случаю. $\square$

Утверждение 6.9. *Для произвольных тензоров $t_1 \in \mathbb{T}^{p_1}(V)$, $t_2 \in \mathbb{T}^{p_2}(V)$ выполнены равенства $a(t_1 \otimes t_2) = a(a(t_1) \otimes t_2) = a(t_1 \otimes a(t_2))$.*

Доказательство. Доказательство аналогично симметричному случаю. $\square$

Определение 6.17. Для произвольных тензоров $t_1 \in \mathbb{T}_0^{p_1}(V)$, $t_2 \in \mathbb{T}_0^{p_2}(V)$ будем обозначать через $t_1 \wedge t_2$ тензор $a(t_1 \wedge t_2) \in \Lambda^{p_1+p_2}$

Утверждение 6.10. *Пусть* $t_1 \in \Lambda^{p_1}, t_2 \in \Lambda^{p_2}, t_3 \in \Lambda^{p_3}$. *Тогда выполнены следующие равенства:*

1. $(t_1 \wedge t_2) \wedge t_3 = t_1 \wedge (t_2 \wedge t_3)$
2. $t_1 \wedge t_2 = t_2 \wedge t_1$

Доказательство. Доказательство аналогично симметричному случаю. □

Определение 6.18. Алгебра $\Lambda := \bigoplus_{p=0}^{\infty} \Lambda^p$ называется *внешней алгеброй*, или *алгеброй Грассмана*, пространства V.

Замечание. Аналогично симметричному случаю, можно показать, что базис в Λ^p — это тензоры вида $e_{i_1} \wedge \dots \wedge e_{i_p}$, $i_1 < \dots < i_p$. Это, в частности, означает, что $\Lambda^p = \{0\}$ при $p > k$ и $\dim \Lambda(V) = 2^k$, где $k := \dim V$.